

AJAY
AGRAWAL

JOSHUA
GANS

AVI
GOLDFARB

LÊ DUNG dịch

AI 5.0

NHANH HƠN, DỄ HƠN,
RẺ HƠN, CHÍNH XÁC HƠN

1. Học tập và nghiên cứu dễ dàng hơn
2. Lập kế hoạch và giải quyết vấn đề nhanh hơn
3. Chi phí đầu tư và sản xuất rẻ hơn
4. Đánh giá và đưa ra dự báo chính xác hơn



NHA XUẤT BẢN
CÔNG THƯƠNG



AI
5.0



POWER AND PREDICTION

The Disruptive Economics of Artificial Intelligence

Copyright © 2022 by Ajay Agrawal, Joshua Gans, and Avi Goldfarb.

All rights reserved.

AI 5.0 – NHANH HƠN, DỄ HƠN, RẺ HƠN, CHÍNH XÁC HƠN

Bản quyền tiếng Việt © 2023,

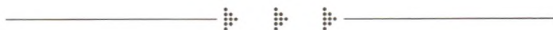
Công ty TNHH Văn hóa và Truyền thông 1980 Books.



Không phần nào trong cuốn sách này được sao chép hoặc chuyển sang bất cứ dạng thức hoặc phương tiện nào, dù là điện tử, in ấn, ghi âm hay bất cứ hệ thống phục hồi và lưu trữ thông tin nào nếu không có sự cho phép bằng văn bản của Công ty TNHH Văn hóa và Truyền thông 1980 Books.



Mọi ấn phẩm của 1980 Books đều được in trên chất liệu giấy cao cấp và an toàn.



Liên hệ về các vấn đề chung và bản dịch, bản thảo: rights@1980books.com

Liên hệ dịch vụ tư vấn và đại diện giao dịch bản quyền: rights@1980books.com

Liên hệ về phát hành, kinh doanh: 1980books@gmail.com

AJAY
AGRAWAL

JOSHUA
GANS

AVI
GOLDFARB

LÊ DUNG dịch

AI 5.0

NHANH HƠN, DỄ HƠN,
RẺ HƠN, CHÍNH XÁC HƠN

1. Học tập và nghiên cứu dễ dàng hơn
2. Lập kế hoạch và giải quyết vấn đề nhanh hơn
3. Chi phí đầu tư và sản xuất rẻ hơn
4. Đánh giá và đưa ra dự báo chính xác hơn

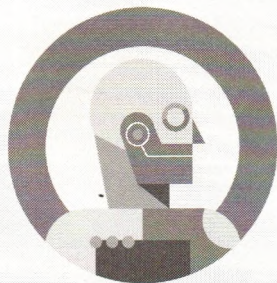


Harvard
Business
Review
Press



NHA XUẤT BẢN
CÔNG THƯƠNG





MỤC LỤC

Phần 1

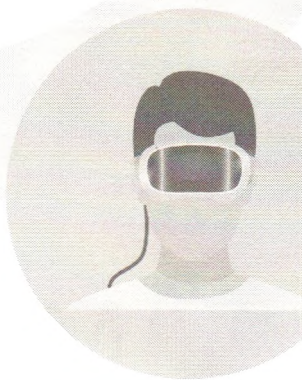
THE BETWEEN TIMES - THỜI GIAN CHUYỂN ĐỔI	23
Chương 1 Ẩn dụ 3 doanh nhân	24
Chương 2 Tương lai của AI	40
Chương 3 AI là Công nghệ dự báo	62

Phần 2

CÁC QUY TẮC	87
Chương 4 Quyết định hay không quyết định	88
Chương 5 Sự không chắc chắn giấu kín	110
Chương 6 Quy tắc là keo dính	127

Phần 3

HỆ THỐNG	145
Chương 7 Dính chặt và bôi trơn	146
Chương 8 Tư duy hệ thống	163
Chương 9 Hệ thống vĩ đại nhất	185



Phần 4

QUYỀN LỰC	199
Chương 10 Sự gián đoạn và quyền lực	200
Chương 11 Máy móc có quyền lực không?	221
Chương 12 Tích lũy quyền lực	241

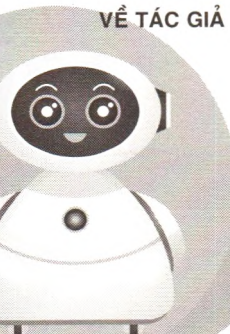
Phần 5

AI GÂY GIÁN ĐOẠN NHƯ THẾ NÀO	263
Chương 13 Một sự tách rời tuyệt vời	264
Chương 14 Tư duy xác suất	282
Chương 15 Những “thẩm phán” mới	305

Phần 6

HÌNH DUNG CÁC HỆ THỐNG MỚI	331
Chương 16 Thiết kế các hệ thống đáng tin cậy	332
Chương 17 Tờ giấy trắng	357
Chương 18 Dự đoán thay đổi hệ thống	383

PHẦN KẾT XU HƯỚNG VÀ HỆ THỐNG AI	407
VỀ TÁC GIẢ	429



LỜI KHEN TẶNG

Đây là cuốn sách dành cho lãnh đạo của mọi doanh nghiệp. Cuốn sách cho thấy tiềm năng khổng lồ của AI và những thách thức đặt ra để đạt đến quy mô đó. Từ ngân hàng đến sản xuất, từ thời trang đến khoáng sản, hệ thống AI sẽ trở nên phổ biến rộng khắp với tầm ảnh hưởng như điện và Internet.

- DOMINIC BARTON, Chủ tịch, Rio Tinto;
cựu đối tác của Global Managing, McKinsey & Company

Tầm ảnh hưởng của AI ở thế kỷ 21 cũng như điện ở thế kỷ 20. Bất kỳ ai cũng nghĩ rằng tương lai của nền kinh tế sẽ xoay quanh và phụ thuộc vào sự phát triển và các ứng dụng của AI. Đây là cuốn sách hay nhất thảo luận về ý nghĩa của điều này với mọi thành phần của nền kinh tế thế giới.

- LAWRENCE H. SUMMERS, Charles W. Eliot,
cựu chủ tịch và giáo sư của Harvard University;
cựu Chủ tịch US Treasury; và cựu Chuyên gia
kinh tế trưởng của World Bank



AI chắc chắn sẽ thay thế nhiều công việc và làm gián đoạn nền kinh tế trong những thập kỷ tới, với sự tiên phong và thúc đẩy của nhiều doanh nhân trẻ. Những thay đổi ở cấp độ hệ thống đã bắt được manh nha cũng chính là chủ đề thảo luận của cuốn sách này, và những thay đổi này sẽ đặt nền móng cho cuộc cách mạng sắp tới.

- VINOD KHOSLA, nhà sáng lập Khosla Ventures;
nhà đồng sáng lập Sun Microsystems

Bạn sẽ cần quyết tâm để đào sâu và sẵn sàng đầu tư thời gian mới có thể gặt hái được thành quả ẩn giấu trong những trang sách này. Thế nhưng phần thưởng sẽ rất xứng đáng. Cuốn sách là sự mở đầu vô cùng gợi mở tư duy và truyền cảm hứng mạnh mẽ về cách định hình chiến lược và cơ cấu tổ chức trong thời đại AI.

- HEATHER REISMAN, nhà sáng lập kiêm CEO của
Indigo Books and Music

Cuốn sách này đúng là “cực phẩm”. Chúng ta thường được nghe rằng AI sẽ trở thành điều vĩ đại nhất mà con người đã tạo nên, thế nhưng có vẻ như ảnh hưởng của nó với thế giới trong hiện tại vẫn rất mơ hồ và nhỏ hẹp. Tác giả cho chúng ta

thấy hai khía cạnh này không hề mâu thuẫn. Với nhiều ví dụ độc đáo và phong phú, cuốn sách là một mô tả hoàn hảo về thời điểm đặc biệt này trong lịch sử của AI. Đây là cuốn sách mà bất kỳ ai muốn tìm hiểu về tương lai của AI cần phải đọc.

- SHIVON ZILIS, Giám đốc điều hành và Chuyên gia dự án của Neuralink; thành viên Hội đồng Quản trị của OpenAI; cựu giám đốc điều hành của Tesla

Không cuốn sách nào có thể giải thích rõ hơn về những yếu tố kinh tế nền tảng của AI và những gì mà AI thực sự có thể làm được. Đó không chỉ là việc sử dụng trường hợp làm cơ sở dự đoán – đó là những hệ thống ra quyết định tốt hơn. Đó là bước tiến dài hơn cho cả doanh nghiệp và nền kinh tế.

- TIFF MACKLEM,
Thống đốc Ngân hàng Trung ương Canada

Agrawal, Gans, và Goldfarb đã làm được, một lần nữa! Cuốn sách mới của họ chắc chắn sẽ trở thành hướng dẫn rõ ràng cho bạn đọc vì sao và bằng cách nào AI có thể biến đổi nền kinh tế.

- ERIK BRYNJOLFSSON, Giám đốc của Stanford Digital Economy Lab; đồng tác giả cuốn sách *The Second Machine Age* và *Machine Platform Crowd*



Dù chúng ta có muốn hay không, trí tuệ nhân tạo được xác định sẽ ảnh hưởng đến mọi mặt của cuộc sống. Làm sao chúng ta có thể đảm bảo mọi cá nhân, doanh nghiệp và tổ chức sẽ được lợi thay vì lãng phí thời gian và tài nguyên đối phó với những hậu quả không mong muốn? Cuốn sách cho chúng ta thấy một khởi đầu xuất sắc, nhấn mạnh đến cách AI có thể cải thiện công việc chúng ta làm nhờ những dự đoán tốt hơn và hỗ trợ tái cơ cấu tổ chức.

- DARON ACEMOGLU, Giáo sư chuyên ngành Kinh tế học của MIT; tác giả cuốn sách *When Nations Fail*

LỜI MỞ ĐẦU

Thành công từ xa?

Sau khi phát hành cuốn sách *Prediction Machines* (AI trong cuộc cách mạng công nghệ 4.0) vào năm 2018, chúng tôi nhớ rằng mình đã nói rằng đó là tất cả những gì chúng ta cần biết về kinh tế học AI. Thế nhưng, chúng tôi đã nhầm.

Mặc dù hoàn toàn thừa nhận rằng công nghệ sẽ liên tục biến đổi – và AI vẫn trong giai đoạn thai nghén, chúng ta cũng phải khẳng định rằng nền tảng kinh tế học sẽ không thay đổi. Đó là vẻ đẹp của kinh tế học. Công nghệ thay đổi, nhưng kinh tế học thì không. Chúng tôi đã thiết lập khung cấu trúc cho kinh tế học AI trong cuốn sách đó và tin rằng nó vẫn hữu ích cho đến hôm nay. Tuy nhiên, khung cấu trúc đó mới chỉ phản ánh một phần câu chuyện, phần giải pháp. Trong những năm sau đó, chúng tôi phát hiện ra rằng một phần quan trọng của câu chuyện AI vẫn chưa được nói đến, phần hệ thống. Chúng tôi sẽ nói đến câu chuyện ở đây.



Làm sao chúng ta đã bỏ lỡ phần này ngay từ đầu. Chúng tôi tua ngược trở lại năm 2017, khi cuốn sách *AI trong cuộc cách mạng công nghệ 4.0* ra đời, để tìm lời giải đáp.

Năm đó, nửa thập kỷ sau các doanh nghiệp ở Canada tiên phong ứng dụng AI vào xử lý và phân loại hình ảnh và đạt được kết quả vượt trội, sự quan tâm và hứng thú với công nghệ này bùng nổ. Người người nói về AI, và người ta dự đoán rằng nhờ vậy Canada có thể bước vào sàn đấu công nghệ của thế giới. Vấn đề không phải là tính khả thi của dự đoán này mà là khi nào điều đó sẽ xảy ra.

Chúng tôi thành lập một chương trình khởi nghiệp hướng đến khoa học với tên gọi Creative Destruction Lab với một bộ phận dành riêng cho AI. Bản khoản của mọi người là “Liệu kỳ lân AI (công ty khởi nghiệp AI đầu tiên được định giá 1 tỉ đô la) đầu tiên của Canada sẽ ở đâu?” Chúng tôi cá là: “Montreal. Hoặc có thể là Toronto, cũng có thể là Edomonton”.

Và không chỉ riêng chúng tôi nghĩ vậy. Đây cũng là ván cược của chính phủ Canada. Vào ngày 26 tháng 10 năm 2017, chúng tôi tiếp đón ông Justin Trudeau, thủ tướng Canada, vào hội nghị thường niên về AI tại Creative Destruction Lab với chủ đề

Máy học và Thị trường trí tuệ (Machine Learning and the Market for Intelligence). Trong bài phát biểu của mình, Thủ tướng nhấn mạnh tầm quan trọng của việc đầu tư theo từng vùng địa lý với đa dạng các đối tượng từ tập đoàn lớn, các công ty khởi nghiệp đến các trường đại học, nhà đầu tư và bao gồm cả đầu tư vào tài năng cá nhân, từ đó thúc đẩy sáng tạo và tạo thêm nhiều việc làm. Ý tưởng trọng tâm là chú trọng vào phân bổ đồng thời. Vài tháng sau đó, chính phủ thông báo đầu tư đáng kể vào năm vùng trọng điểm, trong đó có một vùng tập trung vào AI ở Montreal.

Chúng tôi tự tin với niềm tin của mình về tương lai thương mại hóa AI. Chúng tôi chẳng phải được xem là chuyên gia thế giới ở chủ đề này sao. Chúng tôi là tác giả của cuốn sách bán chạy nhất về nền tảng kinh tế học của AI, chúng tôi đã xuất bản cơ sở các bài nghiên cứu và luận văn về chủ đề này; chúng tôi cũng đồng biên tập tài liệu tham khảo hàng đầu cho nghiên cứu sinh ở lĩnh vực này – *The Economics of Artificial Intelligence: An Agenda* (tạm dịch: Kinh tế học Trí tuệ nhân tạo: Khung chương trình). Hơn nữa, chúng tôi còn xây dựng một chương trình về thương mại hóa AI tập hợp sự quan tâm lớn nhất từ các công ty AI trên thế giới; chúng tôi thực hiện các bài thuyết trình



trên khắp thế giới cho doanh nghiệp và lãnh đạo chính phủ; và chúng tôi tham gia hàng loạt các ủy ban chính sách, các lực lượng chuyên môn hay hội nghị bàn tròn liên quan đến AI.

Chúng tôi cho rằng, sự phát triển và bùng nổ của AI nên được nhìn nhận dưới hình thức dự đoán và quan điểm này của chúng tôi gây được tiếng vang với giới thực hành. Chúng tôi được mời đến thuyết trình và trao đổi tại Google, Netflix, Amazon, Facebook, và Microsoft. Trong một cuộc phỏng vấn, Gustav Söderström, giám đốc sản phẩm, kỹ thuật, dữ liệu và thiết kế tại Spotify – một trong những nhà cung cấp dịch vụ phát nhạc trực tuyến lớn nhất thế giới đều nhắc đến cuốn sách của chúng tôi:

“Các tác giả đã trình bày quan điểm này một cách hoàn hảo trong cuốn sách của họ, *AI trong cuộc cách mạng công nghệ 4.0*. Hãy tưởng tượng sự chính xác trong dự báo của hệ thống máy học giống như nút điều chỉnh âm lượng của radio... Khi bạn vặn nút đến một mức nhất định (giống như khi các dự báo của bạn đủ chính xác), điều gì đó sẽ xảy ra. Bạn vượt qua ngưỡng giới hạn, khi đó bạn phải cân nhắc lại toàn bộ mô hình kinh doanh và sản phẩm dựa trên máy học của mình... Với Discover Weekly, chúng

tôi hoán đổi thuật toán, từ “shopping then shipping”, sang “shipping then shopping”, tương tự như cách mà cuốn sách *AI trong cuộc cách mạng công nghệ 4.0* mô tả. Chúng tôi cũng đạt đến mức độ chính xác dự báo nhất định, khi đó chúng tôi có thể chuyển đổi, từ cung cấp cho người dùng các công cụ tốt hơn để tự tạo danh sách bài hát, đến chỉ cung cấp cho họ một danh sách hằng tháng và để họ lưu bài nhạc yêu thích. Chúng tôi chuyển đổi tầm nhìn, từ “cung cấp công cụ tốt hơn để tự tạo danh sách bài hát” thành “bạn không bao giờ phải tạo lại danh sách”.

Hướng tiếp cận của chúng tôi là thiết kế cho một thế giới mà các dự báo chất lượng không còn đắt đỏ. Điều này vô cùng quan trọng và giúp định hướng nền tảng cho chiến lược AI của chúng tôi.

Vì vậy, vì sao chúng tôi lại tự tin như vậy, rằng kỳ lân AI đầu tiên sẽ đến từ Montreal, Toronto hoặc Edmonton? Chúng tôi liên hệ với hai kỹ sư nhận giải thưởng Turning Award gần đây (tương đương với giải thưởng Nobel cho khoa học máy tính). Chúng tôi đánh giá cao sự tiên phong của họ trong lĩnh vực học sâu (deep learning). Hai người này sống ở Montreal và Toronto. Chúng tôi cũng liên hệ với một kỹ sư khác ở Edmonto,



là một trong những người tiên phong trong lĩnh vực học tăng cường (reinforcement learning). Chính phủ Canada sẽ tăng cường đầu tư cho 3 tổ chức mới chuyên vào nghiên cứu chuyên sâu ở lĩnh vực máy học – tại Montreal, Toronto, và Edmonton. Nhiều tập đoàn lớn nhanh chóng thành lập các phòng thí nghiệm AI ở Montreal (chẳng hạn Ericsson, Facebook, Microsoft, Huawei, Samsung), Toronto (như Nvidia, LG Electronics, Johnson & Johnson, Roche, Thomson Reuters, Uber, Adobe), và Edmonton (như Google/DeepMind, Amazon, Mitsubishi, IBM).

Vì vậy, không có gì không thỏa đáng nếu chúng tôi nói rằng chúng tôi hiểu rất rõ về thương mại hóa AI. Thế nhưng, dự đoán của chúng tôi đã sai – rất sai. Kỳ lân AI đầu tiên của Canada không đến từ Montreal, Toronto, hoặc Edmonton. Thậm chí, kịch bản thứ hai của chúng tôi – Vancouver, Calgary, Waterloo, hoặc Halifax cũng sai bét. Kỳ lân đầu tiên này không xuất hiện ở những thành phố được xem là trung tâm công nghệ của Canada này – vậy nó ở đâu?

Vào tháng 19 tháng 11 năm 2020, *Wall Street Journal* chạy tít “Nasdaq to Buy Anti-Financial Crime Firm Verafin for \$2.75 Billion” (tạm dịch: Nasdaq mua Verafin – công ty chống tội phạm tài chính với giá 2,75 tỉ đô la).

Verafin có trụ sở tại St. John's, Newfoundland. Không có nhiều người, kể cả chúng tôi, có thể dự báo rằng kỳ lân AI đầu tiên của Canada đến từ một thị trấn nhỏ nằm xa ở cực đông bắc của Bắc Mỹ.

St. Johns, Newfoundland là tỉnh cực đông của Canada với dân số chỉ hơn một nửa triệu người. Dĩ nhiên, thị trấn này không rơi vào tầm ngắm của radar công nghệ. Trên thực tế, dù Mỹ là nước láng giềng của Canada, nhiều người Mỹ cũng chỉ lần đầu nghe đến tên thị trấn này khi bản nhạc kịch Broadway, *Come from Away* được đề cử Best Musical và 4 hạng mục khác trong khuôn khổ Tony Awards 2017. Vở nhạc kịch dựa trên câu chuyện có thật về những gì xảy ra trong tuần sau thảm họa 11/9, khi 38 chiếc máy bay theo lệnh phải hạ cánh xuống Newfoundland và người dân tốt bụng, hài hước nơi đây đã tiếp đón 7.000 khách du lịch bị mắc kẹt từ xa đến (from Away). Thế nhưng, chính tại thị trấn này, tại Newfoundland, Brendan Brothers, Jamie King, và Raymond Pretty đã thành lập Verafin – công ty cung cấp phần mềm phát hiện gian lận cho 3.000 tổ chức tài chính tại Bắc Mỹ. Làm thế nào mà chúng ta có thể bỏ qua điều này chứ? Đó có phải hoàn toàn ngẫu nhiên không? Thậm chí các chuyên gia đôi khi cũng mắc sai lầm. Nhận thức muộn màng. Rõ ràng luôn có khả năng xảy ra những việc như thế.



Cái mà NASDAQ mua là AI. Verafin đã đầu tư 1. ạnh và xây dựng các công cụ dự đoán gian lận và xác thực danh tính khách hàng. Đây là 5 chức năng chính của các tổ chức tài chính, xét về cả vận hành và tuân thủ quy định. Để làm được như vậy yêu cầu lượng dữ liệu lớn, và dữ liệu ngân hàng và hợp nhất tín dụng chiếm phần lớn.

Như vậy, không phải ngẫu nhiên mà một doanh nghiệp như Verafin có thể dẫn đầu cuộc đua. Đó là hiển nhiên. Chúng tôi đã quá tập trung vào những khả năng trong dự đoán của hệ thống nên bị “che mắt” trước xác suất triển khai thương mại thực tế. Trong khi tập trung vào các đặc tính kinh tế của bản thân AI (tức là giảm chi phí dự đoán), chúng tôi đã đánh giá thấp yếu tố kinh tế học của việc xây dựng những hệ thống mới tích hợp AI.

Nếu khi đó đã hiểu được như vậy, thay vì đánh giá bối cảnh năng lực để sản xuất các mô hình máy học tối tân, chúng tôi sẽ điều tra bối cảnh để triển khai ứng dụng thực tiễn, tập trung vào các vấn đề dự báo, khi đó các hệ thống tích hợp AI sẽ được thiết kế với mục đích đưa ra các dự đoán “máy móc” nhưng cũng không yêu cầu phải thay thế các dự đoán của con người. Chúng tôi sẽ tìm kiếm các doanh nghiệp sở hữu đội ngũ kỹ sư dữ liệu hùng hậu, triển khai tích hợp các phân tích dự đoán vào

dòng chảy công việc. Chúng tôi cũng sẽ nhanh chóng phát hiện ra các tổ chức tài chính là những đối tượng đó, bởi họ tuyển dụng các đội nhóm kỹ sư dữ liệu chuyên trách dự đoán gian lận, rửa tiền, không tuân thủ quy định và các hành vi vi phạm khác trong giao dịch tài chính. Khi đó, chúng tôi sẽ tìm kiếm các doanh nghiệp nhỏ cởi mở đón nhận và ứng dụng phát triển AI mới nhất vào giải quyết những vấn đề này. Chúng tôi cũng sẽ phát hiện ra rằng chỉ có một số rất nhỏ các công ty như vậy ở Canada lúc bấy giờ. Trong đó số có Verafin với trụ sở tại St. John's, Newfoundland.

Chúng tôi nhận ra rằng, đã đến lúc lùi lại và suy nghĩ nhiều hơn về kinh tế học AI. Hướng tiếp cận của Verafin theo đúng cuốn sách *AI trong cuộc cách mạng công nghệ 4.0*. Không có gì bất ngờ ở đây. Điều ít rõ ràng hơn là vì sao nhiều ứng dụng khác lại mất quá nhiều thời gian triển khai trên diện rộng đến vậy. Chúng tôi nhận ra rằng chúng tôi phải cân nhắc, không chỉ các đặc tính kinh tế học của công nghệ mà còn của những hệ thống mà công nghệ vận hành. Chúng tôi phải hiểu được các lực lượng kinh tế dẫn đến thực trạng ứng dụng AI không đồng đều, cụ thể AI được ứng dụng nhanh chóng và rộng rãi vào tự động phát hiện gian lận tài chính và vào phát triển sản phẩm trong thương mại điện tử, thế nhưng chậm chạp trong ứng dụng vào tự



động hóa hoạt động bảo hiểm và phát hiện thuốc trong ngành dược phẩm.

Chúng tôi không phải là những người duy nhất xem nhẹ những thách thức trong việc triển khai AI vào các mô hình tổ chức hiện tại. Geoffrey Hinton, một đồng nghiệp của tôi tại University of Toronto Geoffrey Hinton, người được mệnh danh là “cha đỡ đầu của AI” với công trình tiên phong về học sâu, cũng có những dự báo đánh giá thấp sự khó khăn trong việc triển khai AI. Trước đó, ông đã từng châm biếm, “Nếu là bác sĩ X-quang, bạn giống như con sói đã vượt qua bờ vực, nhưng chưa nhìn xuống nên không nhận ra rằng phía dưới không phải là mặt đất. Mọi người nên ngừng đào tạo bác sĩ X-quang ngay bây giờ. Rõ ràng là trong vòng 5 năm nữa, học sâu sẽ làm tốt hơn công việc hiện tại của các bác sĩ X-quang. Mặc dù ông ta đã đúng về tốc độ tiến bộ của khoa học kỹ thuật, rằng AI hiện sẽ vượt trội so với các bác sĩ X-quang trong các nhiệm vụ chẩn đoán, thì 5 năm sau phát ngôn đó, theo báo cáo của American College of Radiology, số lượng sinh viên mới đào tạo về X-quang không hề suy giảm.

Chúng tôi bắt đầu nhận ra rằng chúng tôi đã bước vào thời khắc chưa từng có trong lịch sử – The Between Times – sau khi chứng kiến sức mạnh

của công nghệ này và trước khi nó được áp dụng rộng rãi. Chúng tôi một số triển khai là giải pháp điểm, những giải pháp trực diện. Ở đây, áp dụng AI chỉ là sự thay thế đơn giản của các phân tích dự đoán cũ hơn do máy tạo ra bằng các công cụ AI mới hơn (những công cụ này đang phát triển nhanh chóng, như Verafin). Trong khi đó, các triển khai khác yêu cầu thiết kế lại sản phẩm hoặc dịch vụ cũng như tái cơ cấu tổ chức để tận dụng được toàn bộ lợi ích của AI và đủ bù trừ cho chi phí áp dụng. Trong trường hợp cân bằng chi phí, công ty và chính phủ vẫn đang chạy đua để tìm ra giải pháp mang lại lợi nhuận.

Chúng tôi chuyển trọng tâm từ khám phá các mạng lưới thần kinh sang khám phá nhận thức của con người (cách chúng ta đưa ra quyết định), hành vi xã hội (tại sao một số ngành hứng thú hơn với việc nhanh chóng nắm bắt và ứng dụng AI trong khi những ngành khác lại phản đối), hệ thống sản xuất (các quyết định có sự phụ thuộc lẫn nhau), và cấu trúc ngành (cách chúng ta không đưa ra quyết định trong một số trường hợp nhất định khi bản thân không chắc chắn).

Để khám phá những hiện tượng này, chúng tôi đã gặp gỡ các nhà lãnh đạo công ty, giám đốc sản phẩm, doanh nhân, nhà đầu tư, kỹ sư dữ liệu và

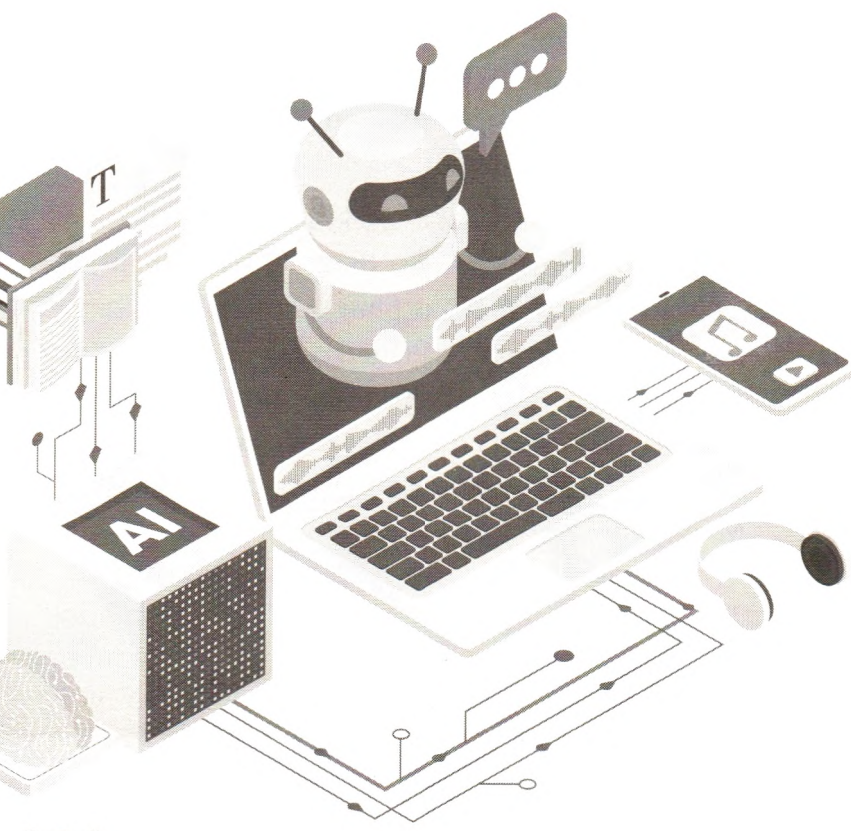


nhà kỹ sư máy tính đang triển khai AI. Chúng tôi đã tổ chức các hội thảo và hội nghị với các chuyên gia và nhà hoạch định chính sách. Đồng thời chúng tôi cũng nghiên cứu sát sao phương thức triển khai nào hiệu quả và phương thức nào không, thông qua hàng trăm thử nghiệm dưới hình thức khởi nghiệp AI.

Tất nhiên, chúng tôi cũng quay trở lại các nguyên tắc đầu tiên trong kinh tế học và nhìn nhận đó là một phần của làn sóng nghiên cứu thực nghiệm về kinh tế học AI, là lĩnh vực bùng nổ hiện tại dù hầu như mới tồn tại chỉ vài năm trước đó khi chúng tôi viết AI trong cuộc cách mạng công nghệ 4.0. Chúng tôi bắt đầu nối các điểm với nhau và hình thành một khuôn mẫu phân biệt giữa giải pháp điểm và giải pháp hệ thống. Điều này không chỉ giúp giải bài toán Verafin mà còn dự báo cho làn sóng áp dụng AI tiếp theo. Bằng cách tập trung vào các giải pháp hệ thống thay vì các giải pháp điểm, chúng tôi có thể giải thích cách thức công nghệ này cuối cùng sẽ lan rộng khắp các ngành, củng cố các công ty đang triển khai nó và gây gián đoạn ở những những công ty khác. Đã đến lúc viết một cuốn sách khác. Đây là cuốn sách đó.

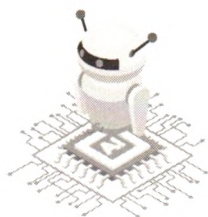
PHẦN 1

THE BETWEEN TIMES THỜI GIAN CHUYỂN ĐỔI



Chương 1

ẢN DỤ 3 DOANH NHÂN



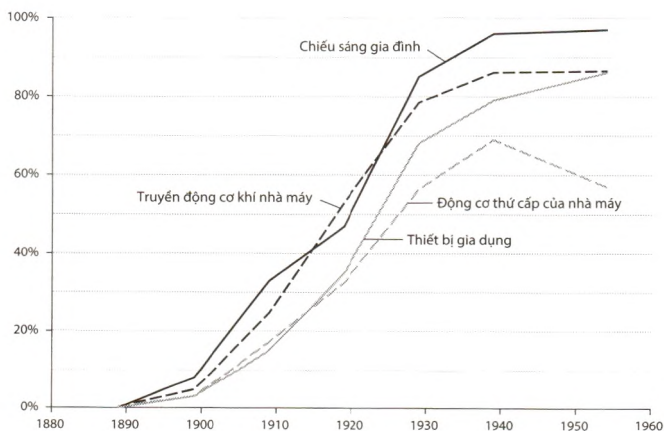
Sự xuất hiện của điện đã thay đổi hoàn toàn thế giới của chúng ta. Nó thay đổi cách chúng ta sống, cung cấp ánh sáng an toàn với giá cả vừa phải chỉ bằng một cái bật công tắc và giảm bớt gánh nặng việc nhà với tủ lạnh, máy giặt và máy hút bụi. Nó cũng thay đổi cách chúng ta làm việc, cung cấp điện năng cho các nhà máy và thang máy. Vậy, làm thế nào để mang đến tất cả những biến đổi này? Thời gian.

Sự phổ biến của điện khiến người ta khó hình dung rằng, vào đầu thế kỷ 20, hai thập kỷ sau khi Thomas Edison phát minh ra bóng đèn, điện gần như chưa xuất hiện ở bất kỳ đâu. Năm 1879, Edison thử nghiệm bóng đèn điện – một màn trình diễn vô cùng nổi tiếng và chỉ vài năm sau đó, đèn điện chiếu sáng nhà ga Pearl Street ở Manhattan và thắp sáng đường phố. Tuy nhiên, 20 năm sau đó, chỉ có 3% hộ gia đình Mỹ có điện. Trong các nhà máy, tỷ lệ này cũng không lớn hơn (xem Hình 1-1). Tuy nhiên, sau hai thập kỷ nữa, con số đó đã tăng lên một nửa dân số. Đối với

điện, 40 năm này chính là The Between Times – Thời gian chuyển đổi.

Có sự quan tâm rất lớn đối với điện nhưng không có nhiều điều để thể hiện. Chúng ta có xu hướng quên đi sự nhiệt tình này khi các công nghệ cấp tiến mới xuất hiện ngày nay. Khi đèn bật sáng, thay vì mọi thứ thay đổi, không có nhiều thay đổi. Bóng đèn AI đã bật. Nhưng chúng ta cần phải làm nhiều hơn nữa. Hiện tại chúng ta đang ở trong khoảng thời gian chuyển đổi của AI, từ chứng minh khả năng của công nghệ đến hiện thực hóa tiềm năng của thông qua áp dụng rộng rãi nó.

Hình 1.1: Triển khai lưới điện ở Mỹ



Nguồn: Data from Paul A. David, "Computer and Dynamo: The Modern Productivity Paradox in a Not-Too-Distant Mirror" (working paper #339, Stanford University, Department of Economics, 1989), [twerp339.pdf\(warwick.ac.uk\)](http://twerp339.pdf(warwick.ac.uk)).



Đối với AI, tương lai đó không phải là điều chắc chắn. Thế nhưng, chúng ta đã chứng kiến con đường phát triển của điện. Vì vậy, để hiểu hơn về những thách thức đối với việc thương mại hóa AI, hãy đặt mình vào vị trí của những doanh nhân vào những năm 1880. Điện là tương lai. Bạn sẽ hình dung tương lai đó như thế nào và làm thế nào để biến nó thành hiện thực?

GIẢI PHÁP ĐIỂM

Hơi nước cung cấp năng lượng cho nền kinh tế trong nửa sau của thế kỷ XIX. Than được đốt cháy để đun nóng nước tạo hơi và hơi nước sinh ra năng lượng. Ngay lập tức, năng lượng đó được sử dụng cho hoạt động của đòn bẩy, ròng rọc và từ đó cho phép hoạt động sản xuất công nghiệp diễn ra. Tất cả đều thừa nhận rằng, hơi nước là phép lạ thúc đẩy cuộc cách mạng lớn nhất trong nền kinh tế kể từ kinh tế nông nghiệp. Vì vậy, một doanh nhân muốn bán điện sẽ phải thúc đẩy khách hàng tiềm năng nghiên cứu kỹ hơn về hơi nước và xác định “mụn cóc” của nó.

Khi so sánh với điện năng, không khó để nhận thấy những điểm hạn chế của dạng năng lượng này. Hơi nước làm tiêu tan nhiệt – đó chính là vấn đề bởi phần lớn lượng nhiệt đó bị lãng phí. Năng lượng hơi nước bị mất từ 30-85% tiềm năng của nó do ngưng

tụ, van bị rò rỉ và ma sát do sử dụng trực và dây kéo để mang năng lượng đó đến khu vực sản xuất. Bạn có thể khó tưởng tượng được hệ thống dây trực này sẽ như thế nào. Hãy tưởng tượng nguồn năng lượng hơi nước ở một đầu, tác động quay một trục dài 3 inch bằng sắt hoặc thép, từ đó tạo lực kéo dây đai và ròng rọc hoạt động. Một số trục có thể nằm ngang, nhưng nhiều nhà máy có nhiều tầng với các trục ở dạng thẳng đứng. Một trục có thể cung cấp năng lượng cho hàng trăm máy.

Cơ hội đặt ra trước mắt là điện sẽ là nguồn năng lượng thay thế tại điểm mà năng lượng hơi nước được sử dụng, tức là ở cuối trục. Frank Sprague, một cựu nhân viên của Edison, đã nhìn ra cơ hội này vào năm 1886 khi ông phát triển một trong những động cơ điện đầu tiên. Trong khi Edison tập trung vào đèn điện, Sprague là một trong số những người nhận ra rằng chi phí của năng lượng điện vào ban ngày thấp và động cơ điện có thể tận dụng điều đó. Sprague đã vận dụng kiến thức và hiểu biết của mình để tạo năng lượng cho xe điện và thang máy xây dựng. Những người khác đưa động cơ vào các nhà máy.

Chúng tôi gọi đây là những “giải pháp điểm” bởi vì hơi nước được sử dụng và hoán đổi với nguồn năng lượng mới, cụ thể là điện, tại thời điểm chúng đi vào các nhà máy. Với phương thức này, vào cuối thế kỷ 19 có hai nhóm khách hàng sẵn



sàng xem điện là một nguồn năng lượng mới. Một nhóm là các nhà máy lớn chạy bằng hơi nước. Một nhà máy dệt ở Columbia, Nam Carolina, đã chuyển từ năng lượng hơi nước sang năng lượng điện vào năm 1893. Thủy điện là sự thay thế đơn giản cho hệ thống truyền tải cáp dài hàng dặm. Nguồn điện được cung cấp là rẻ nhất cả nước. Một nhóm khác là các nhà máy sản xuất quần áo và dệt may. Các “mụn cóc” của hơi nước là nó không phải nguồn năng lượng “sạch” và không có sự nhất quán về tốc độ tạo năng lượng. Năng lượng điện cải thiện được cả hai yếu tố này.

Giá trị hứa hẹn của giải pháp điểm là chi phí thấp hơn và các lợi ích khác dành riêng cho một số nhà máy. Cụ thể, đó là tính nhanh chóng tức thì, plug and play, kết nối điện và máy sẽ hoạt động. Nhưng trong nhiều trường hợp, năng lượng không phải là sản phẩm dễ bán. Chỉ có thể cắt giảm rất nhiều hóa đơn tiền điện bằng cách thay đổi nguồn điện. Điều mà giải pháp điểm không mang lại là nó không thúc đẩy *sử dụng nhiều năng lượng hơn*.

GIẢI PHÁP ỨNG DỤNG

Sau khi khởi động, một động cơ hơi nước bắt đầu chạy. Một động cơ điện có thể được tắt và sau đó bật lại. Do đó, khi năng lượng hơi nước di chuyển xuống cuối trục, kỹ sư điều khiển máy kích hoạt

các đòn bẩy tháo và lắp máy. Như vậy, kỹ sư điều khiển chỉ đơn giản là bật và tắt một động cơ điện gắn trực tiếp với từng máy. Cách này đơn giản hơn và ít yêu cầu bảo trì hơn nhiều. Tuy nhiên, điều này có nghĩa là lượng điện năng mà một nhà máy tiêu thụ sẽ thay đổi tùy theo mục đích sử dụng. Theo nhà sử học kinh tế Nathan Rosenberg, với cách thức này, kỷ nguyên “năng lượng được chia nhỏ” ra đời, cho phép “cung cấp điện theo các đơn vị rất nhỏ, ít tốn kém hơn và cũng dưới hình thức không yêu cầu sản xuất dư thừa chỉ để cung cấp một phần năng lượng nhỏ hoặc không liên tục”.

Dưới góc nhìn doanh nhân, giá trị của điện nằm ở chỗ nó yêu cầu ít năng lượng hơn, hoặc chính xác hơn là chỉ yêu cầu năng lượng khi cần. Quan điểm này bắt đầu tạo nên một số thay đổi trong thiết kế nhà máy, chẳng hạn tách riêng các nguồn năng lượng cho các loại máy khác nhau. Đồng thời, một số kỹ sư bắt đầu hình dung về việc sẵn có động cơ điện cho mỗi máy. Thậm chí, đối với các nhóm máy móc, việc chỉ trả tiền điện khi máy móc được sử dụng cũng có giá trị lớn.

Điểm thay đổi lớn là gắn một ổ điện trên một máy duy nhất. Ngày nay, chúng tôi sẽ gọi đây là một giải pháp ứng dụng. Thay vì chỉ hoán đổi một nguồn điện, toàn bộ thiết bị (tức là ứng dụng) đã được hoán đổi. Hơn nữa, máy móc trở nên di động



hơn nhiều. Không còn phải gắn vào một trục trung tâm, các công cụ có thể di chuyển xung quanh. Công việc không nhất thiết phải đến gần máy móc, mà máy móc có thể di chuyển đến công việc.

Đó là tiềm năng và sự hứa hẹn; còn thực tế là bất kỳ máy công cụ riêng lẻ nào – chẳng hạn như máy khoan, máy cắt kim loại hoặc máy ép – phải được thiết kế lại hoàn toàn để tận dụng lợi thế của việc có một động cơ điện riêng biệt. Hơn nữa, bản thân các động cơ nói chung không có sẵn nhưng phải được điều chỉnh cho phù hợp với máy hoặc mục đích sử dụng cụ thể. Cơ hội cho giải pháp ứng dụng là rất nhiều, nhưng cần thiết kế thiết bị. Hơn nữa, nếu bạn thiết kế một công cụ trong nhà máy với động cơ riêng của nó, điều đó sẽ làm giảm giá trị của bất kỳ động cơ nào cung cấp năng lượng cho các công cụ khác. Để có được sự cân bằng phù hợp rõ ràng sẽ yêu cầu thiết kế lại nhiều công cụ. Tuy nhiên, điều đó có nghĩa là tạo ra một hệ thống mới và cũng đồng nghĩa là mất rất nhiều thời gian.

GIẢI PHÁP HỆ THỐNG

Xuyên suốt toàn bộ cuộc Cách mạng Công nghiệp, các nhà máy được thiết kế để tận dụng hơi nước. Như chúng ta đã thấy, có một nguồn điện duy nhất đi vào nhà máy và sau đó điện được phân phối cho các máy lẻ thông qua một trục trung tâm

trên đó treo dây kéo và ròng rọc. Với cách nhìn hiện đại, đó giống như một cỗ máy lớn với con người bên trong như những bánh răng. Ở dạng lớn, nó là một cỗ máy trong đó hàng trăm bộ phận chuyển động được gắn vào một điểm cấp điện duy nhất. Phương thức hoạt động này không thay đổi khi có loại năng lượng mới thay thế. Thế nhưng, yêu cầu về thiết bị mới khiến một số doanh nhân phải suy nghĩ lại về nhà máy. Giả sử không có trực trung tâm hoặc thậm chí trực dành riêng cho các nhóm máy, khi đó *một nhà máy sẽ như thế nào nếu được thiết kế lại hoàn toàn từ đầu, dựa trên những hiểu biết về điện?*

Các nhà máy được xây dựng sao cho máy móc ở gần nguồn điện. Điều đó có nghĩa là có lợi thế cho những tòa nhà nhiều tầng với thiết kế thẳng đứng. Tuy nhiên, các nhà máy nhiều tầng chật chội vào cuối những năm 1800 cũng phải bỏ ra những chi phí tương ứng như để duy trì điều kiện làm việc, đảm bảo an toàn và hiệu suất máy móc. Sự ra đời của năng lượng điện đã loại bỏ yêu cầu phải dồn tất cả những thứ này vào những không gian nhỏ.

Nhiều nhà quản lý doanh nghiệp nhận ra giá trị thực của điện năng đi cùng với một giải pháp hệ thống – cụ thể, đó là một hệ thống có thể tận dụng toàn bộ điện sẵn có để cung cấp. Nói về hệ thống, đó là tập hợp các quy trình vận hành cùng nhau để đảm bảo hoàn thành một việc cụ thể.



Hãy nghĩ đến những lợi ích kinh tế của quy mô hay không gian trong một nhà máy. Với trục hơi nước trung tâm, không gian gần trục có giá trị hơn không gian ở những nơi khác. Vì vậy, hoạt động sản xuất được tiến hành gần trục, và mọi nhu cầu khác phải để lại và chuyển đi nơi khác. Điều đó có nghĩa là mọi thứ di chuyển qua lại theo yêu cầu về năng lượng.

Điện đã cân bằng giá trị kinh tế của không gian, mang lại sự linh hoạt trong sản xuất. Giờ đây, các nhà máy sẵn sàng tổ chức sản xuất theo dây chuyền sao để cắt giảm quãng đường di chuyển qua lại và thay vào đó sản xuất sẽ đi từ quy trình này sang quy trình tiếp theo. Với năng lượng hơi nước, Henry Ford không thể phát minh ra dây chuyền sản xuất xe Model T. Chỉ khi có điện, nhiều thập kỷ sau đó để hiện thực hóa tiềm năng thương mại của nó, ông mới làm được điều đó. Vâng, Ford là một doanh nhân về xe hơi. Nhưng ông cũng chính là một doanh nhân giải pháp hệ thống. Những thay đổi hệ thống này đã làm thay đổi cục diện công nghiệp. Chỉ khi điện khí hóa cuối cùng đã xuất hiện trong các sổ liệu thống kê hiệu suất, và một cách đáng kể.

GIẢI PHÁP AI

Chúng ta có thể rút ra 3 bài học từ điều này. Đầu tiên, con đường cải thiện hiệu suất vượt trội nằm ở việc hiểu rõ công nghệ mới có thể mang lại những gì.

Một doanh nhân khi quảng cáo về điện vào năm 1890 sẽ tập trung vào yếu tố “tiết kiệm chi phí nhiên liệu” và nhấn mạnh đó là tiềm năng giá trị quan trọng cho công nghệ. Thế nhưng, điện không chỉ là một động cơ hơi nước ít đắt đỏ hơn. Giá trị thực sự của điện là nó có cách để tách rời giữa nguồn năng lượng và nơi sử dụng. Điều đó giúp giải phóng người dùng khỏi sự ràng buộc về khoảng cách, mở ra hàng loạt cải tiến về thiết kế quy trình làm việc và nhà máy. Đến năm 1920, khi quảng cáo về điện, các doanh nhân hẳn đã nhận ra tiềm năng giá trị quan trọng của điện không nằm ở yếu tố “tiết kiệm chi phí nhiên liệu” mà là “cho phép nhà máy được cơ cấu và thiết kế mang lại hiệu suất lớn hơn”.

Chúng tôi kỳ vọng AI sẽ phát triển theo mô hình tương tự. Như chúng tôi đã lưu ý, tiềm năng thương mại ban đầu bao gồm cả những giải pháp điểm như của Verafin, khi công ty này thay thế cách thức dự báo cũ sang cách thức mới tốt hơn, nhanh hơn và rẻ hơn.

Chúng tôi cũng nhận thấy các giải pháp ứng dụng đi kèm yêu cầu thiết kế lại thiết bị hoặc sản phẩm xung quanh AI. Tất cả những máy móc (robot) này đều được hỗ trợ bởi AI và được gọi là ứng dụng. Đây cũng là cách chủ yếu mà AI được triển khai để cải thiện phần mềm trên thiết bị của bạn. Ví dụ với máy ảnh điện thoại của bạn, nó có thể



nhận dạng khuôn mặt của bạn. Để làm được như vậy, đòi hỏi phải có một máy ảnh đặc biệt cũng như phần cứng chuyên dụng để bảo mật thông tin đó. Thế nhưng, sự thúc đẩy rõ ràng nhất của loại hình đổi mới này là hàng tỉ đô la đầu tư cho nỗ lực thiết kế và tung ra các phương tiện tự hành trong điều kiện giao thông hiện tại. Mặc dù bề ngoài những chiếc xe có thể trông giống nhau, nhưng phần cứng bên trong chúng phải được thiết kế lại để đặt cảm biến, xử lý trên bo mạch và sau đó là xử lý máy.

Và những gì chúng ta chưa thấy là rất nhiều giải pháp hệ thống có giá trị cao cho AI có khả năng xuất hiện. Cuốn sách này sẽ chỉ ra tiềm năng cũng như những thách thức trong việc hiện thực hóa những cơ hội đó.

Thứ hai, một khi đã hiểu rõ những điều này, chúng ta cần đặt một câu hỏi khá đơn giản nhưng không dễ để trả lời. Với hiểu biết hiện tại về AI, chúng ta sẽ thiết kế sản phẩm, dịch vụ hoặc nhà máy của mình như thế nào nếu bắt đầu từ con số 0? Kiến trúc nhà máy phẳng mới không xuất hiện đầu tiên trong ngành công nghiệp truyền thống mà đúng hơn là trong những ngành kinh tế mới nổi vào những năm 1900 như thuốc lá, kim loại chế tạo, thiết bị vận tải và cả máy móc điện tử. Chúng ta chứng kiến sự bùng nổ với những ứng dụng ban đầu đặt AI làm trọng tâm thiết kế hệ thống trong các ngành công nghiệp số hóa mới ngày nay:

ngiên cứu, thương mại điện tử, nội dung phát trực tuyến và mạng xã hội.

Đối với AI, hai câu hỏi đặt ra là: (1) AI thực sự mang lại cho chúng ta điều gì? (2) Nếu thiết kế lại doanh nghiệp của mình từ đầu, chúng ta sẽ xây dựng quy trình và mô hình kinh doanh của mình như thế nào? Nếu điện không phải là “năng lượng thay thế với chi phí thấp hơn” mà là “nguồn năng lượng cải thiện thiết kế nhà máy giúp tăng hiệu suất hơn nhiều”, thì có lẽ AI cũng không phải là “chi phí dự đoán thấp hơn” mà là “giải pháp thúc đẩy các sản phẩm, dịch vụ và thiết kế tổ chức hiệu quả hơn nhiều”. Nếu lợi ích chính của điện là tách rời nguồn năng lượng và nơi nó được sử dụng, tạo điều kiện đổi mới thiết kế nhà máy, thì lợi ích chính của AI là nó tách rời dự đoán khỏi phần còn lại của quá trình ra quyết định, thúc đẩy sáng tạo, đổi mới trong cơ cấu tổ chức thông qua tái hình dung mối liên hệ giữa các quyết định với nhau.

Chúng tôi lập luận rằng bằng cách tách dự đoán khỏi các khía cạnh khác của quyết định và chuyển dự đoán từ con người sang máy móc, AI cho phép đổi mới ở cấp độ hệ thống. Quyết định đóng vai trò là yếu tố chủ chốt để xây dựng các hệ thống như vậy và AI nâng cao khả năng ra quyết định.

Bài học thứ 3 và cũng là bài học cuối cùng: các loại giải pháp khác nhau mang đến những cơ



hội khác nhau để giành được năng lượng trên thị trường. Lợi nhuận của các doanh nghiệp đến từ việc họ vừa tạo nên vừa giành lấy giá trị. Với các giải pháp điểm, vấn đề thường là không có nhiều giá trị được tạo ra ngay từ đầu. Điện là năng lượng thay thế cho hơi nước, nhưng hơi nước đã có cơ sở lắp đặt sẵn. Việc đổi cái này lấy cái kia không phải là miễn phí và nếu bạn làm như vậy, đề xuất giá trị cho người tiêu dùng của bạn là giảm hóa đơn tiền điện. Nói cách khác, giải pháp điểm có thể mang lại nguồn lợi nhuận ổn định thông qua cung cấp giải pháp điểm tốt nhất, như cách mà Verafin đã làm, nhưng đó là trường hợp tốt nhất.

Khi chuyển sang giải pháp ứng dụng và sau đó là các hệ thống, giá trị mà được tạo ra được đảm bảo và bảo vệ hơn. Các thiết bị mới được phân biệt với đối thủ cạnh tranh và được bảo vệ bằng bằng sáng chế và các hình thức bảo vệ sở hữu trí tuệ khác. Tuy nhiên, đối với các hệ thống mới, tiềm năng thậm chí còn lớn hơn. Với điện, chủ sở hữu nhà máy có thể quyết định phần lớn đối với thiết kế nhà máy mới. Điều này cũng giúp họ trang bị hiểu biết về tranh giành thị phần và tự bảo vệ mình khỏi sự cạnh tranh. Mặc dù cách bố trí của nhà máy là yếu tố hữu hình, dễ dàng nhìn thấy, các quy trình, khả năng và hoạt động đào tạo bên dưới hệ thống mới là cái khó thấy hơn và cũng khó để bắt chước.

Hơn nữa, các hệ thống mới có thể kích hoạt mở rộng và phát triển về quy mô.

GIÁN ĐOẠN AI VÀ NĂNG LƯỢNG

Điện đã mất hàng thập kỷ để tạo nên cái mà chúng ta gọi là “gián đoạn”. Trong suốt hai thập kỷ đầu tiên, nó được sử dụng như một giải pháp điểm trong một số nhà máy và ứng dụng, và để chiếu sáng ở những nơi khác. Nhưng đến khi các hệ thống mới được phát triển, điện mới thực sự thay đổi nền kinh tế. Đó là sự thay đổi sâu rộng, đưa quyền lực chuyển giao cho những người kiểm soát việc phát điện và lưới điện cũng như cho những người có thể sử dụng điện trên quy mô lớn trong sản xuất hàng loạt. Sau đó, bạn sẽ không còn muốn trở thành nhà sản xuất dây đai và ròng rọc hay người nắm giữ một nhà máy ở trung tâm thành phố nữa.

Chúng tôi thấy những quy trình tương tự đang diễn ra với AI. Đó là những thay đổi thực sự về quyền lực kinh tế, cụ thể quyền kiểm soát đối với những nguồn lực và tài sản khan hiếm được chuyển từ nhóm người này sang nhóm người khác đi kèm với khả năng bảo vệ các doanh nghiệp khỏi áp lực cạnh tranh. Cơ hội để AI thực hiện được điều này là chắc chắn, thế nhưng những cơ hội tạo nên sự gián đoạn – tức là sẽ định hình lại các ngành công nghiệp và người nắm quyền trong đó phải đến từ



các hệ thống mới. Không dễ dàng để phát triển các hệ thống mới và cũng rất khó để sao chép chúng (chúng ta sẽ khám phá điều này sau) vì thường chúng rất phức tạp. Thế nhưng chính điều này mở ra cơ hội cho những người có thể đổi mới hệ thống.

Tuy nhiên, vẫn tồn tại những băn khoăn, nghi ngại. Đối với AI, ai có thể tích lũy sức mạnh từ những công nghệ mới này? Đó vẫn là một câu hỏi còn bỏ ngỏ. Điều đó sẽ phụ thuộc vào liệu những hệ thống mới đó trông như thế nào. Nhiệm vụ của chúng tôi ở đây là giúp bạn “sáng tỏ” để có thể dự đoán ai sẽ giành được và ai có thể mất đi quyền lực khi các hệ thống AI phát triển và được áp dụng.

ĐIỂM CHÍNH

- Câu chuyện về 3 giải pháp, lấy bối cảnh hơn 100 năm trước và tập trung vào thị trường năng lượng, minh họa cách các doanh nhân khác nhau khai thác cùng một sự chuyển đổi công nghệ, từ hơi nước sang điện, có thể khai thác các giá trị tiềm năng khác nhau: giải pháp điểm (chi phí điện năng thấp hơn và tổn thất ít hơn do ma sát – không thay đổi thiết kế đến hệ thống nhà máy); giải pháp ứng dụng (truyền động điện riêng lẻ trên mỗi máy – máy dạng mô-đun, do đó việc dừng máy này không ảnh hưởng đến

máy khác; không thay đổi thiết kế đối với hệ thống nhà máy); và hệ thống các giải pháp (nhà máy được thiết kế lại – xây dựng nhẹ, một tầng, quy trình công việc được tối ưu hóa về mặt bố trí không gian và luồng công nhân và vật liệu).

- Một số đề xuất giá trị hấp dẫn hơn những đề xuất khác. Trong trường hợp điện, các giải pháp điểm và giải pháp ứng dụng dựa trên việc thay thế trực tiếp hơi nước bằng điện mà không sửa đổi hệ thống mang lại giá trị hạn chế, điều này được phản ánh trong việc áp dụng ban đầu chậm chạp của các ngành. Theo thời gian, một số doanh nhân đã nhìn thấy cơ hội cung cấp các giải pháp cấp hệ thống bằng cách khai thác khả năng của điện để tách máy ra khỏi nguồn điện theo cách không thể hoặc quá tốn kém với hơi nước. Trong nhiều trường hợp, đề xuất giá trị của các giải pháp cấp hệ thống vượt xa giá trị từ các giải pháp điểm.
- Giống như điện cho phép tách máy khỏi nguồn điện và do đó tạo điều kiện chuyển đổi đề xuất giá trị từ “chi phí nhiên liệu thấp hơn” sang “thiết kế nhà máy hiệu quả hơn rất nhiều”, AI cho phép tách dự đoán khỏi các khía cạnh khác của một quyết định và do đó tạo điều kiện chuyển đổi đề xuất giá trị từ “chi phí dự đoán thấp hơn” đến “các hệ thống hiệu quả hơn rất nhiều”.

Chương 2

TƯƠNG LAI CỦA AI



Năm 2017 tràn ngập các hội nghị về AI. “Trận đại hồng thủy” đó đã tập hợp các doanh nghiệp và chính phủ lại với nhau, đồng thời cũng khích lệ các học giả. Khi nhận thấy AI có tiềm năng chuyển đổi nền kinh tế, chúng tôi mong muốn khơi gợi chủ đề này với một số nhà nghiên cứu kinh tế giỏi nhất trên thế giới. Chúng tôi đã tổ chức một hội nghị ở Toronto về AI để xây dựng một chương trình nghiên cứu cho các nhà kinh tế học.

Thật ngạc nhiên, chúng tôi không gặp khó khăn để thu hút sự chú ý và tập hợp được một nhóm các chuyên gia. Paul Milgrom của Đại học Stanford, người sẽ tiếp tục giành giải thưởng Nobel cho những sáng kiến, đổi mới trải rộng ở lĩnh vực kinh tế và khoa học máy tính, là một trong số các khách mời. Ông nhớ lại vào năm 1990 ông cũng nhận được lời mời tương tự, tham gia một hội nghị về tiềm năng kinh tế của Internet. Ông đã từ chối tham

gia và cảm thấy hối hận về điều đó. “Tôi nhớ rất rõ, vào năm 1990, NSF hỏi tôi liệu tôi có hứng thú nghiên cứu về tính kinh tế của Internet hay không, và tôi đã quá bận với những nghiên cứu của mình xoay quanh các yếu tố kinh tế của doanh nghiệp và tính siêu mô đun. Vì vậy, tôi đã từ chối”, ông viết. “Lần này không có bất kỳ lý do nào nữa. Vâng, tôi sẽ tham gia”.

Một số người tham gia tỏ ra tin tưởng về tác động của AI. Daniel Kahneman, một chuyên gia khác cũng từng đoạt giải Nobel, nhận xét: “Tôi không nghĩ rằng có nhiều việc chúng ta có thể làm mà máy tính không được lập trình để làm”. Betsey Stevenson, người từng phục vụ trong Hội đồng Cố vấn Kinh tế của Tổng thống Obama, cũng có những nhận xét tóm tắt tâm lý lạc quan này, “Rõ ràng là các nhà kinh tế học tin rằng, trí tuệ nhân tạo đại diện cho cơ hội đạt được những lợi ích kinh tế đáng kể”.

Những người tham gia khác lại tỏ ra hoài nghi hơn. Joseph Stiglitz, một chuyên gia khác đoạt giải Nobel khác, tỏ ra lo lắng về tác động (của AI) đối với sự bất bình đẳng. Tyler Cowen, nhà kinh tế học và từng là người phụ trách chuyên mục của tờ *New York Times*, lo lắng rằng hiệu suất mà AI mang lại sẽ làm gia tăng sự khan hiếm tài nguyên vật chất. Manuel Trajtenberg lưu ý rằng xét về dài hạn những lợi ích của công nghệ sẽ không liên quan



nếu một cuộc cách mạng xảy ra trước, theo đó báo trước sự phản đối ngày càng tăng đối với tự động hóa máy móc và nhận thức phổ biến về tác động đối với việc làm.

Một lo lắng thú vị khác là AI dường như không có nhiều tác động đến nền kinh tế. Cụ thể, các nhà kinh tế học Erik Brynjolfsson, Daniel Rock và Chad Syverson cho rằng:

Chúng ta đang sống trong một thời đại đầy nghịch lý. Ở nhiều lĩnh vực khác nhau, các hệ thống sử dụng trí tuệ nhân tạo đuổi kịp thậm chí vượt qua hiệu suất mà con người đạt được. Những hệ thống mới này tận dụng những tiến bộ vượt bậc trong các công nghệ khác, và từ đó thúc đẩy giá cổ phiếu tăng vọt. Tuy nhiên, tăng trưởng năng suất được đo lường lại giảm một nửa trong thập kỷ qua và thu nhập của phần lớn người dân Mỹ thực tế cũng chững lại kể từ cuối những năm 1990.

Đối với những người nghiên cứu lịch sử công nghệ (và như chúng ta đã quan sát được ở điện), nghịch lý này không phải là chưa từng có tiền lệ. Năm 1987, Robert Solow của MIT đã nhận xét một cách châm biếm và câu nói này trở nên vô cùng nổi tiếng “chúng ta nhìn thấy thời đại máy tính ở mọi nơi trừ trong các số liệu thống kê về năng suất”.

Máy tính xuất hiện khắp nơi nhưng không đo lường được cụ thể về hiệu quả cải thiện năng suất của chúng. Đây không phải là điều gì xa lạ, vì vậy các nhà kinh tế học bắt đầu quan tâm đến điều gì sẽ xảy ra khi “các công nghệ đa năng” – những công nghệ cho phép tăng trưởng năng suất bền vững trong nhiều lĩnh vực – xuất hiện. Các công nghệ đa năng bao gồm động cơ hơi nước và điện, chất bán dẫn và Internet như các khởi tạo gần đây hơn. Đối với những người tham gia hội nghị của chúng tôi, AI có vẻ như là một “ứng viên” hợp lý để thêm vào danh sách. Chúng ta nên mong đợi điều gì? Đúng vậy, xuyên suốt lịch sử, những công nghệ như vậy cuối cùng đã biến đổi nền kinh tế, biến đổi doanh nghiệp và việc làm, thế nhưng những điều gì đã xảy ra trong suốt nhiều thập kỷ là tất cả những gì đang xảy ra? Điều gì đã xảy ra trong khoảng thời gian chuyển giao này?

ĐỐI VỚI HỆ THỐNG CHO AI

Giám đốc điều hành Google Sundar Pichai cho biết, “AI có lẽ là thứ quan trọng nhất mà nhân loại từng nghiên cứu. Tôi nghĩ nó là một cái gì đó sâu rộng, toàn diện hơn điện”. Google đã thấy được lợi ích từ AI. Nhiều công ty không như vậy. Theo một nghiên cứu năm 2020 của *Sloan Management Review* thuộc MIT và BCG, một công ty tư vấn toàn cầu, cho thấy chỉ 11% tổ chức báo cáo những lợi ích tài



chính đáng kể từ AI. Đây không phải là kết quả do thiếu cố gắng. 59% cho biết họ còn xây dựng một chiến lược AI hoàn chỉnh. 57% đã triển khai hoặc thử nghiệm các giải pháp AI.

Nhà tiên phong về AI, Andrew Ng, người sáng lập dự án Google Brain và là nhà khoa học trưởng tại Baidu, đã tuyên bố rằng “AI là nguồn điện mới. Nó có tiềm năng biến đổi mọi ngành công nghiệp và tạo ra giá trị kinh tế to lớn”. Chúng tôi đồng ý (với nhận định này). AI có tiềm năng biến đổi của điện, nhưng nếu lịch sử là kim chỉ nam, thì sự biến đổi đó sẽ là một hành trình dài và gập ghềnh.

Ví dụ về điện cho thấy không có sự mâu thuẫn cố hữu nào giữa sự lạc quan về tương lai của AI và sự thất vọng về kết quả thực tế cho đến hiện tại. Brynjolfsson, Rock và Syverson đã nêu bật thời đại nghịch lý này. Chúng ta có thể dự báo về sự lạc quan đối với tương lai nhưng cũng lưỡng lự trước sự thất vọng về xuất phát điểm của chúng ta ở hiện tại. Thật vậy, có những lý do chính đáng về mặt lý thuyết giải thích cho sự tồn tại đồng thời của hai mặt này khi nền kinh tế trải qua quá trình tái cấu trúc gắn liền với những biến đổi của công nghệ.

Trong làn sóng đầu tiên của điện, bóng đèn thay thế nến và động cơ điện thay thế động cơ hơi nước. Đây là những giải pháp điểm và không đòi hỏi tái cơ cấu. Nền kinh tế không chuyển đổi.

AI cũng đang ở tình trạng tương tự. Nó được áp dụng như một công cụ mới trong phân tích và dự báo. Một số công ty như Verafin nhìn thấy lợi ích từ việc cải thiện và nâng cấp khả năng dự đoán. Đây chính là 11% các công ty đã nhìn thấy lợi ích tài chính từ AI. Họ đã thực hiện công việc dự báo và với AI, dự báo của họ tốt hơn, nhanh hơn và rẻ hơn. Phần “quả tầm thấp” của AI chính là các giải pháp điểm và phần quả này đang được hái rồi.

Giống như tiềm năng thực sự của điện chỉ được giải phóng khi những lợi ích rộng lớn hơn từ sản xuất điện được nhìn nhận và khai thác, AI sẽ chỉ đạt được tiềm năng thực sự khi những lợi thế của nó trong việc đưa ra dự đoán được tận dụng triệt để. Đối với chúng tôi, điều đó chỉ rõ vai trò của dự đoán trong việc nâng cao khả năng ra quyết định. Chúng tôi sẽ chứng minh rằng, trong nhiều trường hợp, dự đoán sẽ thay đổi cách ra quyết định đến mức toàn bộ hệ thống ra quyết định và quy trình của nó trong các tổ chức sẽ phải điều chỉnh theo. Chỉ khi đó, việc áp dụng AI mới thực sự bùng nổ.

Chúng ta đang ở trong Giai đoạn chuyển giao (The Between Times), sau khi tiềm năng của AI đã được chứng minh rõ ràng nhưng vẫn chưa đến lúc nó thực sự có tác động biến đổi. Verafin giống như 11% các tập đoàn lớn đã thành công triển khai AI vì



các dự đoán phù hợp với hệ thống hiện có của họ. Quy trình và dòng chảy công việc đã sẵn sàng để tận dụng những dự đoán đó mà không yêu cầu sửa đổi đáng kể.

Đối với 89% còn lại, hệ thống chưa sẵn sàng. Tiềm năng là rõ ràng và hứa hẹn, nhưng con đường để hiện thực hóa tiềm năng đó thì không. Cần phải có một cách để sử dụng dự đoán của máy móc một cách hiệu quả hơn. Điều đó có nghĩa là sử dụng dự đoán để đưa ra quyết định tốt hơn.

Tác động của AI sẽ xoay quanh những gì con người có thể làm và giúp họ có thể đưa ra quyết định tốt hơn. Đó không chỉ là thách thức về mặt kỹ thuật như thu thập dữ liệu, xây dựng mô hình và đưa ra dự đoán mà còn là thách thức về mặt tổ chức để đúng người đưa ra quyết định đúng vào đúng thời điểm. Và đó là thách thức về mặt chiến lược để làm thế nào xác định được có thể làm gì khác đi nếu được tiếp cận thông tin tốt hơn.

TẠO TIỀN ĐỀ

Đặc trưng của giai đoạn này là sự háo hức và thành công với các giải pháp điểm, nhưng với AI dường như vẫn là một công nghệ “hẹp”. Trong giai đoạn này, giải pháp ứng dụng cũng đã đạt được những phát triển và thử nghiệm thành công nhất định.

Tuy nhiên, về bản chất, các giải pháp này thường rất cụ thể, chẳng hạn để tăng cường các sản phẩm hiện có như điện thoại hoặc các tính năng an toàn cho ô tô.

Cục điều tra dân số Hoa Kỳ đã thực hiện một khảo sát 300.000 công ty về việc sử dụng AI của họ. Các công ty lớn có áp dụng AI đặc biệt nhấn mạnh việc sử dụng AI để tự động hóa và cải thiện các quy trình hiện có. Nói cách khác, AI mà họ triển khai là giải pháp điểm cũng là giải pháp ứng dụng, vì vậy không có sự thay đổi trong hệ thống. Những AI này chỉ tác động khiêm tốn đến năng suất của các tổ chức triển khai chúng. Việc xem xét các quy trình công việc hiện có và xác định AI có thể thay thế con người ở đâu có thể gia tăng các lợi ích thiết thực hơn. Thế nhưng, đó không phải nơi tiềm ẩn những cơ hội lớn nhất.

Trong Giai đoạn chuyển giao, các doanh nhân và nhà quản lý doanh nghiệp đấu tranh để việc áp dụng các ứng dụng trở nên khả thi về mặt kinh tế. Theo Nathan Rosenberg, đối với tất cả các công nghệ, “nhiều trường hợp thất bại trong kinh doanh có thể là do doanh nhân tương lai đó đã không xem xét các điều kiện và sự phụ thuộc lẫn nhau giữa điều mà anh ta quan tâm và phần còn lại của hệ thống lớn hơn”.



Sự chuyển đổi thực sự sẽ chỉ đến khi các nhà đổi mới tập trung vào việc tạo ra các giải pháp hệ thống mới. Bản thân những giải pháp đó sẽ đưa AI lên quy mô toàn nền kinh tế và tạo đà thúc đẩy các giải pháp ứng dụng tiếp theo. Tiềm năng về quy mô và thúc đẩy tiếp tục đổi mới khiến các hệ thống AI trở nên hấp dẫn về mặt kinh tế để theo đuổi.

Với tầm quan trọng của những giải pháp này, chúng ta phải rõ ràng và cẩn thận truyền tải điều chúng ta muốn. Vì vậy, hãy xác định các khái niệm của chúng tôi:

- Một giải pháp điểm cải thiện quy trình hiện có và có thể được áp dụng độc lập mà không làm thay đổi hệ thống mà nó được nhúng vào.
- Một giải pháp ứng dụng cho phép quy trình mới được áp dụng độc lập mà không làm thay đổi hệ thống mà nó được nhúng vào.
- Một giải pháp hệ thống cải thiện các quy trình hiện có hoặc kích hoạt các quy trình mới bằng cách thay đổi các quy trình phụ thuộc.

Phần khó của việc phân biệt các định nghĩa đã được giải quyết bằng trạng từ “độc lập”. Từ này xuất hiện trong định nghĩa giải pháp ứng dụng và điểm nhưng không xuất hiện trong định nghĩa giải

pháp hệ thống. Hãy tưởng tượng chúng ta đang vận hành một quy trình sẵn có hoặc xây dựng một hệ thống, và có thể gia tăng giá trị của nó bằng cách triển khai một công nghệ mới. Nếu giá trị gia tăng lớn hơn chi phí bỏ ra để phát triển và áp dụng giải pháp đó, thì giải pháp đó là khả thi về mặt kinh tế. Hơn nữa, tính khả thi về mặt kinh tế của giải pháp đó không phụ thuộc vào việc có điều gì khác thay đổi hay không. Tuy nhiên, giả sử rằng lợi ích từ công nghệ mới đó là quá thấp và chỉ có thể được cải thiện bằng cách thay đổi những thứ khác. Khi đó, việc áp dụng độc lập mà không có những thay đổi đó là không thể về mặt kinh tế. Việc áp dụng yêu cầu đồng thời thay đổi nhiều quy trình.

Do đó, các nhà máy cảm thấy dễ dàng khi sử dụng điện như một giải pháp điểm và để thay thế cho hơi nước. Và một số ứng dụng cũng có thể được xây dựng để tích hợp với động cơ điện và được sử dụng trong các hệ thống sản xuất hiện có. Tuy nhiên, trong nhiều tình huống, các nhà máy cần phải thiết kế lại và cần một hệ thống điện và lưới điện tập trung lựa chọn năng lượng mới này trở nên khả thi về mặt kinh tế. Nói cách khác, các giải pháp hệ thống đã biến điện từ năng lượng thay thế cho lựa chọn hiện có thành cơ hội sử dụng nguồn năng lượng mới.



Trong chương tiếp theo, chúng ta sẽ lật lại một chủ đề từ cuốn *AI trong cuộc cách mạng công nghệ 4.0*. Cụ thể, những tiến bộ trong AI hiện đại, về bản chất, là sự cải tiến trong công nghệ dự báo. Hơn nữa, các dự báo chỉ có giá trị là thông tin đầu vào cho quá trình ra quyết định. Vì vậy, chúng tôi chỉnh sửa các định nghĩa trước đó theo các mục đích của cuốn sách này:

- **GIẢI PHÁP ĐIỂM AI:** Một dự báo có giá trị như một giải pháp điểm nếu nó cải thiện một quyết định hiện có và quyết định đó có thể được đưa ra một cách độc lập.
- **GIẢI PHÁP ỨNG DỤNG AI:** Một dự báo có giá trị như một giải pháp ứng dụng nếu nó cho phép đưa ra quyết định mới hoặc thay đổi cách đưa ra quyết định và quyết định đó có thể được đưa ra một cách độc lập.
- **GIẢI PHÁP HỆ THỐNG AI:** Một dự báo có giá trị như một giải pháp hệ thống nếu nó cải thiện các quyết định hiện có hoặc thúc đẩy các quyết định mới, nhưng chỉ khi có sự thay đổi đối với cách thức đưa ra quyết định.

Đối với các công nghệ khác, mặc dù chúng ta vẫn hưởng lợi sau khi nhận thức rõ đâu là các yếu tố độc lập và phụ thuộc, thì với AI, chúng ta vẫn phải tìm ra những khía cạnh đó của hệ thống. Cuốn sách này sẽ tập trung để tìm ra chúng.

THAY ĐỔI HỆ THỐNG TẠO SỰ GIÁN ĐOẠN

Sự gia tăng lớn nhất trong việc áp dụng AI, nếu lịch sử là bất kỳ hướng dẫn nào, sẽ đến từ những thay đổi trong hệ thống. Nhưng sự thay đổi như vậy cũng sẽ gây rối. Khi nói đột phá, chúng tôi muốn nói rằng nó thay đổi vai trò của nhiều người và nhiều công ty trong các ngành và cùng với những thay đổi đó, gây ra sự thay đổi quyền lực. Nghĩa là, có khả năng sẽ có người thắng và người thua về mặt kinh tế, đặc biệt nếu sự thay đổi hệ thống diễn ra tương đối nhanh chóng.

Để hiểu được sự gián đoạn đó, hãy xem xét dự đoán trong canh tác. Nông nghiệp là một ngành công nghiệp mà cơ giới hóa làm giảm đáng kể việc làm. Nhưng quản lý trang trại vẫn nằm trong tay người nông dân. Những người nông dân đã sử dụng các dự đoán về thời tiết để hỗ trợ cho những quyết định đó, nhưng bản chất của vùng đất mà họ sở hữu là thứ gắn liền với kỹ năng dự đoán và ra quyết định của họ nói chung.

Nông dân tiếp xúc với các điều kiện thời tiết, nhưng quan trọng là cách họ tiếp xúc khác nhau tùy thuộc vào loại cây trồng và điều kiện đồng ruộng tại địa phương. Rủi ro bổ sung này là điều mà David Friedberg, người đầu tiên cung cấp các dự đoán thời tiết có thể truy cập qua Internet, đã nhận ra khi



ông cố gắng bán bảo hiểm cho nông dân Hoa Kỳ. Giống như dữ liệu thời tiết, chính phủ Hoa Kỳ có dữ liệu – dưới dạng hình ảnh vệ tinh hồng ngoại và dữ liệu về thành phần đất của 29 triệu cánh đồng – cho phép Friedberg tính toán rủi ro liên quan đến thời tiết ở cấp độ cánh đồng hoặc cây trồng.

Friedberg thành lập The Climate Corporation để bán bảo hiểm cho nông dân, nhưng nhanh chóng nhận ra rằng họ cũng quan tâm đến dữ liệu mà ông có về lĩnh vực của họ:

Anh ấy [Friedberg] sẽ cho người nông dân thấy chính xác độ ẩm của cánh đồng tại bất kỳ thời điểm nào – trên một mức nhất định, cánh đồng sẽ bị hư hại nếu tiếp tục. Anh ấy sẽ cho họ xem lượng mưa và nhiệt độ hằng ngày – điều mà bạn có thể nghĩ rằng người nông dân sẽ biết, nhưng khi đó người nông dân có thể đang quản lý 20 hoặc 30 cánh đồng khác nhau, trải rộng trên nhiều quận. Anh ấy sẽ chỉ cho người nông dân biết chính xác giai đoạn phát triển của cây trồng, thời điểm tốt nhất để bón phân, khoảng thời gian 8 ngày tối ưu để gieo hạt và ngày thu hoạch lý tưởng.

Dự đoán là một động lực lớn cho các quyết định quan trọng của nông dân: bón phân, gieo hạt

và thu hoạch. Mục tiêu của những quyết định này gần như mang tính phổ biến – để tối đa hóa năng suất: “Việc trồng trọt luôn liên quan đến các quyết định đánh giá dựa trên bản năng của người nông dân. Tập đoàn Khí hậu đã biến canh tác thành khoa học quyết định và là vấn đề xác suất. Người nông dân không còn chơi roulette nữa mà chơi blackjack. Và David Friedberg đang giúp anh ấy đếm bài”.

Nông dân đã quen nhìn thấy sự thay đổi công nghệ dưới dạng các công cụ mới mà họ có thể sử dụng, nhưng kiến thức này đang thay thế cách họ đưa ra quyết định. Thật vậy, bản thân các quyết định không chỉ thay đổi mà còn thay đổi về mặt vật lý. Ở đâu? Đến San Francisco, xa vùng nông thôn nước Mỹ. Tập đoàn đô thị Bờ Tây này hiện đang nói với nông dân ở Kansas rằng họ không nên trồng ngô nữa.

Tập đoàn Khí hậu hiện không giải quyết tất cả các quyết định canh tác. Người nông dân vẫn đưa ra một số quyết định quan trọng. Tuy nhiên, như Friedberg lưu ý, “theo thời gian, điều đó sẽ trở thành con số 0. Mọi thứ sẽ được quan sát. Mọi thứ sẽ được dự đoán trước”. Nông dân đang nắm bắt điều này từng chút một. Tác giả Michael Lewis kể lại, “[Không] chưa ai từng hỏi Friedberg câu hỏi: Nếu kiến thức của tôi không còn hữu dụng nữa,



thì ai cần tôi?” Nói cách khác, các điều kiện bảo hướng tới sự gián đoạn và quản lý trang trại tập trung. Chúng tôi không biết sẽ mất bao lâu và liệu một số quyết định không thể được tự động hóa hay không. Chúng tôi biết rằng ngành nhìn thấy tiềm năng cao trong các công cụ này.

Monsanto đã mua lại Tập đoàn Khí hậu vào năm 2013 với giá 1,1 tỉ đô la. Từng bước một, khi các máy dự đoán được cải thiện, người nông dân không chỉ đơn giản là lấy những dự đoán đó và đưa ra quyết định mà nhường những quyết định đó cho người khác. Điều này có thể làm cho việc quản lý trang trại trở nên tốt hơn, vì những người có thông tin, kỹ năng, động lực và khả năng phối hợp phù hợp sẽ ngày càng đưa ra nhiều quyết định hơn. Nhưng đồng thời, vai trò của người nông dân sẽ là gì? Họ là chủ đất, nhưng bao lâu nữa thì điều đó cũng thay đổi?

KẾ HOẠCH CHO CUỐN SÁCH NÀY

Mục đích của chúng tôi ở đây là thúc đẩy sự phát triển của các giải pháp hệ thống AI. Trọng tâm của chúng tôi tập trung vào quyết định và vai trò của dự đoán trong đó.

Trong Phần 1, chúng ta đã thảo luận về câu chuyện ngụ ngôn về 3 doanh nhân và giới thiệu

những thách thức trong việc phát triển và triển khai AI trong The Between Times, những thách thức này có khả năng phản ánh những thách thức của điện và các công nghệ có mục đích chung khác trong quá khứ. Là cầu nối để hiểu những thách thức và cơ hội này, trong Chương 3, chúng tôi xem lại luận điểm của cuốn sách trước đó của chúng tôi, Máy dự đoán và mô tả cách AI liên quan đến dự đoán ở cốt lõi của nó.

Trong Phần 2, để xây dựng trường hợp của chúng ta rằng không chỉ cần có các giải pháp điểm cho dự đoán để mang lại giá trị cao, chúng ta sẽ đi sâu vào quá trình ra quyết định. Chúng tôi khám phá 3 chủ đề rộng lớn. Đầu tiên, đưa ra quyết định là khó khăn. Nó liên quan đến chi phí nhận thức liên quan đến việc chỉ tuân theo một quy tắc. Ưu điểm của một quyết định là khả năng thay đổi những gì bạn làm để đáp ứng với thông tin mới. Khi không có dự đoán, những lợi ích đó sẽ bị tắt tiếng. Thứ hai, khả năng dự đoán của AI có thể nghiêng cán cân từ các quy tắc sang các quyết định và cách thức các quy tắc và hành động tiếp theo được thực hiện để bảo vệ một tổ chức khỏi những hậu quả bất lợi của chúng có thể che giấu sự không chắc chắn. Do đó, có thể khó tìm ra nơi áp dụng AI vì sự không chắc chắn bị chôn vùi. Đồng thời, đây là nơi có thể cảm nhận rõ nhất sự gián đoạn. Nếu sự không chắc chắn



đó xuất hiện, các doanh nghiệp cố gắng che giấu nó sẽ gặp nguy hiểm. Thứ 3 là mối quan hệ giữa các quyết định. Khi các quyết định tương tác với nhau, việc chuyển từ một quy tắc sang một quyết định được thúc đẩy bởi dự đoán thực sự làm tăng thêm độ không tin cậy cho hệ thống. Khắc phục điều này thường đòi hỏi sự thay đổi trên toàn hệ thống. Vấn đề là các quy tắc kết dính hệ thống hiện có lại với nhau, thường theo những cách tinh tế và không rõ ràng. Do đó, việc xây dựng một hệ thống mới từ đầu có thể dễ dàng hơn là thay đổi một hệ thống hiện có. Vì vậy, trong lịch sử, những công ty mới tham gia và công ty khởi nghiệp thường hoạt động tốt hơn các doanh nghiệp lâu đời khi cần phải thiết kế lại toàn bộ hệ thống để tối ưu hóa. Do đó, thay đổi ở cấp độ hệ thống là một con đường dẫn đến sự gián đoạn của các công ty đương nhiệm.

Trong Phần 3, chúng ta xem xét quá trình tạo ra các hệ thống mới, không chỉ bao gồm việc thay đổi một quyết định để đáp ứng với dự đoán mà còn để điều đó xảy ra đối với tất cả các quyết định tương tác với nhau. Chúng tôi mô tả giá trị của việc áp dụng tư duy hệ thống và nhìn thấy mối quan hệ tinh tế giữa các quyết định, đặc biệt nếu trước đây phần lớn bị chi phối bởi các quy tắc. Chúng tôi cho thấy rằng dự đoán AI đã có tác động thay đổi hệ thống trong quá trình đổi mới. Điều này cung cấp

một cái nhìn thoáng qua về những thay đổi có thể được yêu cầu ở nơi khác.

Trong Phần 4, chúng tôi trình bày một hệ quả chính của sự thay đổi toàn hệ thống: ý nghĩa đối với quyền lực. Sự gián đoạn là một quá trình liên quan đến việc phân phối lại quyền lực kinh tế – tức là, dưới hệ thống mới, những người tạo ra nhiều giá trị kinh tế nhất sẽ thay đổi. Chúng tôi xem lại lịch sử gần đây hơn để giải thích sự gián đoạn đã làm thay đổi ngành luôn gắn liền với sự thay đổi trên toàn hệ thống như thế nào. Sau đó, chúng tôi xem xét một khía cạnh của nỗi sợ hãi liên quan đến quyền lực nảy sinh khi thảo luận về AI: liệu máy móc có quyền lực hay không. Chúng tôi giải thích rằng khi bạn hiểu rằng AI hoàn toàn là dự đoán và là một đầu vào trong quá trình ra quyết định, quyền lực không đến từ máy móc – mặc dù chúng có vẻ mạnh mẽ – mà từ những người đứng sau máy móc, hướng dẫn cách chúng phản ứng với các dự đoán, cái mà chúng ta gọi là phán đoán. Sau đó, chúng tôi khám phá những lợi thế mà dự đoán tốt hơn – và dữ liệu thúc đẩy chúng – có thể mang lại cho các doanh nghiệp khi cạnh tranh với nhau. Nói cách khác, làm thế nào dự đoán có thể thúc đẩy sự tích lũy quyền lực.

Trong Phần 5, chúng ta sẽ tìm hiểu cơ chế mà dự đoán có thể thay đổi người nắm giữ quyền lực, đó là cách AI gây rối. Chúng tôi giải thích cách thức



áp dụng AI liên quan đến việc tách rời dự đoán và phán đoán mà trước đây được kết hợp với nhau trong cách những người ra quyết định không có máy dự đoán đưa ra quyết định. Điều này đặt ra câu hỏi liệu người ra quyết định hiện tại có thực sự ở vị trí tốt nhất để đưa ra nhận định đó hay không. Sau đó, chúng tôi chuyển sang xem những thẩm phán đó có thể đang theo dõi việc tách rời ai. Cụ thể, chúng tôi khám phá cách thức phán đoán có thể chuyển từ phi tập trung sang quy mô lớn với sự tập trung quyền lực. Tương tự, khi dự đoán liên quan đến sự thay đổi từ một quy tắc thành một quyết định và sau đó là một hệ thống mới, những người mới có vai trò trong việc ra quyết định và do đó trở thành tụ điểm quyền lực mới.

Cuối cùng, trong Phần 6, chúng tôi xem xét thiết kế hệ thống – đặc biệt là đối với các hệ thống đáng tin cậy dựa trên những phát triển AI mới – và cung cấp một công cụ mà bạn có thể sử dụng để hiểu doanh nghiệp và ngành của mình dưới dạng một hệ thống các quyết định (hoặc các quyết định tiềm năng). Điều này liên quan đến việc thực hiện một cách tiếp cận bảng trống để ánh xạ sứ mệnh của bạn thành một nhóm nhỏ các quyết định cơ bản nhất khi bạn có quyền truy cập vào các cỗ máy dự đoán mạnh mẽ. Chúng tôi giải thích cách ngành bảo hiểm gia đình có thể làm điều này. Sau đó chúng tôi

xem xét cách chăm sóc sức khỏe có thể đạt được điều này, cho rằng nó đã phải đối mặt với những thách thức ở cấp độ hệ thống từ việc ứng dụng AI.

Chúng tôi kết thúc cuốn sách với ví dụ về sự thiên vị của AI, điều mà nhiều người quan tâm. Chúng tôi lập luận rằng khi được coi là một giải pháp điểm, sự thiên vị của AI là một vấn đề và có thể tạo ra sự phản kháng chính đáng đối với việc áp dụng các máy dự đoán. Nhưng sự thiên vị được nhìn nhận đúng đắn hơn từ tư duy hệ thống. Khi chúng tôi hiểu cách hệ thống có thể điều chỉnh để phù hợp với dự đoán của AI, sẽ dễ dàng nhận thấy rằng việc loại bỏ sự thiên vị là một cơ hội do AI mang lại chứ không phải là cơ hội bị nó lật đổ.

Nhìn chung, chúng tôi nhấn mạnh rằng việc chuyển đổi ngành dựa trên AI cần có thời gian. Nó không rõ ràng làm thế nào để làm điều đó lúc đầu. Nhiều người có thể sẽ thử nghiệm và thất bại vì họ hiểu sai về nhu cầu hoặc họ không thể vận dụng kinh tế học đơn vị. Cuối cùng, ai đó sẽ thành công và thiết lập một con đường dẫn đến lợi nhuận. Những người khác sẽ cố gắng bắt chước. Người dẫn đầu ngành sẽ cố gắng tạo ra những con hào để bảo vệ lợi thế của mình. Đôi khi nó sẽ thành công. Bất kể thế nào, ngành công nghiệp này sẽ chuyển đổi và như mọi khi, sẽ có người thắng và người thua.



ĐIỂM CHÍNH

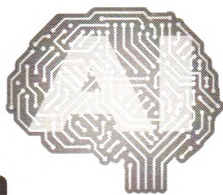
- Bất chấp khả năng dự đoán tuyệt vời của AI, tăng trưởng năng suất đo được đã giảm một nửa trong thập kỷ qua và thu nhập thực tế đã đình trệ kể từ cuối những năm 1990 đối với đa số người Mỹ. Nghịch lý năng suất này không phải là mới. Chúng tôi đã trải nghiệm điều tương tự vào những năm 1980 với máy tính. Chúng tôi gọi đây là The Between Times: sau khi chứng kiến sức mạnh của AI và trước khi nó được áp dụng rộng rãi. Mặc dù các giải pháp điểm và giải pháp ứng dụng có thể được thiết kế và triển khai một cách hợp lý một cách nhanh chóng, nhưng các giải pháp hệ thống sẽ khai phá tiềm năng to lớn của AI cần nhiều thời gian hơn.
- Khái niệm chính trong định nghĩa của 3 loại giải pháp AI – giải pháp điểm, giải pháp ứng dụng và giải pháp hệ thống – là tính độc lập. Nếu dự đoán AI tạo ra giá trị bằng cách nâng cao quyết định đầu mối và việc tạo ra giá trị đó không phụ thuộc vào bất kỳ thay đổi nào khác đối với hệ thống, thì giải pháp điểm (quyết định hiện có nâng cao) hoặc giải pháp ứng dụng (quyết định mới) đều khả thi. Tuy nhiên, nếu giá trị

của quyết định nâng cao không độc lập mà đòi hỏi những thay đổi cơ bản khác đối với hệ thống để tạo ra giá trị, thì cần phải có giải pháp hệ thống.

- Các giải pháp hệ thống thường khó triển khai hơn các giải pháp điểm hoặc giải pháp ứng dụng vì quyết định do AI tăng cường tác động đến các quyết định khác trong hệ thống. Trong khi các giải pháp điểm và giải pháp ứng dụng thường củng cố các hệ thống hiện có, các giải pháp hệ thống, theo định nghĩa, hỗ trợ các giải pháp hiện có và do đó thường dẫn đến sự gián đoạn. Tuy nhiên, trong nhiều trường hợp, các giải pháp hệ thống có khả năng tạo ra lợi nhuận tổng thể lớn nhất cho các khoản đầu tư vào AI. Hơn nữa, các giải pháp hệ thống có khả năng gây ra sự gián đoạn trong một số ngành, tạo ra kẻ thắng người thua.

Chương 3

AI LÀ CÔNG NGHỆ DỰ BÁO



Trong cuốn sách đầu tiên của chúng tôi, *AI cuộc cách mạng công nghệ trong thời đại 4.0*, chúng tôi đã tiến hành nghiên cứu các yếu tố kinh tế đơn giản của AI. Chúng tôi loại bỏ tất cả sự phức tạp tiềm ẩn và sự cường điệu liên quan đến AI và rút gọn nó thành một yếu tố duy nhất: dự báo. Đưa một chủ đề mới thú vị về đúng bản chất và ít giật gân hơn là một phần việc quan trọng của những nhà kinh tế học.

Khi nghĩ về AI, mọi người nghĩ về những cỗ máy thông minh rải rác trong nền văn hóa đại chúng. Họ nghĩ đến những robot hữu ích như R2-D2 hoặc WALL-E. Họ nghĩ về những người đồng đội xuất sắc như Data trong *Star Trek* hay J.A.R.V.I.S. từ *Iron Man*. Họ cũng nghĩ về những nhân vật trở nên bất hảo như HAL 9000 trong *2001* hay Ultron từ *The Avengers*. Dù mang những tính cách hay đặc điểm kỳ quặc hay mang theo ý định gì, những đại diện của AI này đều có một điểm chung: không ai nghĩ

ngờ thực tế họ có thể suy nghĩ, lập luận và có quyền tự quyết, giống như chúng ta.

Chúng ta có thể phát triển những công nghệ có thể thực hiện được tất cả những điều đó, nhưng đó không phải là những gì chúng ta có ngày nay. Những gì chúng ta có là sự tiến bộ trong kỹ thuật thống kê vượt quá suy nghĩ của chúng ta. Thế nhưng bước tiến của kỹ thuật thống kê này có ý nghĩa quan trọng. Khi tiến bộ đó đạt đến tiềm năng của nó, chi phí dự báo sẽ được giảm đáng kể và chúng ta có thể thấy dự báo ở khắp mọi nơi.

Cột mốc nổi bật trong những phát triển gần đây của AI là sự ra đời của deep learning, hay học sâu, chứng minh sự ưu việt của các kỹ thuật mới trong máy học. Vào năm 2012, một nhóm từ Đại học Toronto, do Geoffrey Hinton đứng đầu, đã sử dụng học sâu để cải thiện đáng kể khả năng của máy móc trong việc xác định những gì đang diễn ra trong hình ảnh. Sử dụng một bộ dữ liệu gồm hàng triệu hình ảnh được gọi là ImageNet, trong suốt hơn một thập kỷ, các nhóm đã cố gắng tạo ra các thuật toán xác định chính xác hình ảnh đang hiển thị. Bộ dữ liệu đó đã gắn nhãn cho mỗi hình ảnh với sự phân loại của con người về những gì có trong đó. Ý tưởng là lấy tập dữ liệu đó, sử dụng nó để phát triển một thuật toán và



sau đó cung cấp cho thuật toán những hình ảnh mới. Tiếp theo, sẽ có một cuộc đua giữa các thuật toán đó và với con người để xác định hình ảnh hiển thị cái gì. Con người không hoàn hảo trong nhiệm vụ này, nhưng trước năm 2012, họ vượt trội hơn nhiều so với bất kỳ thuật toán nào. Năm 2012, điều đó bắt đầu thay đổi.

Phương pháp học sâu tiếp nhận nhiệm vụ (xác định đối tượng trong hình ảnh) như một bài toán dự báo. Mục tiêu là, khi được cung cấp một hình ảnh mới, nó có thể dự đoán con người sẽ nói gì về hình ảnh đó. Khi được nhận hình ảnh một con chó con, nhiệm vụ không phải là hiểu rõ cái gì thực sự tạo nên hình ảnh đó. Thay vào đó, đó là đoán xem thứ hiển thị trong ảnh nhiều khả năng nhất sẽ là nhãn hiệu nào trong tất cả những nhãn hiệu hiện có. Do đó, mục tiêu là đoán nhãn hiệu có khả năng đúng nhất, nhãn này đã trở thành dự đoán. Bằng cách cho phép một số lượng lớn các thuộc tính và sự kết hợp của chúng được tiếp nhận và xử lý – một phép tính toán vô cùng khó, nhóm Toronto đã chứng minh được học sâu có thể vượt trội hơn bất kỳ thuật toán nào khác và cuối cùng là hầu hết con người.

Dự báo của máy hữu ích bởi nó chính xác hơn bất kỳ thứ gì khác. Lý do là dự báo là đầu vào quan trọng trong quá trình ra quyết định của chúng ta.

BỔ SUNG DỰ BÁO

Dự báo không phải là đầu vào duy nhất trong quá trình ra quyết định. Để hiểu tầm quan trọng của dự báo, cần phải hiểu hai yếu tố đầu vào quan trọng khác trong các quyết định: phán đoán và dữ liệu. Phán quyết được giải thích tốt nhất với một ví dụ.

Trong bộ phim *I, Robot*, thám tử điều tra Del Spooner sống trong một tương lai nơi người máy phục vụ con người. Vị thám tử ghét người máy, và sự căm ghét đó thúc đẩy phần lớn cốt truyện. Bộ phim cũng giải thích nguyên nhân cho sự thù hận của Spooner đối với người máy.

Trong một vụ tai nạn, xe của Spooner va chạm với một chiếc xe khác chở một bé gái 12 tuổi. Hai chiếc xe văng khỏi cây cầu, sau đó cả thám tử và bé gái đều rơi vào tình trạng sắp chết đuối. Một người máy xuất hiện và cứu thám tử, thay vì bé gái. Anh ta cho rằng lẽ ra người máy phải cứu bé gái. Sau đó, anh ta mang mối hận với người máy.

Vì đó là người máy, nên Spooner có thể dễ dàng kiểm tra các quyết định của nó. Anh ta biết rằng người máy đã dự đoán rằng anh ta có 45% cơ hội sống sót và cô gái chỉ có 11%. Do đó, nếu chỉ có thời gian để cứu một người, thì nó cứu anh ta. Spooner



nghĩ rằng 11% là quá đủ để cố gắng cứu bé gái, và nếu là con người thì anh ta sẽ biết điều đó.

Có lẽ vậy. Đó là một tuyên bố về sự phán xét – quá trình xác định phần thưởng cho một hành động cụ thể trong một môi trường cụ thể. Nếu cứu bé gái là quyết định đúng đắn, thì chúng ta có thể suy ra rằng anh ta tin rằng mạng sống của bé gái đáng giá gấp 4 lần mạng sống của anh ta. Nếu bé gái có 11% cơ hội sống sót và anh ta có 45% cơ hội, thì một người có thông tin đó khi buộc phải đưa ra lựa chọn thì anh ta phải xác định giá trị tương đối của mạng sống của từng người. Người máy rõ ràng đã được lập trình để đánh giá tất cả mạng sống của con người đều có giá trị như nhau. Khi sử dụng máy dự đoán, chúng ta cần phải rõ ràng về phán đoán.

TƯƠNG QUAN VÀ NHÂN QUẢ

Dữ liệu cung cấp thông tin cho phép dự đoán. Khi AI thu được nhiều dữ liệu chất lượng cao hơn, các dự đoán sẽ được cải thiện. Theo chất lượng, chúng tôi muốn nói rằng bạn có dữ liệu về bối cảnh mà bạn đang cố gắng dự đoán. Các nhà thống kê gọi đây là nhu cầu dự đoán điều gì đó dựa trên “sự hỗ trợ” của dữ liệu của bạn. Ngoại suy quá nhiều từ dữ liệu bạn có và dự đoán có thể không chính xác.

Dự đoán về mức độ hỗ trợ của dữ liệu của bạn không đơn giản bằng việc thu thập dữ liệu từ nhiều cài đặt khác nhau để đảm bảo bạn không ngoại suy quá nhiều hoặc tránh dự đoán quá xa về tương lai. Đôi khi dữ liệu bạn cần không tồn tại. Điều này làm cơ sở cho điệp khúc được lặp đi lặp lại trong mọi khóa học thống kê trên toàn thế giới: tương quan không nhất thiết là nhân quả.

Trong ngành công nghiệp đồ chơi của Hoa Kỳ, có một mối tương quan chặt chẽ giữa quảng cáo và doanh thu. Chi tiêu quảng cáo tăng mạnh vào cuối tháng 11 và duy trì ở mức cao trong khoảng một tháng. Trong thời gian này khi quảng cáo cao, đồ chơi bay khỏi kệ. Chỉ nhìn vào dữ liệu, có thể rất hấp dẫn để tăng quảng cáo trong thời gian còn lại của năm. Chắc chắn có thể tăng doanh thu trong tháng 4 nếu ngành này chỉ quảng cáo vào đầu mùa xuân giống như trong tháng trước lễ Giáng sinh.

Ngành công nghiệp đã không làm điều đó. Quảng cáo đồ chơi trong tháng 4 thấp hơn nhiều so với quảng cáo đồ chơi trong tháng 12. Điều đó có nghĩa là bất kỳ dự đoán nào về điều gì sẽ xảy ra nếu quảng cáo tăng trong tháng 4 đều không có dữ liệu hỗ trợ. Từ mối tương quan hằng tháng giữa quảng cáo và doanh thu, bạn không thể biết quảng cáo tạo ra doanh thu hay Giáng sinh tạo ra



cả hai. Có thể mỗi tương quan là nhân quả, vì vậy việc tăng chi tiêu quảng cáo trong tháng 4 sẽ dẫn đến doanh số bán đồ chơi tăng mạnh. Tất nhiên, cũng có thể không phải quảng cáo gây ra phần lớn doanh số bán hàng trong tháng 12. Thay vào đó, có thể chính sự chờ đón Giáng sinh là nguyên nhân thúc đẩy cả quảng cáo và bán hàng. Cũng có thể là quảng cáo gây ra doanh số bán hàng trong tháng 12, nhưng vì có rất ít người Mỹ mua đồ chơi vào tháng 4 nên sẽ không có nhiều tác dụng vào thời điểm đó trong năm.

Nói cách khác, chỉ riêng máy dự đoán sẽ không cung cấp thông tin về điều gì sẽ xảy ra với doanh số bán đồ chơi của tháng 4 nếu chiến lược quảng cáo của ngành thay đổi. Để khám phá mối quan hệ đó, bạn cần sử dụng một nhánh thống kê khác gọi là “suy luận nhân quả”. Giống như AI, nhánh này cũng đã đạt được những tiến bộ lớn trong vài năm qua (với giải Nobel Kinh tế năm 2021 được trao cho những tiến bộ trong phân tích các mối quan hệ nhân quả) và ngày càng rõ ràng rằng bản thân những công cụ này là một phần bổ sung cho AI, cung cấp cho AI dữ liệu cần thiết để cho phép dự đoán hiệu quả trong nhiều cài đặt. Các công ty AI hàng đầu trên thế giới nhận ra điều này. Ví dụ, trong số 3 người đoạt giải Nobel năm 2021,

2 người làm việc cho Amazon. Ngoài các vị trí học thuật của họ, Guido Imbens còn là một nhà khoa học trong nhóm Core AI và David Card là một Học giả của Amazon. Những thách thức suy luận nhân quả hạn chế tính hữu ích của AI ở những nơi có thể thu thập dữ liệu liên quan. AI đã rất hiệu quả khi chơi các trò chơi, bao gồm cờ vua, cờ vây và Super Mario Brothers. Cài đặt của trò chơi giống nhau mỗi khi bạn chơi, vì vậy không cần phải ngoại suy nhiều từ dữ liệu trong quá khứ sang trò chơi hiện tại. Hơn nữa, đối với các tình huống không có trong dữ liệu, vì trò chơi là phần mềm nên có thể chạy thử nghiệm mô phỏng. Những thử nghiệm này cho phép AI điền vào phần còn lại của dữ liệu, để khám phá điều gì sẽ xảy ra nếu nhấn một nút khác hoặc thử một chiến lược mới. Đó là cách mà AlphaGo và AlphaGoZero của DeepMind đã khám phá ra các chiến lược chiến thắng trong trò chơi cờ vây chưa từng được sử dụng thành công trong các trò chơi cạnh tranh cấp độ cao. DeepMind đã chạy hàng triệu thí nghiệm mô phỏng và cỗ máy đã học cách dự đoán các chiến lược chiến thắng bằng cách mô phỏng điều gì sẽ xảy ra nếu nó thử một số cách tiếp cận khác nhau.

Trong nhiều tình huống kinh doanh, dữ liệu có sẵn. Khi không có trong tay, thường có thể thu thập nó thông qua thử nghiệm. Thử nghiệm trong



kinh doanh mất nhiều thời gian hơn so với trong trò chơi vì nó diễn ra với tốc độ của con người hơn là tốc độ mà máy tính có thể chạy mô phỏng. Tuy nhiên, nó có thể là một công cụ mạnh mẽ để thu thập dữ liệu liên quan có thể là đầu vào hữu ích cho AI.

Các thí nghiệm ngẫu nhiên là công cụ chính để các nhà thống kê khám phá nguyên nhân gây ra cái gì. Chúng là tiêu chuẩn vàng để phê duyệt các phương pháp điều trị y tế mới. Một nhóm được chỉ định ngẫu nhiên để được điều trị. Người kia nhận được một giả dược. Mặc dù các nhóm không giống nhau – chúng bao gồm những người khác nhau – sự khác biệt là do ngẫu nhiên. Với đủ số người được chỉ định cho mỗi nhóm, bạn có thể kết luận liệu việc điều trị có gây ra kết quả hay không. Bằng cách chạy đúng loại thử nghiệm, thường có thể điền vào dữ liệu cần thiết để rút ra kết luận là quan hệ nhân quả chứ không chỉ là mối tương quan.

Đôi khi dữ liệu mô phỏng, ngẫu nhiên hoặc thậm chí gần như ngẫu nhiên như vậy sẽ khó hoặc không thể thu thập được. Cuối cùng là ứng dụng của AI trong bối cảnh quân sự. Lúc đầu, chiến tranh có vẻ là nơi lý tưởng để áp dụng các công cụ AI. Như nhà lý luận quân sự Carl von Clausewitz đã viết vào thế kỷ 19, “Chiến tranh

là lĩnh vực của sự không chắc chắn”. Dự đoán có thể làm giảm sự không chắc chắn và do đó tạo ra lợi thế quân sự đáng kể. Tuy nhiên, thách thức là các cuộc chiến liên quan đến kẻ thù. Trong chiến tranh, “nếu AI trở nên giỏi trong việc tối ưu hóa giải pháp cho bất kỳ vấn đề nào, thì kẻ thù thông minh sẽ có động lực để thay đổi vấn đề”. Kẻ thù sẽ vượt ra ngoài tập huấn luyện và dữ liệu thời bình sẽ ít được sử dụng.

Ý tưởng này cũng áp dụng trong bối cảnh kinh doanh. Các dự đoán sẽ hoạt động khi không có đối thủ cạnh tranh nào có động cơ làm suy yếu dự đoán của bạn hoặc khách hàng có động cơ tìm cách vượt qua chúng. Nếu khách hàng có thể làm tốt hơn bằng kỹ thuật đảo ngược các khía cạnh chính của AI của bạn và cung cấp cho nó thông tin sai lệch, thì AI sẽ chỉ phục vụ mục tiêu của bạn chừng nào khách hàng chưa khám phá ra cách thức hoạt động của nó. Khi các dự đoán không hỗ trợ dữ liệu của bạn và các vấn đề suy luận nguyên nhân phát sinh, điều có vẻ giống như một giải pháp điểm thường yêu cầu thay đổi ở cấp hệ thống. Tuy nhiên, đối với 11% công ty đã nhìn thấy giá trị từ AI, dự đoán thường dựa trên sự hỗ trợ của dữ liệu họ có trong tay, vì vậy giải pháp điểm AI vẫn hoạt động tốt.



DỰ BÁO LÀ TRỌNG TÂM

Xem xét quyết định chấp nhận hoặc từ chối một giao dịch tài chính. Chìa khóa cho quyết định này là dự báo về các gian lận, và đó cũng là trọng tâm hoạt động kinh doanh của Verafin. Ví dụ phát sinh giao dịch yêu cầu thanh toán, cụ thể là chuyển từ tài khoản này sang tài khoản khác. Nếu giao dịch được chấp thuận, tiền sẽ đổi chủ, chính điều này kích hoạt việc trao đổi hàng hóa và dịch vụ thực. Nếu giao dịch không được chấp thuận, tiền sẽ không được chuyển đi, điều này có thể cản trở các nghĩa vụ thực tế tiềm ẩn. Thực tế là tất cả các giao dịch đều yêu cầu phê duyệt thanh toán bởi nếu có sai lầm, hậu quả là vô cùng lớn. Nếu phê duyệt một giao dịch mà giao dịch đó không xuất phát từ người không sở hữu tài khoản thì một loạt các khoản nợ và vấn đề sẽ phát sinh. Nhưng nếu từ chối một giao dịch không có vấn đề như vậy và bạn làm gián đoạn các hoạt động trong thế giới thực nằm bên dưới toàn bộ quy trình.

Bạn có thể nghĩ rằng điều này sẽ khiến bạn muốn có một hệ thống có thể tránh tất cả sai lầm. Thách thức ở đây không phải là tránh những sai lầm là không thể. Với thời gian và sự xem xét cẩn thận, một ngân hàng có thể làm được điều đó.

Vấn đề đặt ra là để hoàn toàn tránh khỏi sai lầm sẽ rất tốn kém. Nó làm chậm quá trình, tăng phí giao dịch, và mặt khác sẽ lấy đi mọi tiện lợi mong muốn trong giao dịch ngay từ đầu. Xét cho cùng, nếu giao dịch bằng cách gửi tin nhắn số để phê duyệt các thay đổi trong mục nhập tài khoản sẽ tốn quá nhiều chi phí, có lẽ tốt hơn là sử dụng tiền mặt cũ và hiển thị tiền ngay lúc đó.

Thay vào đó, để hệ thống hoạt động, các ngân hàng tham gia trò chơi dự đoán. Họ phải cân bằng các lỗi đoán mà họ có thể mắc phải. Nếu quá khắt khe với những giao dịch mà họ chấp thuận, họ có nguy cơ từ chối nhiều giao dịch hợp pháp và kết thúc bằng những khách hàng không hài lòng. Nếu họ dễ dãi trong việc phê duyệt, thì cuối cùng họ sẽ tạo cơ hội cho những kẻ lừa đảo thực hiện giao dịch bất hợp pháp và khiến việc thu hồi tiền bị thất lạc trở nên khó khăn, điều này gây tổn hại trực tiếp đến lợi nhuận của ngân hàng. Vì vậy, họ đoán và thiết lập ngưỡng siêu âm để cân bằng hai lỗi mà họ chắc chắn sẽ mắc phải.

AI là phương tiện giúp các ngân hàng trở nên tốt hơn trong trò chơi đoán và giảm thiểu sai sót. Khi các nhà kinh tế học nghiên cứu những phát triển mới của AI trong thập kỷ qua, chúng tôi đã nhận thấy vai trò của mình là cắt đứt sự



cường điệu. Trí tuệ nhân tạo thu hút sự chú ý của các nhà triết học, nhà làm phim, nhà tương lai học, người tiên đoán ngày tận thế và nhiều người khác, những người có thể làm sôi nổi cuộc trò chuyện trong bữa tiệc tối của bạn. Chúng tôi đóng vai trò ngược lại. Lấy gợi ý từ nơi thực sự đạt được những thành công trong khoa học máy tính, tất cả sự phát triển của AI với những cái tên lạ mắt chẳng hạn như mạng thần kinh, học máy, học sâu hoặc tối ưu hóa đối thủ, chúng tôi tập trung vào một bước tiến – mặc dù là một bước tiến lớn – trong thống kê, cụ thể là, thống kê dự đoán. Do đó, thay vì AI tham gia vào cuộc thập tự chinh chống tội phạm, chống gian lận, những gì AI đang thực sự làm là cải thiện khả năng của các ngân hàng trong việc phân loại các giao dịch hợp pháp khỏi các giao dịch gian lận với chi phí thấp hơn nhiều – tức là khả năng dự đoán.

AI ngày nay là một cỗ máy dự đoán, và tất cả chỉ có vậy. Đối với Verafin, điều đó hóa ra lại chính xác như những gì nó mong muốn. Để làm cho hệ thống thanh toán hiện đại hoạt động đòi hỏi mức độ tự động hóa cao. Bạn muốn có sự tự tin cao trong những phê duyệt đó. Đó là nơi AI tham gia. Nó sử dụng lượng thông tin phong phú mà các ngân hàng có về khách hàng của họ, các kiểu hành vi, thời gian và địa điểm các giao dịch đang diễn ra, đồng thời

chuyển những thông tin đó thành dự đoán về việc liệu một giao dịch có hợp pháp hay không. Trong hai thập kỷ qua, độ chính xác đã được cải thiện. Hiện nay, các công cụ AI được sử dụng rộng rãi để phát hiện gian lận trong các dịch vụ ngân hàng và tài chính, và các viện này yêu cầu những lợi ích đáng kể về độ chính xác.

Dự đoán là công việc kinh doanh của Verafin và vì AI là một công nghệ tiên đoán tiến bộ vượt bậc, nên các công ty như Verafin chắc chắn sẽ là những người sớm được hưởng lợi từ nó. Các ngân hàng và các tổ chức tài chính khác được sử dụng để thực hiện các chức năng dự đoán của riêng họ. Phê duyệt những gì là kinh doanh của họ. Những quyết định được đưa ra càng tốt thì họ càng hoàn thành tốt công việc của mình. Và họ có thể sử dụng tất cả thông tin họ có thể nhận được. Hóa ra, Verafin có thể cung cấp thông tin đó bằng cách tận dụng khả năng tìm hiểu và tinh chỉnh các thuật toán của mình từ các giao dịch mà hàng nghìn tổ chức tài chính và khách hàng của họ đã xử lý. Điều này không có nghĩa là đạt được vị trí dẫn đầu thị trường như một miếng bánh ngọt. Verafin đã có gần hai thập kỷ kinh nghiệm. Nhưng vấn đề ở đây là dự đoán đã và luôn là công việc của nó, và AI đã tạo ra một cơ hội mới để nó nâng cao trò chơi của mình.



NGOÀI DỰ ĐOÁN

Cuốn sách này không nói về những công ty như Verafin, nhưng Verafin là ưu tiên hàng đầu vì nó minh họa một ngoại lệ hơn là một quy tắc cho việc áp dụng và tác động của AI. Mọi thứ rơi vào vị trí cho Verafin. Đầu tiên, dự đoán – đầu ra chính của AI – là trọng tâm hoạt động kinh doanh của công ty. Thứ hai, rất ít phải thay đổi để khách hàng của mình – các tổ chức tài chính – chấp nhận các sản phẩm của Verafin vì dự đoán cũng là trọng tâm của các doanh nghiệp đó. Thứ ba, những doanh nghiệp đó đã đưa ra quyết định dựa trên dự đoán, họ biết phải làm gì với những dự đoán đó và họ đã quen với việc giải quyết hậu quả của các lỗi dự đoán để AI có thể được triển khai một cách an toàn. Các ngân hàng đã chín muồi để đổi mới giải pháp điểm.

Điểm mấu chốt là Verafin đã hoạt động trong một hệ thống được thiết kế để áp dụng AI. Hệ thống không phải thay đổi để sử dụng dự đoán. Một cách mới để đưa ra quyết định không cần phải được tạo ra. Verafin đã sẵn sàng cung cấp dự đoán cho các doanh nghiệp biết họ cần dự đoán để làm gì, đã tự thiết lập để tận dụng những dự đoán đó và quan trọng nhất là có thể tự định hướng theo các hướng khác nhau dựa trên những dự đoán đó.

Verafin là một ví dụ về bước cuối cùng trong một quy trình sẽ khó khăn hơn đối với hầu hết các doanh nghiệp có thể hưởng lợi từ việc áp dụng AI ngay bây giờ và trong tương lai. Nếu doanh nghiệp của bạn muốn áp dụng AI, nó có thể sẽ phải dọn sạch bụi rậm và có thể là cả một khu rừng trước khi có thể triển khai nó. Cuốn sách này nói về quá trình thanh toán bù trừ đó – xác định những gì cần thay đổi cũng như những tình huống khó xử và thách thức mà bạn sẽ phải đối mặt khi thực hiện thay đổi đó. Chúng tôi đang đề cập đến sự thay đổi ở cấp độ hệ thống, chứ không phải là giải pháp điểm hoặc giải pháp ứng dụng mà bạn có thể triển khai trong khi vẫn giữ nguyên hệ thống hiện có. Biết bạn đang làm gì là một bước quan trọng để tìm hiểu xem liệu nó có xứng đáng hay không.

MINH HỌA THỬ THÁCH

Một trong những phần được tham khảo nhiều nhất trong cuốn sách trước đây của chúng tôi đến từ một thí nghiệm về suy nghĩ. Amazon sử dụng AI để dự đoán những gì người tiêu dùng cụ thể có thể muốn mua. Khi bạn mua sắm trên trang web của Amazon, những dự đoán đó cho biết quyết định nên đề xuất mặt hàng nào từ danh mục hàng chục triệu lựa chọn. Bạn sắp xếp các đề xuất của nó và



đặt hàng một số; những thứ đó được gửi đến cho bạn. Đó là một vài ngày giữa thời gian bạn bắt đầu mua sắm và sự xuất hiện của hàng hóa.

Trong bối cảnh này, chúng tôi tự hỏi, điều gì có thể thay đổi nếu dự đoán của Amazon về thứ bạn muốn mua tốt hơn nhiều? Amazon có thể muốn trả lại cho bạn vài ngày đó bằng cách dự đoán những gì bạn muốn, vận chuyển nó cho bạn và mời bạn chấp nhận hoặc không nhận nó tại cửa của bạn. Nói cách khác, Amazon giao hàng cho bạn dựa trên dự đoán của nó và sau đó bạn mua sắm từ các hộp được giao đến trước cửa nhà bạn. Chúng tôi gọi đây là bước chuyển từ shop-then-ship (mua-rồi-giao hàng) sang ship-then-shop (giao hàng-rồi-mua). Mặc dù một số người có thể cảm thấy rùng rợn khi có các sản phẩm xuất hiện trong nhà của họ, nhưng không khó để tưởng tượng nó có thể tiện lợi như thế nào. Chuyển sang giao hàng sau đó mua sắm là điều mà chúng tôi dự tính là một giải pháp ứng dụng. Nó sẽ đưa ra dự đoán và đưa cho Amazon quyết định có vận chuyển thứ gì đó hay không, thay vì để khách hàng quyết định có vận chuyển mặt hàng hay không. Nhiều người coi việc mua sắm là một gánh nặng, do đó, bằng cách có thể cung cấp điều đó với giá rẻ, khả năng dự đoán tốt hơn đã đưa ra giải pháp cho ứng dụng đó. Chúng tôi vẫn chưa thấy Amazon

làm điều này. Điều đó đã không ngăn nó đặt chân xuống nước. Nó đã được cấp bằng sáng chế cho ý tưởng “chuyên hàng trước”, nhưng việc triển khai nó có phần bị tắt tiếng. Ví dụ, nó thường cung cấp cho người tiêu dùng các tùy chọn đăng ký mua một sản phẩm thay vì chủ động đặt hàng. Làm sao? Nó thông báo lượng giấy vệ sinh mà gia đình bạn sử dụng và hứa sẽ ngừng cung cấp sản phẩm đều đặn. Điều này mang lại cho Amazon sự chắc chắn về nhu cầu; nó chuyển khoản tiết kiệm được cho người tiêu dùng dưới dạng giảm giá khi đăng ký. Tuy nhiên, khi bạn vượt ra ngoài thử nghiệm tưởng tượng, bạn có thể thấy tại sao việc triển khai giao hàng rồi mua sắm là một thách thức đáng kể. Nếu dự đoán là hoàn hảo, đây không phải là một giải pháp ứng dụng khó khăn. Nhưng những dự đoán không hoàn hảo và có thể không bao giờ đúng như vậy. Để làm được điều đó, Amazon cần có cách thu thập các sản phẩm mà bạn chọn từ chối. Đó là những sản phẩm đủ cứng để giao một cách an toàn, chứ đừng nói đến việc mang chúng ra công để trả lại. Lợi nhuận cũng có thể là một nỗi đau cho người tiêu dùng. Vì vậy, nếu không có một hệ thống trả lại hàng gần như miễn phí, mô hình giao hàng-rồi-cửa hàng của Amazon khó có thể xuất hiện từ công xuất phát. Thật vậy,



Amazon đã phải vật lộn với việc trả lại hàng đến mức họ không bao giờ bán lại nhiều mặt hàng bị trả lại mà gửi thẳng chúng vào thùng rác. Với hệ thống hiện có của mình, việc Amazon vứt bỏ hàng trả lại sẽ rẻ hơn là đưa những sản phẩm đó trở lại hệ thống hậu cần của riêng mình. Bài học ở đây là vận chuyển-rời-mua sắm, trong khi nó có vẻ là một giải pháp ứng dụng, là một giải pháp đòi hỏi những thay đổi ở những nơi khác trong hệ thống để mang lại hiệu quả kinh tế. Mặc dù chúng tôi không đánh giá cao điều đó khi viết *AI trong cuộc cách mạng công nghệ 4.0*, vận chuyển-rời-mua sắm là một giải pháp hệ thống vì nó tác động đến các quyết định quan trọng khác và yêu cầu thiết kế hệ thống của Amazon để tạo điều kiện thuận lợi cho cách xử lý hàng trả lại tiết kiệm chi phí hơn nhiều.

GIỜ THÌ SAO?

“Được rồi, bây giờ thì sao?” Đó là câu hỏi mà nhiều doanh nghiệp và tổ chức bắt đầu triển khai công nghệ AI đã hỏi chúng tôi. Các công ty này đã nghe thấy sự cường điệu về AI và sử dụng cảm nang chúng tôi đặt ra trong *AI trong cuộc cách mạng công nghệ 4.0*, họ đã bắt đầu hành trình AI của mình. Họ đã thành lập các nhóm để xem xét các nhiệm vụ của mình và xác định các cơ hội để

tận dụng những gì AI đã cung cấp: dự đoán. Dự đoán là quá trình bạn chuyển đổi thông tin bạn có thành thông tin bạn cần. Và như chúng tôi đã ghi lại trong cuốn sách đó, những đổi mới gần đây về AI đều là những tiến bộ trong việc đưa ra dự đoán tốt hơn, nhanh hơn và rẻ hơn.

Kết quả của những tiến bộ đó phổ biến đến mức chúng ta thậm chí không nghĩ về chúng nữa. Điện thoại của bạn được tích hợp nhiều công nghệ AI. Khi bạn mở khóa điện thoại, nó sẽ dễ dàng nhận ra khuôn mặt của bạn. Bạn thậm chí không cảm thấy như thể điện thoại nằm sau một hàng rào an toàn cho phép bạn và chỉ bạn đi qua cửa trước. Bạn đến một màn hình nơi các ứng dụng được sắp xếp cho bạn dựa trên dự đoán của điện thoại về những gì bạn có thể muốn vào thời điểm chính xác đó. Bạn đang ở gần quán cà phê yêu thích của mình và muốn đặt hàng? Bạn đang ở trong xe của bạn và cần chỉ đường? Điện thoại biết, nhưng có vẻ như nó chỉ tiện lợi. Vấn đề là tất cả những kết quả thấp cho việc sử dụng AI prediction đều đã được chọn. Câu hỏi mà các doanh nghiệp đặt ra là “Có phải nó không?”

Cuốn sách này là câu trả lời cho câu hỏi đó: không. Mặc dù AI dường như ở khắp mọi nơi, giống như nhiều công nghệ đột phá khác trước đây,



nhưng nó chỉ mới bắt đầu. Các cuộc cách mạng công nghệ quan trọng như điện, động cơ đốt trong và chất bán dẫn đều bắt đầu chậm chạp, mất nhiều thập kỷ để đạt được bước tiến của chúng. Dự đoán của AI sẽ không có gì khác biệt, bất chấp sự cường điệu rằng nó đại diện cho một số mô hình tăng tốc thay đổi công nghệ khác biệt.

Thay vì ngồi trên một chuyến tàu lượn siêu tốc, nơi chúng ta đang ở trong những thế lực bất chợt nằm ngoài tầm kiểm soát của mình, chúng ta đang ngồi trên bờ vực của cơ hội – trong The Between Times. Những người và doanh nghiệp có thể tìm thấy câu trả lời cho câu hỏi “now what?” sẽ thiết lập đường dẫn cho AI.

Là những nhà kinh tế học, chúng tôi trông chờ vào các lực lượng kinh tế để hướng dẫn chúng tôi trả lời những câu hỏi như vậy. Tuy nhiên, chúng tôi vượt ra ngoài kinh tế học đơn giản, nơi các quan sát về chi phí dự đoán giảm cho thấy nhiều ứng dụng dự đoán hơn. Thay vào đó, chúng tôi đi sâu vào thực tế rõ ràng là cách mọi người và doanh nghiệp quyết định nên làm gì.

Nên làm gì không phải là một bài tập kỳ diệu để nhanh chóng đi đến câu trả lời tối ưu mà thay vào đó, liên quan đến việc cân nhắc, quy trình và

chi phí của chính nó. Để tận dụng lợi thế của dự đoán, bạn cần xem xét việc sử dụng dự đoán đó để làm gì và thực tế là những người ra quyết định đã làm mà không có dự đoán đó trước đây. Khi bạn không có thứ gì đó, bạn không nên bỏ cuộc. Bạn bồi thường cho nó. Nếu bạn không có thông tin cần thiết để đưa ra lựa chọn sáng suốt, bạn sẽ tự bảo vệ mình khỏi những hậu quả của việc phải làm mọi việc một cách mù quáng. Do đó, khi dự đoán AI xuất hiện, không nên ngạc nhiên khi các cơ hội sử dụng nó không rõ ràng ngay lập tức. Những người sắp ra quyết định đã xây dựng một giàn giáo dựa trên việc không có thông tin đó.

Tất cả điều này có nghĩa là việc xác định điều gì sẽ xảy ra tiếp theo sẽ đòi hỏi phải xem xét kỹ lưỡng không chỉ những gì dự đoán có thể làm mà còn cả những bức tường đã được dựng lên để ngăn cản việc đặt ra câu hỏi đó. Chúng tôi sẽ giải cấu trúc các quyết định để cung cấp bộ công cụ giúp bạn nhìn xa hơn những cơ hội rõ ràng về dự đoán AI đến những cơ hội không rõ ràng nhưng có tiềm năng quan trọng hơn mà dự đoán AI có thể mang lại.



NHỮNG ĐIỂM CHÍNH

- Những tiến bộ gần đây trong AI đã làm giảm chi phí dự đoán. Chúng tôi sử dụng dự đoán để lấy thông tin chúng tôi có (ví dụ: dữ liệu về việc liệu các giao dịch tài chính trong quá khứ có gian lận hay không) và tạo dữ liệu chúng tôi cần nhưng không có (ví dụ: liệu giao dịch tài chính hiện tại có gian lận hay không). Dự đoán là một đầu vào để ra quyết định. Khi chi phí của một đầu vào giảm, chúng ta sẽ sử dụng nhiều hơn. Vì vậy, khi dự đoán trở nên rẻ hơn, chúng tôi sẽ sử dụng nhiều AI hơn. Khi chi phí dự đoán giảm, giá trị của các dự đoán thay thế cho máy móc (ví dụ: dự đoán của con người) sẽ giảm. Đồng thời, giá trị của các phần bổ sung cho dự đoán của máy sẽ tăng lên. Hai trong số những bổ sung chính cho dự đoán máy là dữ liệu và phán đoán. Chúng tôi sử dụng dữ liệu để đào tạo các mô hình AI. Chúng tôi sử dụng phán đoán cùng với dự đoán để đưa ra quyết định. Trong khi dự đoán là một biểu hiện của khả năng xảy ra, phán đoán là một biểu hiện của mong muốn – những gì chúng ta muốn. Vì vậy, khi đưa ra quyết định, chúng ta dự tính khả năng xảy ra của từng kết quả có thể xảy ra từ quyết

định đó (dự đoán) và mức độ chúng ta đánh giá từng kết quả (phán đoán).

- Có lẽ sự lạm dụng lớn nhất đối với các dự đoán của AI là coi các mối tương quan mà chúng xác định là nhân quả. Thông thường, các mối tương quan là đủ tốt cho một ứng dụng. Tuy nhiên, nếu chúng ta cần AI cho mối quan hệ phi chính thức, thì chúng ta sẽ sử dụng các thử nghiệm ngẫu nhiên để thu thập dữ liệu liên quan. Những thí nghiệm này là công cụ tốt nhất để các nhà thống kê khám phá nguyên nhân gây ra cái gì.
- Trong *AI trong cuộc cách mạng công nghệ 4.0*, chúng tôi đã giới thiệu một thử nghiệm tương tự liên quan đến công cụ đề xuất của Amazon. Chúng tôi tưởng tượng những gì sẽ xảy ra ngày càng chính xác. Lúc đầu, công cụ này thực hiện công việc tốt hơn là giới thiệu các mặt hàng cho khách hàng. Tuy nhiên, tại một số điểm, nó vượt qua ngưỡng mà nó trở nên tốt đến mức những người ở Amazon có thể hỏi: “Nếu chúng tôi rất giỏi trong việc dự đoán những gì khách hàng muốn, thì tại sao chúng tôi lại đợi họ đặt hàng? Chúng ta hãy vận chuyển nó. Mặc dù Amazon đã đăng ký bằng sáng chế trong lĩnh vực “vận chuyển dự



kiến”, nhưng họ vẫn chưa áp dụng mô hình kinh doanh mới này. Tại sao không? Giải pháp điểm bán đầu – AI cung cấp đề xuất tốt hơn trên nền tảng hiện có – tận dụng nguyên trạng hệ thống Amazon. Mô hình mới sẽ yêu cầu Amazon thiết kế lại hệ thống của mình, đặc biệt là về cách nó xử lý hàng trả lại. Hiện tại, hệ thống trả lại hàng của họ rất tốn kém nên việc vứt bỏ các mặt hàng bị trả lại thường tiết kiệm hơn là tích trữ chúng để bán cho các khách hàng khác. Ngưỡng trong thử nghiệm suy nghĩ của chúng tôi yêu cầu chuyển từ giải pháp điểm sang giải pháp hệ thống. Trong *AI trong cuộc cách mạng công nghệ 4.0*, chúng tôi đã đánh giá thấp sự khác biệt.

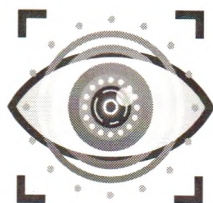
PHẦN 2

CÁC QUY TẮC



Chương 4

QUYẾT ĐỊNH HAY KHÔNG QUYẾT ĐỊNH



Bạn có muốn biết bí mật này không? Các nhà kinh tế không thực sự tin rằng con người lý trí một cách hoàn hảo. Bạn biết đấy, có một bức tranh biếm họa thế này. Trong tranh là một đơn vị kế toán đang cẩn thận đặt tất cả các lựa chọn trước mắt – hàng triệu lựa chọn về thời gian và không gian. Các tính toán cho thấy chính xác mục tiêu của họ là gì, dù là lợi nhuận, niềm hạnh phúc hay thứ gì khác. Sau đó, họ đưa ra lựa chọn và đính kèm theo kế hoạch hành động. Sự tính toán logic hoàn hảo đó thường là những gì được mô tả trong các mô hình của các nhà kinh tế học, nếu bạn hiểu theo nghĩa đen. Và các nhà kinh tế thực sự coi trọng những dự báo mà các mô hình đó đưa ra. Thế nhưng, từ kinh nghiệm của bản thân, họ biết rằng con người thực tế không đạt được sự hoàn hảo về logic như trong bức tranh đó. Các nhà kinh tế học sẽ tròn mắt khi được gắn với nhận xét “các nhà kinh tế học tin rằng con

người lý trí một cách hoàn hảo”. Họ không tin như vậy. Thật là phi lý nếu tin vào điều đó.

Tuy nhiên, việc nhìn nhận con người như thế chúng ta tính toán, nhất quán và hành động theo một nhóm lợi ích là tiền đề hữu ích để hiểu được hành vi của hàng nghìn hoặc hàng triệu người. Bạn có muốn biết liệu đánh thuế lên thuốc lá có làm giảm hút thuốc hay không? Một tác động là nếu chi phí cho một việc gì đó tăng hơn, mọi người sẽ giảm bớt nó. Tuy vậy, mức độ cắt giảm là bao nhiêu và liệu mức độ đó có đủ lớn hay không lại là một vấn đề khác. Bạn sẽ phải hiểu rõ lịch sử loài người, sự căng thẳng, các nhóm xã hội và những phương thức tiếp thị mà các công ty thuốc lá triển khai. Thế nhưng, xuất phát điểm tuyệt vời cho nhiều ngành khoa học xã hội là sự thừa nhận con người đưa ra quyết định một cách có chủ ý.

Mọi người đưa ra quyết định về việc sẽ mặc gì hằng ngày. Steve Jobs nổi tiếng với chiếc áo cổ lọ màu đen đặc trưng và quần jean, bất kể dịp nào hay thời tiết ra sao. Mark Zuckerberg vẫn diện quần jean nhưng chọn áo thun xám. Barack Obama, với tư cách là tổng thống, chỉ mặc bộ vest xám hoặc xanh lam. Ông giải thích lý do tại sao với Michael Lewis của cuốn sách *Vanity Fair*:

Ông nói: “Bạn sẽ thấy tôi chỉ mặc những bộ đồ màu xám hoặc xanh lam. Tôi đang cố



gắng giảm bớt các quyết định. Tôi không muốn phải cân nhắc quyết định về việc ăn gì hay mặc gì. Bởi vì tôi còn rất nhiều quyết định khác phải đưa ra”. Ông đề cập đến kết quả nghiên cứu cho thấy chính hành động đưa ra quyết định sẽ làm suy giảm khả năng đưa ra quyết định tiếp theo của một người. Đó là lý do tại sao mua sắm rất mệt mỏi. “Bạn cần tập trung năng lượng ra quyết định của mình. Bạn phải tạo thành thói quen cho chính mình. Bạn không thể trải qua cả ngày bị phân tâm bởi những chuyện vặt vãnh được”.

Một trong số chúng tôi (Joshua) đã từng mua hết nguồn cung cấp toàn cầu cho một đôi giày mà anh ấy thích (tổng cộng là 6 đôi, nếu bạn muốn biết) vì anh ấy không muốn phải đối mặt với việc mua sắm giày trong một thập kỷ nữa. Mục đích của tất cả những lựa chọn này chính xác là để tránh phải đưa ra quyết định. Khi mọi người hình thành thói quen hoặc tuân theo các quy tắc, họ đang thừa nhận rằng chi phí cho những nỗ lực tối ưu hóa là quá cao. Vì vậy, họ, thực tế, quyết định sẽ không quyết định. Điều này xảy ra ở khắp mọi nơi. Hãy nghĩ về chính mình và bạn nhận ra rằng hầu hết những gì bạn quyết định không phải là những quyết định thực tế mà là những quyết định

tiềm ẩn, tức là những điều bạn có thể chọn nhưng lựa chọn không chọn.

Điều này đặt ra một thách thức đáng kể cho một cuốn sách về AI. Dự báo bằng AI chỉ hữu ích nếu bạn đang đưa ra quyết định. Nhưng nó còn nhiều hơn thế. Thông thường, cách chúng ta xây dựng các hệ thống gồm các thành phần phụ thuộc lẫn nhau chính là đầu tư vào độ tin cậy. Chắc chắn bạn sẽ không mong muốn trong một hệ thống có một bộ phận làm điều gì đó mà các bộ phận khác không mong đợi hoặc đoán trước. Bạn muốn sự tin cậy. Quy tắc là cách sự tin cậy được hình thành và đảm bảo trong các hệ thống. Tuy nhiên, nếu dự báo bằng AI sẽ phá vỡ các quy tắc và biến chúng thành các quyết định, thì một hệ quả sẽ là phát sinh sự thiếu tin cậy đối với các hệ thống hiện có. Cũng vì thế, có thể sẽ không đáng để sử dụng AI trừ phi bạn có thể thiết kế lại hệ thống để phù hợp với các quyết định mà AI đang kích hoạt.

Đó là lý do tại sao chúng tôi sẽ bắt đầu với những quyết định mà chúng ta đã quyết định không đưa ra. Chúng tôi muốn biết vì sao chúng ta làm như vậy, với mục tiêu đánh giá xem việc áp dụng AI có thể thay đổi suy nghĩ của chúng ta và biến những quyết định tiềm ẩn đó thành quyết định thực tế hay không. Như bạn sẽ thấy trong chương này, chúng tôi tin rằng AI có thể, với những lợi ích tiềm năng to



lớn và ý nghĩa đối với cách các tổ chức có thể phải điều chỉnh để đáp ứng.

THIẾT LẬP NÓ VÀ BỎ QUÊN NÓ

Việc không đưa ra quyết định luôn dễ dàng hơn việc phải quyết định một việc cụ thể. Cụ thể hơn, chúng ta luôn có xu hướng tránh phải thu thập thông tin, xử lý chúng, cân nhắc các phương án và sau đó đi đến quyết định. Nền kinh tế vận hành chủ yếu dựa trên tiền đề này, trong đó người khác là những quyết định được phân bổ ngay cả khi chúng ta không hoàn toàn tin tưởng họ sẽ làm những gì chúng ta sẽ làm.

Không ai hiểu điều này hơn Herbert Simon, người đồng thời đoạt giải Nobel Kinh tế cho công trình nghiên cứu về đề tài tính hợp lý có giới hạn và Giải thưởng Turing cho nghiên cứu tiên phong của ông trong lĩnh vực AI. Theo quan sát của ông trong quá trình làm việc tại bộ phận công viên Milwaukee, ngân sách cho các hoạt động không được phân bổ một cách tối ưu; tức là mọi người đã không tối ưu hóa theo cách mà các nhà kinh tế học lập mô hình với sự xuất hiện của máy tính. Vào những năm 1950, khi Simon cố gắng lập trình máy tính trở thành những người ra quyết định thông minh, ông đã nhận thấy cái giá phải trả của việc tối ưu hóa. Bị buộc phải sử dụng các tài nguyên máy tính hạn chế, mọi người sẽ

làm điều mà Simon phải làm với những chiếc máy tính sơ khai vào thời của ông: “make do”.

“Make do”, một thuật ngữ nghệ thuật mà Simon cảm thấy “hài lòng”, không phải là biến cái hoàn hảo thành kẻ thù của cái tốt. Thay vì tìm kiếm các giải pháp mà họ biết là tốt hơn, họ sẽ thực hiện các hành động đủ tốt. Thay vì đối phó với một môi trường phức tạp, mọi người sẽ thu hẹp phạm vi lựa chọn được xem xét. Thay vì liên tục cập nhật các lựa chọn của họ dựa trên thông tin mới nhận được, họ sẽ áp dụng các quy tắc và thói quen không hòa nhập với thông tin mới và do đó, cho phép họ hoàn toàn bỏ qua thông tin.

Tuy nhiên, chỉ cần lưu ý rằng đôi khi mọi người mặc định tuân theo các quy tắc hơn là đưa ra quyết định. Chúng ta cần hiểu khi nào các quyết định được đưa ra. Điều gì quyết định khi nào một vấn đề cụ thể sẽ được đặt trong quy tắc mặc định thay vì trong cơ chế quyết định hiện tại?

HỆ QUẢ THẤP

Hai yếu tố chính tác động đến các quyết định bao gồm: mức độ hậu quả (tác động) và chi phí thông tin. Chúng ta sẽ nói về thông tin ở phần sau. Trước hết, hãy xem xét hậu quả của một quyết định. Trong triết học có quan điểm cho rằng chúng ta không nên yêu cầu cao ở bản thân khi kết quả



đạt được không nhiều. Trong câu chuyện ngụ ngôn thâm thúy của nhà triết học người Pháp Jean Buridan, con lừa khi được đặt giữa đồng cỏ khô và một thùng nước sẽ muốn cái nào ở gần nó hơn. Đối mặt với việc không thể làm đứt sợi dây và đưa ra lựa chọn, con lừa chết đói. Ta có thể tưởng tượng tình trạng nan giải tương tự khiến máy tính bị khóa trong một vòng lặp. Tuy nhiên, trong khuôn khổ mục đích của cuốn sách này, vấn đề là thời gian dành cho việc ra quyết định không nên tương xứng với hệ quả của nó.

Hãy quay trở lại với các quy tắc về trang phục của Jobs, Zuckerberg, Obama được hình thành để hạn chế tải trọng nhận thức của họ. Điều mà mỗi người nhận ra là kết quả đem lại từ việc chọn trang phục này thay vì trang phục khác là không lớn. Đứng trước một tủ quần áo với một loạt các lựa chọn, họ sẽ buộc phải đối mặt với những lựa chọn không mấy quan trọng hằng ngày. Họ có thể chỉ cần nhắm mắt và lấy món đồ đầu tiên họ chạm vào. Nhưng họ không tin tưởng chính mình. Vì vậy, họ hạn chế một cách có ý thức các lựa chọn của mình.

Đối với hầu hết chúng ta, hệ quả không quá thấp. Chắc chắn, Jobs và Zuckerberg có thể xuất hiện ở nơi làm việc trong bất kỳ trang phục nào họ muốn. Chắc chắn, Obama phải mặc vest hầu

hết mọi ngày, và không ai thực sự quan tâm đến màu sắc của nó. Thế nhưng, phần lớn chúng ta không có được sự xa xỉ đó. Bạn có thực sự nhìn qua toàn bộ tủ quần áo của mình mỗi ngày hay phân loại những bộ mà bạn thường xuyên mặc không? Ngẫm lại xem, không phải nhiều người trong chúng ta giới hạn các lựa chọn của mình để có thể đưa ra các lựa chọn dễ dàng hơn sao? Cuối cùng, chúng ta cố gắng làm cho những điều này ít để lại hậu quả hơn để phù hợp với mong muốn giảm bớt sự phức tạp trong các quyết định.

Ví dụ về lựa chọn quần áo trở nên nổi tiếng như một quyết định có hệ quả thấp và tải trọng nhận thức tương đối cao nếu bạn cố gắng tối ưu hóa. Tuy nhiên, cả hệ quả và tải trọng nhận thức đi đôi với nhau. Bạn hãy tưởng tượng về việc cam kết với một người bạn đời hoặc có một đứa con. Hậu quả của việc lựa chọn sai là rất lớn, vì vậy cần phải dành thời gian và công sức để cân nhắc. Do đó, nếu chúng ta xem các quyết định tiềm ẩn là những thứ mà chúng ta cần dành thời gian và công sức để cân nhắc trước khi bắt đầu thực hiện chúng thay vì gạt chúng sang một bên bằng cách trì hoãn hoặc chọn một quy tắc mặc định, vì những hậu quả quan trọng hơn được dự báo trước, thì chúng ta sẽ chủ động tham gia hơn là lựa chọn không quyết định.



THÔNG TIN ĐẮT ĐỎ

Yếu tố thứ hai ảnh hưởng đến việc liệu bạn có chủ động đưa ra quyết định hay không là liệu bạn có thông tin hay cụ thể là chi phí của thông tin mà bạn cần để đưa ra quyết định. Thông tin tốn kém có thể khiến một quyết định trông giống như một quyết định có hậu quả thấp và do đó khiến bạn áp dụng các quy tắc mặc định thay vì cân nhắc kỹ lưỡng.

Bạn có nên mang theo một chiếc ô ngày hôm nay? Mặc dù không làm rung chuyển trái đất về mặt lợi ích của nó đối với người khác, nhưng lựa chọn này có thể mang lại những hậu quả lớn. Nếu bạn chọn không mang ô và bị mắc mưa, đó là một ngày tồi tệ. Bạn có thể đảm bảo điều này không xảy ra nhờ mang theo một chiếc ô, nhưng điều đó cũng có chi phí riêng của nó. Tất nhiên, nếu có thông tin chính xác (cụ thể là liệu trời có mưa hay không và có khả năng bạn sẽ dính mưa hay không), bạn sẽ mang theo một chiếc ô nếu khả năng đó rất cao và bỏ nó ở nhà nếu không. Nhưng nếu xác suất là 50-50 thì sao?

Để hiểu rõ bản chất của xác suất như tung đồng xu, hãy giả sử nếu bị ướt, bạn sẽ mất 10 đô la chi phí cá nhân, nhưng nếu mang theo ô và trời không mưa, bạn cũng phải chịu 10 đô la từ một trở ngại không cần thiết. Xét về chi phí dự tính, dù theo cách nào thì bạn cũng có thể phải chịu $10 \text{ đô la} \times 50\%$,

tức 5 đô la. Điều này khiến bạn thờ ơ với việc phải mang và không mang.

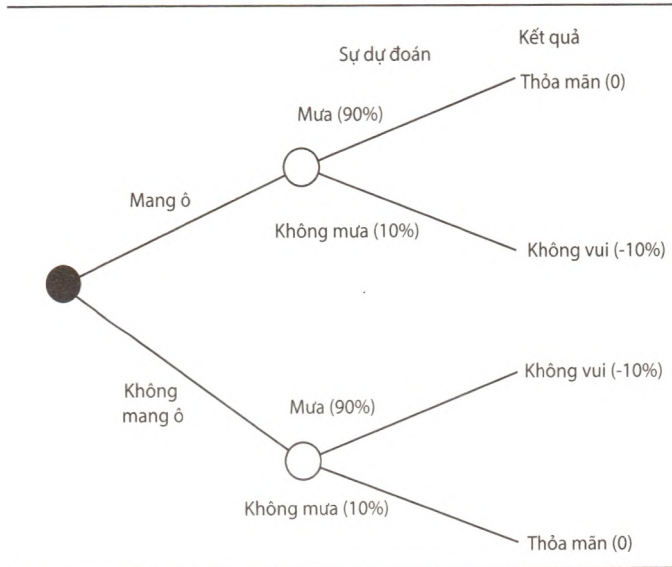
Trước khi đứng ở cửa dẫn đo giải quyết tất cả những điều này, bạn có thể xem dự báo. Nếu dự báo thời tiết cho biết khả năng có mưa lớn hơn 50%, thì bạn mang ô vào ngày hôm đó, còn nếu ít hơn 50% thì bạn không cần. Tuy nhiên, ở đây chúng tôi đã đơn giản hóa vấn đề và loại bỏ ngữ cảnh có thể khiến thông tin đó trở nên không đủ. Nếu dự báo thời tiết nói rằng trời sẽ khô ráo với độ tin cậy 90%, thì mọi thứ đã rõ ràng, nhưng dự báo thời tiết không phải lúc nào cũng rõ ràng như thế. Chúng ta hiếm khi có thông tin mang sắc thái để quyết định nếu khả năng có mưa là 40% hoặc 30%, có khác gì với khả năng 50%. Hơn nữa, vào thời điểm bạn tìm ra tất cả những điều đó, chẳng hạn như bằng cách xem xét các thông tin về sức gió hoặc áp suất khí quyển, bạn lại rơi vào tình huống mà chi phí nhận thức của việc đưa ra quyết định này vượt quá những gì bạn có thể thu được từ quyết định đó.

Chúng ta có thể biểu diễn quyết định này dưới dạng cây quyết định, một phần chính của các khóa học MBA về kinh tế học và phân tích quyết định. Ý tưởng của cây quyết định là các nhánh của cây đại diện cho các lựa chọn. Ví dụ, trong Hình 4-1, lựa chọn (tại nút liền màu đen) là cầm theo hay để lại chiếc ô. Quyết định này được đưa ra trong các



điều kiện không chắc chắn, và kết quả của sự không chắc chắn đó cũng được thể hiện bằng các nhánh trời mưa hoặc quang đãng (tại nút tròn “chọn”). Nếu bạn không có bất kỳ dự đoán nào, xác suất liên quan đến cả hai nhánh đó có thể là 50%. Tuy nhiên, ở đây có một dự báo nói rằng khả năng có mưa là 90%. Các mẹo cuối cùng của các nhánh cuối cùng là kết quả. Gắn liền với mỗi kết quả (có bốn kết quả: mang theo + mưa, mang theo + trời quang, để lại + mưa và để lại + trời quang) là sự đánh đổi mà chúng tôi đã mô tả là giá trị tiền tệ tương đương của từng kết quả.

Hình 4-1: Cây quyết định



Ở đây chúng tôi trình bày những kết quả này như chi phí bỏ ra nếu những điều tồi tệ xảy ra. Mức độ được đánh giá và quyết định bởi người lựa chọn (trong trường hợp này là bạn). Đây là lý do tại sao chúng tôi đề cập đến mức độ này như đánh giá (judgment). Đánh giá là một khái niệm quan trọng được nhắc lại xuyên suốt cuốn sách này. Đặc biệt, người đưa ra đánh giá ở nhiều khía cạnh sẽ kiểm soát quyết định, và hơn nữa, điều mà các máy (hay hệ thống) dự báo cho phép là dự báo được tách rời khỏi đánh giá, vì nếu không có máy, người ra quyết định thường làm cả hai việc cùng nhau. Ở đây, chúng tôi đã chọn làm cho chi phí của các kết quả xấu tương đương nhau, nghĩa là cả hai đều là 10 đô la. Điều này có nghĩa là với dự báo về mưa này, nếu mang theo ô, bạn dự báo chi phí là 1 đô la và nếu không, bạn dự báo chi phí là 9 đô la. Một người thông thường sẽ chọn lấy mang theo ô khi có dự báo đó.

Điều mà nhiều người làm là không đưa ra quyết định, đặc biệt nếu họ không có sẵn dự báo. Ví dụ, bạn có thể linh hoạt lựa chọn và hành động trong ngày của mình, theo đó bạn có thể điều chỉnh thời gian ra ngoài nếu trời có vẻ sắp mưa. Trong trường hợp này, quy tắc mặc định của bạn có thể là không bao giờ mang ô. Ngoài ra, bạn có thể mua một chiếc ô nhỏ có thể gấp gọn, khá đắt tiền vì chúng không



quá bền, nhưng chúng dễ mang theo hơn. Sau đó, quy tắc mặc định của bạn sẽ là luôn mang theo một chiếc ô và thậm chí không bao giờ nghĩ đến việc trời có mưa hay không.

Vấn đề ở đây là, thay vì thu thập thông tin để đưa ra lựa chọn tối ưu, khi làm điều đó tốn kém, chúng ta chọn những thói quen hoặc quy tắc để loại bỏ nhu cầu phải xem xét thông tin. Chúng ta chỉ làm điều tương tự mỗi lần mà không cần phải suy nghĩ về nó.

CÁC QUYẾT ĐỊNH BAO GỒM NHỮNG GÌ

Nếu bạn đang kiểm chế việc đưa ra quyết định và tuân theo các quy tắc, thoát nhìn thì có vẻ như dự báo của AI sẽ không hữu ích cho bạn. Công việc của AI là cung cấp thông tin để ra quyết định. Nếu bạn không đưa ra quyết định, thì thông tin đó không có giá trị.

Chức năng của AI là cung cấp dự báo tốt hơn, về cơ bản có nghĩa là bạn có thông tin cần thiết để đưa ra quyết định tốt hơn. Quyết định mang theo ô được kích hoạt nếu chúng ta có dự báo thời tiết mà chúng ta có thể tin tưởng thay vì dự báo mà chúng ta không quá chắc chắn. Với thông tin tốt hơn, bạn có thể loại bỏ các quy tắc của mình và cảm thấy đáng để dành thời gian đưa ra quyết định thực tế.

Trong ví dụ về chiếc ô, thông tin tuyệt vời sẽ giải phóng bạn khỏi các chi phí. Cho dù đó là chi phí khi bị ướt hay chi phí phát sinh khi mang theo một cách không cần thiết, nếu biết điều gì sẽ xảy ra, bạn sẽ không phải chịu bất kỳ chi phí nào trong đó. Để được như vậy, bạn sẽ phải từ bỏ thói quen của mình, chẳng hạn như luôn mang theo ô. Thay vào đó, bạn cần kiểm tra một ứng dụng dự báo thời tiết – kho lưu trữ các máy dự báo về lượng mưa – rồi quyết định có nên mang ô hay không. Bạn cũng có thể hình dung rằng chỉ cần ứng dụng gợi ý bạn nên cầm theo ô và làm theo mà thậm chí không cần nghĩ về lý do tại sao. Điều đó vẫn được tính là chuyển từ một quy tắc (mặc định) sang một quyết định, mặc dù các nhà lập trình của ứng dụng đã xác định các ngưỡng cho việc nên mang theo ô. Chúng ta có thể chưa hiểu được điều đó trong trường hợp mang ô, nhưng một minh họa khác là danh sách phát bài hát được đề xuất hoặc các tin tức phương tiện truyền thông được cung cấp. Luôn có một quyết định được đưa ra ở đây. Lựa chọn làm theo các đề xuất có nghĩa là không dựa vào các quy tắc (ví dụ: đọc báo từ trang đầu đến trang cuối) mà cho phép đưa ra quyết định.

Bạn có thể nhận được những điều giá trị từ việc chuyển sang ra quyết định dựa trên dự đoán của AI. Chúng ta có một ý niệm mơ hồ về điều đó từ



“thử nghiệm bắt buộc” trước đây, một điều quen thuộc với bất kỳ ai phải làm việc tại nhà trong thời kỳ Covid-19. Trước khi dịch bệnh bùng nổ, chúng ta không thực sự biết mình có thể làm việc hiệu quả như thế nào khi ở nhà, nhưng việc buộc phải làm điều gì đó khác biệt đã cho chúng ta bài học đó. Chúng ta học được điều gì đó mới mẻ đến từ việc phá vỡ những thói quen trong quá khứ của mình. Nếu sau dịch Covid-19, chúng ta không trở lại làm việc như trước đây, điều này cho thấy rằng quyết định về nơi làm việc hiện có ý nghĩa với chúng ta.

Một trường hợp tương tự khác là cuộc đình công kéo dài 2 ngày năm 2014, ảnh hưởng đến hoạt động của mạng lưới đường sắt London. Hơn 60% các nhà ga phải đóng cửa, do đó thói quen đi lại của nhiều người cũng thay đổi. Trong tình trạng đó, đối với hầu hết mọi người, lựa chọn thuận tiện nhất tiếp theo là nhà ga cách họ với khoảng cách gần như tương đương. Ngoài ra, bởi vì ở nước Anh, trời liên tục mưa vào thời điểm đó. Thời tiết không thuận tiện để đi bộ hoặc sử dụng xe đạp. Một nghiên cứu về sự kiện này cho thấy rằng, mặc dù thời gian đình công diễn ra trong thời gian ngắn, nhưng hơn 5% hành khách đã thay đổi phương thức đi lại sau sự kiện đó. Những người có nhiều khả năng thay đổi nhất là những người mà bản đồ chui với thông tin bị bóp méo

nhieu nhất, nghĩa là, dường như khoảng cách giữa các trạm khác với những gì được thể hiện trên bản đồ đúng. Ước tính cho thấy, những người thay đổi phương tiện đi lại tiết kiệm được hơn 6 phút mỗi ngày. Với thời gian đi làm trung bình khoảng 30 phút hoặc lâu hơn, tiết kiệm được 20% thời gian, lượng thời gian bạn có thể sử dụng để quyết định nên mặc gì.

Nghiên cứu này đã cho thấy các quy tắc không thích ứng và không tối ưu. Thế nhưng, đến khi cuộc chiến giá xăng bán lẻ kéo dài 3 tuần nổ ra ở Perth, Úc, vào tháng 5 năm 2015, tính công khai và biến động của giá cả trở nên rõ ràng. Điểm thú vị là một nền tảng (và sau đó là ứng dụng) hiển thị giá tại các trạm xăng khác nhau trong thành phố từ năm 2001. Cuộc chiến giá cả ngay lập tức nâng cao giá trị của việc sử dụng ứng dụng. Các nhà nghiên cứu xác định ứng dụng này đã được sử dụng nhiều hơn 70% cả trong năm xảy ra khủng hoảng giá và cả năm sau đó. Trên thực tế, cuộc chiến giá trở thành động lực đủ lớn, buộc mọi người thay đổi thói quen trước đây là không tìm kiếm mức giá thấp nhất và bắt đầu làm như vậy như một phần trong quá trình ra quyết định của họ.

Vấn đề là khi tuân theo một quy tắc, bạn có thể không nhận thức được giá trị của việc thu thập



thông tin và đưa ra quyết định. Những ví dụ này cung cấp bằng chứng cho thấy có những lợi ích tiềm ẩn và chưa được khai thác trong quá trình ra quyết định. Như vậy, chúng ta có thể dự đoán rằng một số hình thức dự báo của AI có thể mở khóa những khả năng đó theo cách tương tự.

ĐẦU TƯ VÀO VIỆC KHÔNG ĐƯA RA QUYẾT ĐỊNH

Jobs, Zuckerberg hay Obama có thực sự né tránh các quyết định về thói quen ăn mặc hằng ngày của họ không? Có, nếu bạn quan sát theo đơn vị ngày, nhưng không phải nếu bạn nhìn vào toàn bộ bức tranh. Nếu định mặc cùng một kiểu trang phục mỗi ngày, tốt nhất bạn nên chọn thứ gì đó phù hợp với mục đích đó. Bạn không thể chọn trang phục không thoải mái trong các điều kiện khác nhau hoặc không phù hợp với nhiều tình huống. Không dễ gì tìm được kiểu trang phục đó. Xác suất là mỗi người phải dành thời gian đáng kể để đi đến lựa chọn cuối cùng.

Nhìn nhận dưới khía cạnh này, các quy tắc thực tế không phải là thiếu đi phần đưa ra quyết định mà thay vào đó, là đưa ra quyết định phủ đầu. Chúng ta làm điều này mọi lúc khi tiến hành lập kế hoạch. Rất ít người đi du lịch mà không đặt trước

chỗ ở, chọn chuyến bay khứ hồi và nỗ lực đóng gói hành lý. Những người thường xuyên đi du lịch cắt giảm nhận thức liên quan đến việc xếp đồ bằng cách chuẩn bị sẵn các vật dụng (đồ vệ sinh cá nhân và bộ sạc) mà họ chỉ sử dụng khi đi du lịch. Đây chính là bài tập quản lý thời gian và tần suất quyết định điều đó, nhờ vậy, cho phép bạn tiết kiệm thời gian bằng cách dành thời gian trước.

Khi bạn đầu tư cho việc không ra quyết định, những thói quen hình thành sau đó sẽ rất khó để phá vỡ. Nếu những thói quen đó cho hiệu quả, bạn thậm chí không ý thức rằng bản thân có một thói quen và nó có thể được cải thiện nhờ các quyết định. Khi đó, nếu bạn đang đầu tư phát triển AI với giá trị là thúc đẩy ra các quyết định mà hiện tại chưa được đưa ra, thì khó khăn mà bạn phải đối mặt là làm thế nào để giành được sự chấp nhận.

Những khoản khoản đầu tư mà các nhân vật của công chúng thực hiện liên quan đến thói quen ăn mặc của họ trở nên nhạt nhòa khi so sánh với những khoản đầu tư mà hầu hết các doanh nghiệp và tổ chức dành ra cho việc không đưa ra quyết định. Mặc dù duy trì danh tiếng cho điều ngược lại, hầu hết các tổ chức không phải là những cỗ máy ra quyết định. Trọng tâm của hệ thống đó là các quy trình vận hành tiêu chuẩn (hay SOP). Đó là



những tài liệu mô tả chi tiết các quy trình để thực hiện mọi việc trong một tổ chức. Rõ ràng, những quy trình này khác nhau giữa các doanh nghiệp, nhưng không doanh nghiệp nào hoạt động mà không có chúng.

Một SOP có thể tiết kiệm khi không phải lật lại quy trình để ra quyết định, và nhờ vậy nó đóng vai trò như một khoản đầu tư để cắt giảm nhu cầu nhận thức tương tự như đối với các lựa chọn cá nhân mà chúng tôi đã mô tả trước đó. Tuy nhiên, các SOP này còn mang lại một lợi ích khác: sự tin cậy. Khi mọi người trong một tổ chức tuân theo các quy tắc, họ đang làm những việc giúp người khác làm việc của họ dễ dàng hơn mà không cần phải tham gia vào các hoạt động giao tiếp tốn kém chẳng hạn như các cuộc họp.

Trong ngành xây dựng, toàn bộ quy trình thường được phân nhỏ thành các nhiệm vụ đơn giản hơn. Kế hoạch thi công xây dựng sẽ bao gồm danh sách từng dòng và từng ngày cho từng nhiệm vụ cần hoàn thành và theo thứ tự cụ thể. Kết quả của những nhiệm vụ đó cũng được lên kế hoạch trước. Không ai trên công trường phải lo nghĩ về những việc khác ngoài nhiệm vụ được giao. Vai trò duy nhất của họ khi hoàn thành công việc là báo cáo rằng họ đã hoàn thành, đánh dấu, và sau đó

chuyển sang nhiệm vụ mới. Vẫn phát sinh những trường hợp ngoại lệ yêu cầu thay đổi và xem xét, nhưng phần lớn, mọi thứ đều diễn ra theo kế hoạch. Mỗi người làm phần của mình và ghi chú khi họ đã hoàn thành nhiệm vụ được giao.

Những quy tắc như vậy tạo ra sự tin cậy, giảm sự không chắc chắn và nhu cầu làm bất cứ điều gì để phối hợp các nhiệm vụ. Trên thực tế, các quyết định được đưa ra trước và sau đó đưa vào kế hoạch. Tuy nhiên, bản thân kế hoạch có nghĩa là bất kỳ sự thay đổi định hướng nào cũng sẽ rất tốn kém. Nếu phát sinh một vấn đề nhỏ, mọi thứ có thể tiến hành. Tuy nhiên, một vấn đề lớn phát sinh có thể làm hỏng toàn bộ kế hoạch. Và một bộ SOP cố định có thể khiến bạn khó thay đổi và thích ứng. Như chúng ta sẽ thấy ở phần sau của cuốn sách, nếu bạn muốn đưa AI vào sự kết hợp của hệ thống các quy tắc tinh vi này, đó là một thách thức rất lớn. Mục đích của AI là cho phép và thúc đẩy đưa ra quyết định, nhưng khi quyết định được đưa ra, việc phối hợp trở nên khó khăn.

NHỮNG QUYẾT ĐỊNH MỚI

Dự báo bằng AI có thể đóng vai trò vô cùng quan trọng, cung cấp thông tin cần thiết cho quá trình đưa ra quyết định. Khi đó quá trình này sẽ dự trên các dự báo thay vì bám vào quy tắc.



Quyết định mới thay thế quy tắc cũ. Thế nhưng, các quy tắc cũ không chỉ tồn tại một cách cô lập. Thay vào đó, “tòa tầng và giàn giáo” được dựng lên để tách biệt các quy tắc đó khỏi sự không “rung lắc” vẫn còn tồn tại. Có những doanh nghiệp và ngành công nghiệp toàn bộ nguồn lực cho sự tách biệt đó, do đó cũng không nhìn thấy cơ hội cho các quyết định trước mắt. Thách thức đặt ra là nhận thức được điều đó, tìm kiếm những quyết định tiềm ẩn và đưa ra những quyết định mới có thể thay thế các quy tắc hiện có. Chúng ta sẽ xem xét thách thức đó tiếp theo.

NHỮNG ĐIỂM CHÍNH

- Quy tắc là những quyết định mà chúng ta đưa ra trước. Đưa ra quyết định, không giống như tuân theo quy tắc, cho phép chúng ta tính đến thông tin có sẵn tại thời điểm và địa điểm đưa ra quyết định. Do đó, các hành động bắt nguồn từ các quyết định thường tốt hơn các hành động bắt nguồn từ các quy tắc vì chúng có thể phản ứng với tình huống. Vì vậy, tại sao chúng ta lại sử dụng các quy tắc thay vì đưa ra quyết định? Các quyết định phải chịu chi phí nhận thức cao hơn. Khi nào chi phí đáng giá? Khi hậu quả là đáng kể và khi chi phí thông tin là nhỏ.

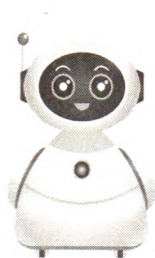
Giới thiệu AI không thay đổi kết quả, nhưng nó làm giảm chi phí thông tin.

- Sự đánh đổi giữa các quy tắc và quá trình ra quyết định là rất quan trọng trong bối cảnh của các hệ thống AI vì lợi ích chính của AI là tăng cường quá trình ra quyết định. AI cung cấp ít giá trị cho các quy tắc. AI tạo ra dự đoán và dự đoán là đầu vào thông tin chính để ra quyết định. Vì vậy, khi AI trở nên mạnh mẽ hơn, chúng sẽ giảm chi phí thông tin (dự đoán) và tăng lợi nhuận tương đối cho quá trình ra quyết định so với việc sử dụng các quy tắc. Do đó, những tiến bộ trong AI sẽ giải phóng một số quá trình ra quyết định khỏi việc tuân theo quy tắc.
- Tuy nhiên, các quy tắc không chỉ giảm chi phí nhận thức mà còn cho phép độ tin cậy cao hơn. Một quyết định thường ảnh hưởng đến những quyết định khác. Trong bối cảnh một hệ thống có các quyết định phụ thuộc lẫn nhau, độ tin cậy có thể rất quan trọng. Ví dụ: hầu hết các tổ chức đều dựa trên các quy trình vận hành tiêu chuẩn (SOP), đó là các quy tắc. SOP giảm tải nhận thức và nâng cao độ tin cậy. Nếu bạn định sử dụng khả năng dự đoán của AI để biến các quy tắc thành quyết định, thì bạn có thể cần phải thiết kế lại hệ thống để giải quyết tình trạng giảm độ tin cậy.

Chương 5

SỰ KHÔNG CHẮC CHẮN

GIẤU KÍN



Nhà kinh tế học George Stigler từng nhận xét: “Nếu chưa bao giờ bị lỡ chuyến bay, có nghĩa là bạn đang dành quá nhiều thời gian ở sân bay”. Ông ấy đã nói như vậy từ nhiều thập kỷ trước, nhưng liệu ngày nay ông ấy có đưa ra tuyên bố tương tự không?

Các kiến trúc sư thiết kế Nhà ga số 2 mới của Sân bay quốc tế Incheon Hàn Quốc hy vọng là không. Đến sân bay sớm và bạn có nhiều việc để làm hơn là chỉ chờ đợi. Bạn có thể đến spa, đánh bạc, tham gia triển lãm nghệ thuật, xem biểu diễn khiêu vũ hoặc trượt băng. Bạn cũng có thể mua sắm thỏa thích, ăn một bữa hoặc chợp mắt một chút tại “khu vực NAP”. Đối với các nhà ga sân bay mới, đây không phải là ngoại lệ mà là tiêu chuẩn. Singapore gần đây đã lắp đặt một khu vườn với thác nước năm tầng. Doha xây dựng một bể bơi và trung tâm giải trí dành cho trẻ em. Vancouver có một thủy cung. Amsterdam thường xuyên giới thiệu các bộ sưu tập

nghệ thuật được chuyển đến từ các bảo tàng nổi tiếng của thành phố.

Đối với kiến trúc sư của Incheon, mục tiêu là biến sân bay thành “điểm đến du lịch”:

Đối với thế hệ sân bay, các nhà ga hành khách không chỉ là một điểm tập trung. Chúng tôi đang nhận ra một thực tế mới: để yên tâm, hành khách đến sớm và dành nhiều thời gian hơn ở nhà ga, vì vậy địa điểm này trở nên quan trọng đối với tăng trưởng doanh thu và danh tiếng của các sân bay, đồng thời mở ra những tiềm năng mới cho sân bay. Với nhận thức này, các sân bay ngày càng xem nhà ga là điểm đến du lịch, nơi hành khách cũng chi tiền.

Và quả thật như vậy:

Tom Theobald, kiến trúc sư và hiệu trưởng tại Fentress Architects, chuyên về thiết kế sân bay, cho biết: “Hành khách đang ở trong sân bay lâu hơn 1 giờ so với họ đã làm chỉ một thập kỷ trước đó. Ông lưu ý rằng, ngay cả khi việc di chuyển hàng không đã thay đổi đáng kể, thì thời gian tăng thêm đó thường được dành cho các sân bay được xây dựng vào những năm 60 và 70.



Thế nhưng điều gì đến trước? Các sân bay hiện chỉ được thiết kế như một “điểm đến”. Mặc dù các sân bay vẫn như trước đây, mọi người vẫn đang dành nhiều thời gian hơn ở đó. Đó là sự lựa chọn. Tại sao? Bởi vì việc di chuyển hàng không ngày càng bao gồm nhiều yếu tố không chắc chắn. Giao thông, bãi đậu xe và hệ thống kiểm tra an ninh. Các hãng bay tính phí thay đổi chuyến bay, hay đặt chỗ quá nhiều, vấn đề kết nối. Việc lên chuyến bay đúng giờ trở nên khó khăn hơn và hậu quả của việc muộn giờ tăng lên. Ngay cả khi trong sân bay không có cả một sân golf 9 lỗ như ở sân bay Bangkok), bạn có thể vẫn muốn đến sớm hơn một chút và đọc một cuốn sách. Tuy nhiên, với mỗi tiện nghi mới, bạn quên mất lý do tại sao bạn đến sân bay sớm hơn 1 giờ so với giờ bay. Nó trở thành quy tắc mới của bạn.

Hãy suy nghĩ về tình trạng kỳ lạ này. Kể từ năm 1992, Incheon đã chi 10 tỉ đô la để mở rộng sân bay. Phần lớn chi phí đó là để xây dựng và mở rộng các nhà ga hành khách với thiết kế đầy thách thức bên ngoài các chốt kiểm tra an ninh. Thế nhưng, hãy xem lại tuyên bố sứ mệnh của sân bay Incheon. Mục tiêu của nó là “đảm bảo vận chuyển hàng không thông suốt”. Bạn sẽ không tìm thấy một sân bay nào tuyên bố sẽ làm bất cứ điều gì khác ngoài việc đưa hành khách qua đó. Tuy nhiên, người thiết kế

sân bay đang nghĩ cách giữ chân hành khách ở lại sân bay. Hơn nữa, khoảng 40% doanh thu của sân bay hiện nay đến từ các loại phí phi hàng không, thành phần lớn nhất trong số đó là tiền thuê mặt bằng từ các nhà bán lẻ. Các nhà thiết kế đang làm công việc của họ, đó là đảm bảo các sân bay có thể tạo ra nhiều doanh thu hơn, đồng thời khiến hành khách không phải đắn đo về thời gian họ dành thêm tại các sân bay.

Các sân bay hiện đại là biểu tượng cho cái mà chúng ta gọi là “sự không chắc chắn tiềm ẩn”. Khi mọi người không có thông tin họ cần để đưa ra quyết định tối ưu – ví dụ, về thời điểm xuất phát đến sân bay – họ sẽ áp dụng các quy tắc. Những thay đổi trong di chuyển hàng không cũng như trong quá trình di chuyển đến và rời sân bay thúc đẩy bạn muốn lựa chọn các quy tắc và các quy tắc này khiến bạn phải chờ ở sân bay lâu hơn. Các sân bay biết rằng, nếu chờ đợi là điều khó chịu thì việc đi lại cũng không hề dễ chịu, kết quả là bạn sẽ ít di chuyển hơn. Vì vậy, khi thực hiện các khoản đầu tư lớn vào cơ sở hạ tầng, họ không chỉ nghĩ đến việc giúp mọi người thuận tiện “trung chuyển” tại sân bay. Họ đang nghĩ về việc làm cho việc chờ đợi trở nên dễ chịu hơn và trong quá trình đó, họ kiếm được nhiều tiền hơn.



Một khi đến sân bay sớm hơn, bạn sẽ sẵn sàng trả tiền cho một bữa ăn hoặc các hoạt động khác giống như cách mà bóng ngô, thậm chí là những chiếc hộp đất tiền, trở nên hấp dẫn hơn khi bạn chuẩn bị đi xem phim. Nếu bạn không cảm thấy đó là một quy tắc đắt đỏ và hiếm khi lỡ chuyến bay, bạn sẽ không suy ngẫm hay đắn đo về các quy tắc và thói quen của mình. Sự không chắc chắn lùi về hậu trường trong khi hậu quả của toàn bộ hệ thống lựa chọn đó có thể được nhìn thấy dưới dạng các tòa nhà mới lấp lánh với thác nước năm tầng tuyệt đẹp.

Để tìm kiếm cơ hội sử dụng AI để đưa ra các quyết định mới, chương trước đã chỉ ra rằng, bạn nên xem xét kỹ các quy tắc và xem liệu chúng có thể được chuyển thành các quyết định tận dụng AI thay vì chấp nhận sự không chắc chắn hay không. Trong chương này, chúng tôi sẽ chỉ ra rằng, không chỉ đơn giản là bản thân các quy tắc đại diện cho mục tiêu cơ hội cho các quyết định mới do AI hỗ trợ mà còn là các tầng nhà và giàn giáo được xây dựng để che giấu sự không chắc chắn dẫn đến cảm giác lãng phí và kém hiệu quả trong các quy tắc chúng ta đã thông qua. Chúng không chỉ là dấu hiệu cho thấy có cơ hội cho AI; chúng cũng đại diện cho tầm quan trọng của cơ hội đó. Thật vậy, đối với các sân bay, một vài ứng dụng rất đơn giản

của AI có thể là mối đe dọa đối với tất cả những gì chúng hiện đang đại diện.

VŨ TRỤ SÂN BAY THAY THẾ

Trước khi xem xét mối đe dọa mà dự báo AI có thể gây ra cho các sân bay, giống như mọi thứ khác, có một hệ thống thay thế có thể cho chúng ta thấy mặt còn lại sẽ như thế nào. Một ví dụ là vũ trụ của giới siêu giàu. Họ không bay thương mại và vì vậy không có cơ hội tiếp xúc với các nhà ga sân bay công cộng cũ hoặc mới được thiết kế. Thay vào đó, họ bay riêng và đi qua các nhà ga riêng. Thông thường, những nhà hàng hào nhoáng, xa xỉ, đẹp đẽ và những phòng trưng bày nghệ thuật sẽ là nơi mà những người vô cùng giàu có xuất hiện. Thế nhưng, trong thế giới của sân bay, nhà ga tư nhân lại đơn giản một cách tích cực.

Lý do không có sự đầu tư nào nhằm cải thiện các nhà ga tư nhân trở nên tốt hơn là bởi sự không chắc chắn đang ảnh hưởng đến phần còn lại của thế giới không ảnh hưởng đến những người giàu. Với chuyến bay thương mại, bạn bị ràng buộc vào một lịch trình và sẽ bị bỏ lại nếu đến muộn. Với máy bay tư nhân, lịch trình linh hoạt hơn hoặc thậm chí không có bất kỳ lịch trình cố định nào. Nếu hành khách không ở đó, máy bay sẽ không rời đi cho đến khi họ đến nơi. Nếu hành khách có mặt sớm hơn,



máy bay sẽ rời đi ngay lập tức. Toàn bộ hệ thống được thiết kế để không phải chờ đợi – ít nhất là về phía hành khách. Không phải chờ đợi có nghĩa là không cần đầu tư để việc chờ đợi trở nên thú vị hơn. Đồng thời, người giàu không có quy tắc về thời điểm họ cần khởi hành đến sân bay. Họ xuất phát khi họ muốn. Nếu nhiều người hơn có thể có trải nghiệm tương tự, thì chắc chắn nhà ga tối ưu sẽ thiên về sự đơn giản hơn xa hoa.

Tuy nhiên, bạn không cần phải vô cùng giàu có để nhìn thấy một mặt khác này. Thay vào đó, chỉ cần so sánh với thế giới ở phía bên kia, cổng đến, với thế giới của hành khách khởi hành. Khi khu vực khách đến tách riêng với khu vực khách đi, chúng đơn giản hơn rất nhiều. Bạn có thể tìm thấy một số cửa hàng thức ăn nhẹ, nhưng mọi thứ khác được thiết kế để giúp hành khách nhanh chóng rời khỏi sân bay. Vấn đề quan trọng là khoảng cách giữa xe taxi và bãi đậu xe, mặc dù bạn có thể không quá vội. Bạn có nhớ bất kỳ chi tiết nào về sảnh đến sân bay thông thường không, ngoài làm thế nào để ra ngoài một cách nhanh nhất?

MỐI ĐE DỌA SÂN BAY AI

Các sân bay không xa lạ gì với AI. Kiểm soát không lưu đã áp dụng các hệ thống dựa trên AI để có thể đưa ra các dự báo tốt hơn về chuyến bay đến

và tình trạng tắc nghẽn. Tại sân bay Eindhoven, một hệ thống xử lý hành lý bằng AI mới đang được thử nghiệm, theo đó hành khách chỉ cần chụp ảnh túi của họ, thả xuống và nhận chúng tại điểm đến – không cần dán nhãn. Vẫn tuân thủ các yêu cầu về quyền riêng tư, hãng hy vọng sẽ làm điều tương tự với hành khách. Tất cả những thay đổi này sẽ giúp hành khách có thể nhanh chóng lên máy bay.

Tuy nhiên, không có thay đổi nào trong số này đánh vào những nguyên nhân chính dẫn đến sự không chắc chắn trong quá trình di chuyển đến chuyến bay của hành khách – giao thông và an ninh. Tuy nhiên, thay đổi đã có ở đây liên quan đến giao thông. Các ứng dụng điều hướng như Waze cung cấp cho người dùng thông tin tình trạng giao thông hiện tại và có thể ước tính hợp lý thời gian cần thiết để di chuyển đến sân bay tùy theo từng thời điểm cụ thể trong ngày. Các ứng dụng không hoàn hảo, nhưng chúng ngày càng tốt hơn.

Các ứng dụng giúp hành khách không còn phụ thuộc vào các quy tắc nhắc nhở họ cần khởi hành sớm đến sân bay. Thay vào đó, họ có thể thêm thời gian của chuyến bay đó vào lịch trình và một ứng dụng sẽ cho họ biết thời điểm tốt nhất để khởi hành và sắp xếp thời gian phù hợp. Tốt hơn nữa, trong tương lai gần, sự không chắc chắn



về thời gian thực tế mà một chuyến bay khởi hành cũng sẽ được nghiên cứu và tối ưu. Thay vì chỉ cho bạn biết thời điểm bạn cần khởi hành dựa trên thời gian bay theo lịch, ứng dụng sẽ cho bạn biết thời điểm rời nhà tùy thuộc vào thời điểm chuyến bay khởi hành thực tế được dự đoán. Chúng tôi nhắc lại lần nữa rằng vẫn còn sự không chắc chắn, nhưng đã có bước nhảy vọt từ không có thông tin sang có thông tin chính xác hơn và nhờ vậy giúp tiết kiệm hàng giờ chờ đợi. Tương tự như vậy, nhiều tài xế Uber trước đây nghĩ rằng họ sẽ không quan tâm đến ước tính thời gian xe sẽ đến nơi. Giờ đây, họ đánh giá thông tin này là một trong những tính năng có giá trị nhất của dịch vụ. Uber sử dụng AI để đưa ra dự báo đó.

AI cũng có thể dự báo thời gian chờ ở đường kiểm tra an ninh. Kết hợp các yếu tố với nhau, bạn có thể sử dụng AI để quyết định khi nào nên xuất phát đến sân bay thay vì dựa vào các quy tắc mặc định. Cũng giống như với mọi thứ khác, sẽ có một số vượt trội hơn ở khả năng này so với một số khác. Tại Incheon và nhiều sân bay khác, việc chờ đợi không còn là điều tồi tệ nữa, vì vậy có lẽ bạn không cần phải đưa ra quyết định sáng suốt.

Những người phát triển ứng dụng điều hướng hoặc công cụ dự báo thời gian khởi hành dựa trên

AI không nhận lợi ích trực tiếp từ doanh thu của các hoạt động tại sân bay trong nhà ga. Tuy nhiên, giá trị của các ứng dụng AI mà họ phát triển phụ thuộc rất nhiều vào việc có bao nhiêu người không muốn phải chờ đợi ở sân bay. Do đó, nếu thời gian đợi ở sân bay hiện ít tốn kém hơn, thì giá trị của các ứng dụng đó sẽ giảm đi.

Dự đoán dòng thời gian kiểm tra an ninh là một vấn đề khác. Các sân bay tuyên bố rằng, họ muốn cải thiện thời gian kiểm tra an ninh và giảm bớt sự không chắc chắn. Nhưng với tư cách là các nhà kinh tế học, chúng tôi không nghĩ rằng các sáng kiến của họ phù hợp với hành khách. Cải thiện thời gian kiểm tra an ninh đồng nghĩa với việc nhiều thời gian hơn để khu vực chờ sau khi hoàn thành thủ tục. Đồng thời, nó cũng làm giảm sự không chắc chắn và khiến hành khách thất vọng thời gian đến sân bay. Kết hợp với AI để giải quyết những sự không chắc chắn khác cho hành khách khi đến nhà ga, liệu các sân bay có muốn loại bỏ sự không chắc chắn trong tầm kiểm soát của chính họ?

CÁC QUY TẮC THÍCH ỨNG

Góc nhìn rộng hơn của chúng tôi không phải về sân bay mà về các quy tắc. Các quy tắc hình thành bởi vì việc chấp nhận sự không chắc chắn sẽ kéo



theo những chi phí tốn kém. Thế nhưng, các quy tắc cũng tạo ra một loạt vấn đề.

Theo Nguyên tắc Shirky của nhà văn công nghệ Clay Shirky, “các tổ chức sẽ cố gắng duy trì vấn đề mà họ chính là giải pháp”. Điều tương tự đối với các doanh nghiệp. Nếu doanh nghiệp của bạn đang cung cấp phương thức giúp hành khách khi họ đợi máy bay, thì liệu bạn có khả năng bạn sẽ làm gì đó để họ không phải đợi máy bay nữa không?

Nếu muốn tìm kiếm cơ hội bằng cách tạo ra các quyết định mới được AI thúc đẩy, bạn cần nhìn xa hơn các hàng rào bảo vệ các quy tắc khỏi hậu quả của sự không chắc chắn và hướng đến các hoạt động giúp gánh chịu các chi phí đó dễ dàng hơn hoặc giảm khả năng xảy ra các kết quả xấu.

Chúng ta có thể thấy điều này trong các biện pháp bảo vệ lâu đời mà người nông dân ở Anh sử dụng – xây hàng rào (hedgerow). Hàng rào được hình thành từ các loại cây khỏe mạnh như một bức tường chắn giữa các cánh đồng. Hàng rào sẽ là công cụ cực kỳ hữu ích nếu nông trại của bạn nuôi rất nhiều động vật và bạn không muốn phải thuê thêm người để đảm bảo vật nuôi của bạn không đi lang thang. Hàng rào cũng giúp ích ngăn lượng mưa lớn làm xói mòn đất hoặc ngăn gió lớn làm ngã cây. Với nhiều ý nghĩa bảo vệ trước các sự kiện rủi ro, chúng ta không ngạc nhiên khi thông lệ này là nguồn

gốc của thuật ngữ “phòng ngừa rủi ro” (hedging), tương tự như bảo hiểm với nghĩa rộng hơn.

Thế nhưng, bản thân các hàng rào cũng kèm theo chi phí. Khi xây dựng hàng rào, đất canh tác bị chia tách, do đó người nông dân không thể sử dụng một số kỹ thuật canh tác (kể cả cơ giới hóa) vốn chỉ hiệu quả đối với những vùng đất rộng lớn. Sau Thế chiến II, chính phủ Anh đã trợ cấp để người nông dân loại bỏ hàng rào, mặc dù trong một số trường hợp, việc loại bỏ là không nên bởi vai trò phòng ngừa rủi ro của chúng. Ngày nay, phong trào khôi phục hàng rào nổi lên, với người khởi xướng nổi bật nhất là Hoàng tử xứ Wales.

Trong nhiều tình huống, các khoản đầu tư tốn kém được thực hiện để phòng bị hoặc bảo vệ cho người sẽ ra quyết định khỏi các rủi ro tiềm ẩn. Hàng dặm đường cao tốc được rào bằng lan can để ngăn ô tô mất lái lao xuống hoặc đi vào dòng xe ngược chiều. May mắn thay, hầu hết đều không bao giờ được sử dụng, nhưng mỗi hàng rào đều là điều kiện để con đường đủ an toàn, do khả năng dễ mắc sai lầm của người lái xe.

Một cách khái quát hơn, các quy tắc trong xây dựng được hình thành để chỉ định chính xác các biện pháp khác nhau để bảo vệ những người bên trong tòa nhà khỏi các sự kiện không chắc chắn. Các sự kiện đó bao gồm hỏa hoạn nhưng cũng có thiệt



hại do thời tiết, nền móng yếu và các hiện tượng tự nhiên khác như động đất.

Điểm chung của các biện pháp bảo vệ này là chúng thường tạo ra thứ giống như các giải pháp được thiết kế quá mức. Chúng được thiết kế đặc biệt cho tập hợp các sự kiện nhất định – một cơn bão chỉ xuất hiện một lần trong đời hay trận lũ lớn nhất trong một thế kỷ. Khi những sự kiện đó xảy ra, tất cả những dự phòng kỹ thuật đó trở nên đáng giá. Thế nhưng, khi chúng chưa xuất hiện, sự băn khoăn, hoài nghi là có lý do. Trong nhiều năm, các tác giả của *Freakonomics*, Steven Levitt và Stephen Dubner, đã chỉ ra rằng áo phao và phao cứu sinh trên máy bay – không nói đến các minh chứng về độ an toàn của từng loại – dường như là lãng phí, do có rất ít máy bay hạ cánh thành công trên mặt nước. Sau đó, vào năm 2009, Cơ trưởng Sullenberger đã hạ cánh thành công chiếc máy bay của US Airways khi động cơ không hoạt động trên sông Hudson. Liệu một ví dụ về một sự kiện có xác suất xảy ra thấp có khiến việc dự phòng áo phao trở nên có giá trị không? Thật khó để biết. Nhưng chúng ta không thể kết luận rằng việc chưa có kết quả xảy ra khiến chúng ta đánh giá xác suất của kết quả đó bằng không.

Tuy nhiên, trọng tâm của Levitt và Dubner là mặc dù thường có thể áp dụng các biện pháp bảo vệ để đánh giá khả năng xảy ra hay đánh giá thay

đôi trong khả năng xảy ra sự không chắc chắn tiềm ẩn theo thời gian, nhưng không thể đo lường liệu các khoản đầu tư được thực hiện để giảm xác suất xảy ra một hậu quả là thừa thãi hay không, vì chính chiến lược quản lý rủi ro được triển khai sẽ lấy đi thông tin đó. Hoàn toàn có thể xảy ra trường hợp lãng phí vào thứ gì đó mà vì những lý do khác, không còn rủi ro cao nữa.

HỆ THỐNG NHÀ KÍNH

Các sân bay có vẻ là một cơ hội dễ dàng để AI có thể nâng cao chi phí của sự không chắc chắn. Nhưng cơ hội có thể ở ngay trong các hoạt động của bạn. Tìm kiếm sự không chắc chắn tiềm ẩn và tạo các dự báo AI để thúc đẩy các quyết định mới có thể dẫn đến những thay đổi đáng kể trong cách bạn thực hiện công việc kinh doanh của mình.

Trồng trọt nông nghiệp luôn rất bất ổn, chủ yếu là do thời tiết. Nếu thời tiết quá nóng, quá lạnh, quá ẩm, không đủ ẩm hoặc quá gió, thì năng suất sẽ thấp. Đây chính là nguyên nhân khiến bạn muốn nuôi trồng bên trong nhà, nơi các điều kiện thời tiết có thể được kiểm soát. Vấn đề là cây trồng cần ánh sáng để sinh trưởng. Do đó, nhà kính, nơi cây trồng được trồng bên trong nhưng vẫn nhận được ánh sáng mặt trời. Nhà kính mang lại cho người nông dân khả năng kiểm soát đặc biệt về nhiệt độ, độ ẩm



và tươi tiêu. Dĩ nhiên, cái giá cho sự kiểm soát này không hề rẻ. Hệ thống tưới, làm mát và ánh sáng bổ sung đều cần năng lượng. Năng lượng cần thiết có thể ước tính trước và có thể được quản lý.

Tuy nhiên, vấn đề là cây trồng không phải là thứ duy nhất mong muốn các điều kiện khí hậu được kiểm soát. Sâu bệnh cũng phát triển mạnh trong điều kiện đó. Do đó, rệp, muỗi, giun, ve và những loài khác phát triển và sinh sôi nhanh hơn ở ngoài trời. Massachusetts có sổ tay hướng dẫn quản lý nhà kính, 1/3 trong số đó dành cho việc kiểm soát sâu bệnh. Người nông dân mất nhiều thời gian cho việc này. Họ kiểm tra thực vật, loại bỏ các vũng nước đọng, khử trùng dụng cụ và sử dụng thuốc trừ sâu. Phần lớn công việc liên quan đến việc quản lý nhà kính liên quan đến việc bảo vệ nhà kính khỏi khả năng côn trùng xâm nhập hoặc giảm tác động của bất kỳ loài gây hại nào lọt qua cửa.

AI có thể giúp đỡ. Ecoation là một công ty khởi nghiệp đang sử dụng AI để cải thiện việc quản lý sâu bệnh trong nhà kính. Ecoation là một hệ thống do thám. Người vận hành điều khiển máy xung quanh trong nhà kính. Các hệ thống thị giác máy tạo ra các dự báo về sự xâm nhập có thể xảy ra và các khu vực rủi ro, từ đó đưa ra các dự báo về thực trạng: cho người nông dân biết nơi cần thuốc trừ sâu hoặc

các công cụ kiểm soát sâu bệnh khác. Dữ liệu này cũng cho phép AI dự báo về áp lực dịch hại trên toàn bộ nhà kính trong một tuần. Thời gian một tuần là khoảng thời gian có thể đặt hàng và triển khai các công cụ kiểm soát dịch hại. Ưu điểm hàng đầu là tiết kiệm chi phí: với AI, các công cụ kiểm soát dịch hại phù hợp được đặt hàng vào đúng thời điểm. Đó là cách Ecoation hiện đang quảng bá dịch vụ của mình.

Nhưng nhìn vào toàn bộ hệ thống sẽ cho bạn biết có những lợi ích lớn hơn là tiết kiệm chi phí. Nông dân tuân thủ rất nhiều quy tắc nhằm giảm thiểu các vấn đề về sâu bệnh, bao gồm trồng cây trồng kháng sâu bệnh, giữ nhà kính nhỏ để có thể kiểm tra, điều chỉnh điều kiện khí hậu theo những cách nhất định,... Có giá trị thực sự trong việc có thể nới lỏng các quy tắc đó. Nếu AI để dự đoán dịch hại đủ tốt, thì nhà kính có thể hoạt động theo cách khác.

Nông dân có thể trồng các loại cây trồng nhạy cảm với sâu bệnh. Nhà kính lớn hơn trở nên có thể. Các chiến lược thay thế để tiết kiệm năng lượng có thể phát sinh. Nếu các công ty AI như Ecoation làm tốt công việc kiểm soát sâu bệnh, thì chúng ta có thể thay thế các quy tắc hiện có và xây dựng một hệ thống mới. Trong nông nghiệp, giống như trong sân bay, AI có thể cho phép chuyển từ quy tắc sang quyết định.

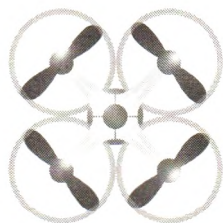


NHỮNG ĐIỂM CHÍNH

- Không chỉ bản thân các quy tắc đại diện cho mục tiêu cơ hội cho các quyết định do AI hỗ trợ mà còn là các tòa nhà và giàn giáo được xây dựng để che giấu sự không chắc chắn dẫn đến lãng phí và kém hiệu quả trong các quy tắc mà chúng tôi đã áp dụng.
- Các sân bay hiện đại là một ví dụ về các tòa nhà và giàn giáo đắt tiền được xây dựng để che giấu sự không chắc chắn. Nguồn gốc của sự không chắc chắn là sự chậm trễ tiềm ẩn do giao thông và an ninh. Các sân bay mới xa hoa được thiết kế để giúp mọi người quên rằng họ đang hoạt động theo một quy tắc buộc họ phải đến sân bay trước giờ khởi hành theo lịch trình rất lâu.
- Trong nhà kính, AI có thể sử dụng các dự đoán về sự xâm nhập của sâu bệnh để nâng cao khả năng phòng ngừa của người trồng trọt. Đây là một giải pháp điểm. Nếu AI để dự đoán dịch hại trở nên đủ tốt, thì thay vì được sử dụng như một giải pháp điểm, AI có thể tạo ra sự thay đổi ở cấp độ hệ thống. Toàn bộ thiết kế cấu trúc và quy trình làm việc của nhà kính bị ảnh hưởng bởi nguy cơ dịch hại xâm nhập. Với dự đoán tốt hơn, nông dân có thể trồng các loại cây trồng khác nhau (nhạy cảm với sâu bệnh hơn), vận hành các nhà kính lớn hơn và theo đuổi các chiến lược tiết kiệm năng lượng thay thế mới.

Chương 6

QUY TẮC LÀ KEO DÍNH



Checklist, hay danh sách kiểm tra là một phần của đời sống công việc hiện đại. Khi tìm kiếm một máy bay ném bom mới, ban đầu quân đội Hoa Kỳ đã từ chối mô hình Boeing 299 so với lựa chọn thay thế của McDonnell Douglas, mặc dù mô hình Boeing này có khả năng mang tải trọng nặng hơn gấp 5 lần, bay nhanh hơn và đi được quãng đường dài gấp đôi. Đúng vậy, nó bị rơi. Vụ tai nạn không phải do vấn đề thiết kế mà do lỗi của phi công. Đó là một chiếc máy bay không dễ dàng cho phi công điều khiển.

Quân đội vẫn quyết định mua một vài chiếc. Tuy nhiên, thay vì để phi công được đào tạo cẩn thận hơn, như Gawande lưu ý, thay vào đó, họ đưa ra một giải pháp đơn giản hơn: họ xây dựng và phân phát cho các phi công một danh sách kiểm tra các bước cần thiết trong các hoạt động khác nhau như cất cánh và hạ cánh:

Sự tồn tại đơn thuần của danh sách này cho thấy ngành hàng không đã tiến xa đến thế nào. Trong những năm đầu, việc đưa một chiếc máy bay



lên không trung có lẽ vô cùng căng thẳng, nhưng dường như không quá phức tạp. Sử dụng một danh sách kiểm tra cho cất cánh sẽ không áp dụng với một phi công nhiều như khi một người lái xe lùi xe ra khỏi nhà để xe. Thế nhưng, máy bay mới này quá phức tạp và việc ghi nhớ quy trình bay là quá khó đối với bất kỳ phi công nào, dù là chuyên gia.

Với danh sách kiểm tra trong tay, các phi công vẫn tiếp tục bay với mô hình 299, đạt tổng cộng quãng đường 1,8 triệu dặm mà không có bất kỳ tai nạn nào. Quân đội cuối cùng đã đặt mua gần 13.000 máy bay, được mệnh danh là B-17.

Gawande lập luận, một cách thuyết phục, rằng y học hiện đại đã phát triển phức tạp đến mức cùng một cách tiếp cận có thể cho nhiều lợi ích khác nhau. Ông biết đó là điều không dễ dàng. Các bác sĩ phẫu thuật hàng đầu vẫn kháng cự với việc phải rửa tay và chà mạnh. Nhưng danh sách kiểm tra áp dụng cho các môi trường phức tạp ở khắp mọi nơi, từ các công trường xây dựng đến bếp bánh. Nếu điều đó có thể cứu lấy nhiều mạng sống, các bác sĩ chắc chắn chấp nhận yêu cầu này.

Chúng tôi sẽ tranh luận với Gawande về giá trị của danh sách kiểm tra, mà thông cảm với những người sử dụng chúng. Danh sách kiểm tra tồn tại vì sự không chắc chắn. Một hệ thống phức tạp bao gồm nhiều phần liên quan đến nhau và nhiều người

thực hiện các nhiệm vụ trong đó để mọi thứ đi vào hoạt động, vì vậy danh sách kiểm tra không chỉ đơn giản là các chỉ số cho thấy một cái gì đó đã được thực hiện. Thay vào đó, chúng là biểu hiện của các quy tắc và sự cần thiết phải tuân theo chúng. Các quy tắc tồn tại để đảm bảo độ tin cậy và giảm lỗi. Phương án thay thế là các chuyên gia đưa ra quyết định dựa trên quan sát của bản thân, điều này tạo ra vấn đề và sự không chắc chắn cho người khác.

Các doanh nghiệp lớn có danh sách kiểm tra. Họ cũng có các quy trình vận hành tiêu chuẩn (SOP), đóng vai trò tương tự. Như đã thảo luận trong Chương 4, SOP là các hướng dẫn sử dụng xác định tất cả các bước mà mọi người cần tuân theo, bao gồm kiểm tra xem họ có thực hiện chúng không. Các SOP giúp các tổ chức phức tạp hoạt động trơn tru. Tuy nhiên, chúng ta phải nhìn nhận chúng như những gì chúng đại diện. Đó là những quy tắc để tuân theo hơn là các quyết định để đưa ra.

SOPS và danh sách kiểm tra là những mảnh vụn của sự không chắc chắn ẩn giấu đã tạo ra vô số quy tắc ăn sâu vào tĩnh mạch của các tổ chức. Luôn có sự không chắc chắn đã dẫn đến mỗi quy tắc. Và đối với mỗi quy tắc, chúng ta có thể đặt câu hỏi: nếu chúng ta có dự báo AI, chúng ta có thể nâng cao năng suất bằng cách biến quy tắc thành một quyết định và loại bỏ nó khỏi hướng dẫn SOP không?



NHỮNG NGƯỜI KHÁC NHAU KHÁC NHAU

Các quy tắc mặc định mọi người cùng một việc như thế tất cả họ đều giống nhau. Nhưng thực tế họ không giống nhau. Đó có lẽ là bài học cơ bản của marketing. Vì vậy, những người làm marketing cố gắng phân chia dân số thành các nhóm và nhắm sản phẩm đến những nhóm mục tiêu có thể thấy sản phẩm hấp dẫn.

Khi các nhà tiếp thị nhìn nhận mọi người như nhau, đó là vì họ thiếu thông tin. Nếu có thông tin, họ sẽ cung cấp các sản phẩm và dịch vụ được cá nhân hóa. Họ có thể chuyển từ các quy tắc nhìn nhận mọi người như nhau sang các quyết định cho phép họ cung cấp đúng sản phẩm cho đúng người vào đúng thời điểm.

Radio là một ví dụ điển hình của quy tắc. Đài phát thanh đã thuê các DJ phát các bài hát giống nhau cho mọi người nghe. Các dịch vụ phát nhạc trực tuyến như Spotify, Apple Music và Pandora cho phép tạo danh sách phát được cá nhân hóa.

Vậy có những thách thức nào liên quan đến việc tạo ra giá trị từ danh sách phát được cá nhân hóa? Các nhà nghiên cứu David Reiley và Hongkai Zhang tại Pandora đã hỏi câu hỏi này khi họ tập trung vào việc kiểm tra các quy tắc khác tại công ty. Mặc dù danh sách phát nhạc đã được cá nhân

hóa, hoạt động kinh doanh vẫn được vận hành theo quy tắc. Pandora áp dụng một mô hình được gọi là freemium. Một số khách hàng trả phí và nhận được trải nghiệm nghe nhạc mà không có quảng cáo. Những người khác nghe miễn phí miễn là họ xem hết một số lượng quảng cáo nhất định mỗi giờ.

Reiley và Zhang, sau khi làm việc với giáo sư Ali Goli của Đại học Washington, đã nhận ra rằng, họ có thể áp dụng AI vào dữ liệu từ một thử nghiệm xác định mức độ thích và không thích của người nghe đối với quảng cáo. AI cung cấp các dự báo được cá nhân hóa, từ đó họ có thể đánh giá không chỉ mức độ trung bình mà mọi người không thích quảng cáo mà còn cả sự khác biệt giữa mọi người. Với thông tin đó, họ không còn cần tuân theo quy tắc về cách chèn quảng cáo nữa. Thay vào đó, sẽ có nhóm người nhận được nhiều quảng cáo hơn và một số ít hơn. Khi cá nhân hóa số lượng quảng cáo, họ nhận thấy lợi nhuận có thể tăng đáng kể. AI dự đoán nhóm khách hàng nào sẽ nghe nhiều hơn nếu giảm quảng cáo. Nó cũng dự đoán nhóm khách hàng nào có thể được thuyết phục chuyển sang phiên bản trả phí.

Với thông tin này trong tay, họ không còn cần quy tắc hiển thị cùng một số lượng quảng cáo cho mọi người nữa. Họ có thể quyết định hiển thị ít quảng cáo hơn với nhóm sẽ tăng thời gian nghe nếu số lượng quảng cáo mỗi giờ giảm xuống. Họ cũng sẽ



hiển thị nhiều quảng cáo hơn với nhóm sẽ chuyển sang phiên bản trả phí. Bộ phận nghiên cứu của Pandora đã chỉ ra cách AI có thể đưa ra các quyết định mới.

Điều này không hề đơn giản. Để mở rộng khả năng quảng cáo đòi hỏi phải tìm nhà quảng cáo. Goli, Reiley và Zhang ước tính rằng, chỉ 2/3 vị trí quảng cáo sẽ được lấp đầy. Thách thức đặt ra với họ là cần thêm các nhà quảng cáo mới để tránh gửi đi gửi lại cùng một quảng cáo cho cùng một khách hàng. Để triển khai thành công đòi hỏi một chiến lược bán hàng quảng cáo mới.

Nó cũng đòi hỏi sự hiểu biết về phản ứng của khách hàng. Khía cạnh sinh lợi nhất của AI là tăng cường quảng cáo cho những khách hàng đang phân vân giữa phiên bản miễn phí và trả phí. Bằng cách làm giảm chất lượng của phiên bản miễn phí, những khách hàng này sẽ chuyển sang phiên bản trả phí. Tuy nhiên, khách hàng có thể sẽ khó chịu nếu đây là cách Pandora sử dụng dữ liệu của họ. Do đó, chiến lược này có nguy cơ khiến người dùng từ bỏ sử dụng dịch vụ hoàn toàn.

Những hạn chế này có nghĩa là Pandora vẫn chưa triển khai AI. Nó vẫn sử dụng một quy tắc để xác định số lượng quảng cáo sẽ hiển thị. Xây dựng AI là bước đầu tiên để loại bỏ một quy tắc. Các quy trình vẫn cần phải thay đổi để có thể ra quyết định.

MỘT VIÊN GẠCH KHÁC TRÊN BỨC TƯỜNG

Giáo dục chứa đầy các quy tắc. Ngồi ở đâu. Cư xử ra sao. Hành động thế nào. Một trong số chúng tôi (Avi) đã nhận được cuốn “hướng dẫn dành cho phụ huynh về các chính sách và thực hành” dài 59 trang từ trường học của con anh ấy. Cuốn hướng dẫn bao gồm các quy tắc về sức khỏe và an toàn liên quan đến dị ứng, chảy rận, thương tích và tiêm chủng. Nó cũng bao gồm quy định về bài tập về nhà, cách tổ chức sinh nhật, điện thoại di động, đưa đón và sắp xếp lớp học. Và đó mới là hướng dẫn dành cho cha mẹ!

Những quy tắc này phục vụ một mục đích. Chúng hỗ trợ xây dựng một hệ thống giáo dục an toàn và hiệu quả. Như Cosmo Kramer đã nói trên *Seinfeld*, “Quy tắc là quy tắc và hãy đối mặt với nó, không có quy tắc thì sẽ hỗn loạn”.

Dĩ nhiên, có thể có những quy tắc thừa thãi. Nỗi lo lắng rằng giáo dục tạo ra sự đồng nhất đến mức rập khuôn đã mạnh mún từ rất lâu. Năm 1859, John Stuart Mill đã viết trong *On Liberty*, “nền giáo dục phổ cập của nhà nước chỉ là phương tiện để uốn nắn con người trở nên giống hệt nhau”.

Các nhà giáo dục nhận thức rõ ràng về sự căng thẳng này giữa các quy tắc và tính linh hoạt. Tài liệu giáo dục mô tả vấn đề này và cố gắng tìm ra



giải pháp. Các Tiêu chuẩn Học tập Mẫu giáo của Tiểu bang New York nhấn mạnh:

Thay vì quy định tiến trình bài học hoặc chương trình giảng dạy theo từng bước cho tất cả trẻ em ở mọi môi trường, các tiêu chuẩn làm rõ những kỳ vọng mà trẻ em có thể học và làm nhờ sự hướng dẫn không được tiêu chuẩn hóa mà được cá nhân hóa, được phân biệt, thích nghi, phù hợp về mặt văn hóa và ngôn ngữ, và dựa trên ngữ cảnh. Mặc dù có thể đặt ra các mục tiêu học tập giống nhau cho tất cả trẻ em, phương tiện chúng ta áp dụng để đạt được những mục tiêu này sẽ thay đổi rất lớn đối với từng trẻ.

Vì vậy, có những tiêu chuẩn đặt ra cho tất cả mọi người, nhưng giáo dục mà mỗi học sinh nhận được là khác nhau. Đó là một tầm nhìn tuyệt vời và cũng đầy thách thức. Những giáo viên có thể hiện thực hóa điều này chính là những người thầy tốt nhất. Họ điều chỉnh bài học theo từng học sinh trong lớp. Những giáo viên khác sẽ thấy khó khăn hơn để thực hiện điều này. Từ góc độ toàn cầu, nó thậm chí còn khó hơn. Ở những quốc gia có thu nhập cao, bố mẹ chi hàng nghìn đô la cho mỗi đứa trẻ. Nhiều quốc gia có thu nhập thấp chỉ chi 50 đô la cho mỗi đứa trẻ mỗi năm. Với nguồn lực ít ỏi, rất khó để loại bỏ các quy tắc.

Chúng ta có thể bắt đầu từ giáo dục khởi nghiệp. Các cơ quan viện trợ như Ngân hàng Thế giới và các

chính phủ trên khắp thế giới hàng năm chi hơn 1 tỉ đô la để đào tạo khoảng 4 triệu doanh nhân và tài năng kinh doanh ở các nước đang phát triển. Nhiều chương trình đào tạo trong số này hướng đến mục đích cải thiện hoạt động kinh doanh và lợi nhuận, nhưng chúng rất tốn kém và không phải lúc nào cũng thấy được lợi tức đầu tư. Đào tạo trực tuyến hứa hẹn những lợi ích to lớn, nhưng không phải là hình thức có thể đồng nhất. Thông tin thương mại được thiết kế chung cho mọi đối tượng không tạo được nhiều sự khác biệt. Một trong những bài học lớn rút ra ở đây là đào tạo cá nhân chuyên sâu cho hiệu quả tốt nhất. Thách thức đặt ra là làm thế nào để cung cấp nền giáo dục được cá nhân hóa đó trên quy mô lớn.

Các nhà kinh tế Yizhou Jin và Zhengyun Sun nghĩ rằng AI có thể giúp ích. Họ đã làm việc với một nền tảng thương mại điện tử lớn để đào tạo về kinh doanh cho hàng trăm nghìn người bán hàng mới. Chương trình đào tạo liên quan đến hàng chục mô-đun khả thi và tập trung vào việc thiết lập trang web, chiến lược tiếp thị và dịch vụ khách hàng. Ví dụ: khóa đào tạo có thể cung cấp danh sách kiểm tra các phương pháp hay nhất trong phần mô tả sản phẩm để khách hàng hiểu họ đang mua gì. Một khía cạnh khác của khóa đào tạo tập trung vào việc tối ưu hóa công cụ tìm kiếm và lựa chọn từ khóa.



Không phải tất cả các mô-đun đều phù hợp với mọi người bán và người bán mới có thể không biết loại hình đào tạo nào sẽ hữu ích. AI kích hoạt cá nhân hóa. Nó lấy dữ liệu về các hoạt động và sản phẩm thực tế của người bán và phát triển một trình tự đào tạo. Sau đó, nó đề xuất các mô-đun cho người bán. Sau đó, người bán đã triển khai các mô-đun. Điều này có nghĩa là giáo dục tinh thần kinh doanh được cá nhân hóa cho hàng trăm nghìn người bán. Thay vì quy tắc rằng mọi người bán đều nhận được thông tin giống nhau, AI đã kích hoạt các quyết định mới về việc doanh nhân nào được đào tạo.

Chương trình được chạy bằng thử nghiệm kiểm soát ngẫu nhiên để có thể đo lường hiệu quả của nó. Trong số 8 triệu doanh nghiệp mới trên nền tảng này, 2 triệu doanh nghiệp đã được cung cấp khóa đào tạo. Trong số này, khoảng 500.000 đã sử dụng nó. Các doanh nghiệp sử dụng khóa đào tạo đã tăng 6,6% doanh thu. Trong suốt một năm, chương trình đã tăng doanh thu của người bán lên khoảng 6 triệu đô la. Điều đó nghe có vẻ không nhiều: 12 đô la cho mỗi doanh nghiệp mỗi năm.

Nhưng đó là bắt đầu từ cơ sở tổng doanh thu là 200 đô la. Một chương trình đào tạo được cá nhân hóa với giáo viên là con người không bao giờ có thể tiết kiệm chi phí. AI cho phép quyết định nên gửi khóa đào tạo nào cho doanh nhân nào trong

khi tiếp cận hàng trăm nghìn doanh nghiệp. AI kích hoạt các quyết định đối với các quy tắc và tạo ra giá trị trên quy mô lớn.

GỖ BỎ CÁC QUY TẮC

Khi các quy tắc đã tồn tại trong một thời gian dài, khó có thể nhìn thấy hệ thống mà trong đó chúng được tích hợp. Vì các quy tắc đáng tin cậy nên vô số quy tắc và thủ tục có thể gắn kết với nhau. Nếu một yếu tố thay đổi, cả khối theo đó sẽ phải thay đổi theo.

Trên phiên bản miễn phí của Pandora, mọi người dùng sẽ nhận được lượng quảng cáo như nhau. Quy tắc này thường được áp dụng với các phương tiện hỗ trợ quảng cáo. Các đài truyền hình sử dụng 8 phút sau mỗi nửa giờ để quảng cáo. Đây là quy tắc thúc đẩy doanh thu của cả hệ thống. Sau đó, một loạt các quy trình khác đã được phát triển xoay quanh quy tắc này. Các chương trình được thiết kế dài 22 phút hoặc 44 phút, có nghĩa là các nhà biên kịch cần viết mỗi tập của một chương trình có cùng thời lượng, với các khoảng dừng tự nhiên để chèn quảng cáo. Quy tắc này đã gắn chặt vào hệ thống.

YouTube là ví dụ về một phương thức khác cho thiết kế nội dung cho hệ thống. Không giống như truyền hình, người sáng tạo nội dung trên



YouTube có thể tạo nội dung với độ dài bất kỳ. Trí tuệ nhân tạo của hệ thống có thể dự đoán người xem nào sẽ bị thu hút nhất bởi nội dung nào. AI điều khiển công cụ tìm kiếm và công cụ đề xuất cho phép người xem tìm thấy nội dung phù hợp trong danh mục tùy chọn dài bất tận. Hơn nữa, AI có thể dự đoán người dùng nào sẽ bị thu hút nhiều nhất bởi quảng cáo nào. Điều quan trọng là khả năng dự đoán này có giá trị hơn nhiều trong một hệ thống cho phép những người dùng khác nhau xem các nội dung khác nhau. Ngay cả khi truyền hình triển khai AI và có thể tạo ra các dự đoán tương tự, thì giá trị sẽ thấp hơn nhiều vì hệ thống của nó buộc mọi người xem phải xem cùng một nội dung. Vì vậy, điều tốt nhất nó có thể làm là dự đoán quảng cáo nào sẽ hấp dẫn nhất đối với nhiều người xem nhất.

Nói cách khác, cùng là triển khai AI để dự đoán mức độ thu hút của người xem đối với nội dung và quảng cáo, nhưng giá trị mà nó mang lại lớn hơn nhiều trong hệ thống YouTube so với hệ thống truyền hình. Và mặc dù AI *trực tiếp* cho phép khám phá trong một danh mục nội dung rộng lớn và cho phép quảng cáo tương ứng, nhưng nó *gián tiếp* cho phép độ dài nội dung linh hoạt bởi khi đó các giải pháp khám phá và quảng cáo giải quyết vấn đề của việc vô số kết hợp nội dung, quảng cáo và thời

gian khiến hạn chế sự linh hoạt của độ dài nội dung truyền hình.

Trong hệ thống trường học, học sinh cùng một cấp sẽ học cùng một chương trình. *Có chương trình học cố định*. “Học sinh được học theo nhóm, chia theo độ tuổi, như thể điểm chung quan trọng nhất của chúng là *ngày sản xuất*”. Ví dụ, ở Ontario, nơi chúng tôi sống, hầu hết trẻ em sinh năm 2009 đều sẽ vào lớp 1 năm 2015 và vào trung học năm 2023. Những quy tắc này được áp dụng để quản lý sự không chắc chắn về trình độ của học sinh, xét về lĩnh vực học tập và xã hội. Đổi lại, những quy tắc này gắn kết với nhau trong một hệ thống: đào tạo giáo viên để quản lý các nhu cầu học tập không đa dạng; không có nhiều sự trợ giúp cũng như nguồn lực hỗ trợ cho học sinh bị tụt lại phía sau. Và, ở cấp trung học, có những chương trình mang tính hình thức dành cho học sinh không tuân theo quy trình tiêu chuẩn áp dụng nhóm (tuổi), bao gồm các trường học thay thế phù hợp, các chương trình vừa học vừa làm và quy trình để lấy chứng chỉ tương đương với trường trung học.

Khi AI dự đoán chương trình học tốt nhất tiếp theo cho mỗi học sinh, khi đó giáo dục được cá nhân hóa, cho phép học sinh nắm vững một chủ đề nhanh chóng chuyển sang một chủ đề mới trước khi chúng cảm thấy nhàm chán, đồng thời



cho phép học sinh có thêm thời gian thực hành với một chủ đề nếu cần, đồng thời bổ sung các bài tập để phát triển năng lực trong lĩnh vực đó trước khi chuyển sang chủ đề hay môn học mới. Là một giải pháp điểm, AI này có thể nâng cao hiệu quả học tập trong hệ thống trường học hiện tại ở một mức độ nào đó, mặc dù tác động sẽ bị hạn chế vì một khi hoàn thành chương trình học tiêu chuẩn theo độ tuổi (cấp học), học sinh sẽ hoàn thành chương trình học cho cả năm, và khi đó việc học nâng cao sẽ không nhận được nhiều hỗ trợ từ giáo viên vì giáo viên thường được đào tạo cho một cấp học cụ thể (ví dụ: toán trung học cơ sở). Trong hệ thống hiện tại, vấn đề này sẽ ngày càng trở nên nghiêm trọng ở các lớp sau này khi sự chênh lệch giữa học sinh tiếp thu nhanh và chậm hơn trong một môn học ngày càng tăng theo thời gian. Để hỗ trợ học sinh của mình, giáo viên sẽ cần thành thạo việc giảng dạy đối với một loạt các chủ đề khác nhau và sự đa dạng này sẽ ngày càng lớn.

Thay vào đó, hãy tưởng tượng một hệ thống giáo dục mà trong đó học sinh vẫn học dần lên theo lớp (sự phát triển về thể chất và xã hội của các em sẽ dựa vào mặt sinh học), nhưng người hướng dẫn và giáo viên sẽ có nhiệm vụ hỗ trợ các học sinh khác nhau, tùy thuộc vào nhu cầu học tập của từng cá nhân. Người hướng dẫn và giáo viên

hỗ trợ học sinh không phụ thuộc vào độ tuổi của học sinh đó mà thay vào đó được xác định bởi bản chất của các câu hỏi và khả năng của các em trong một lĩnh vực cụ thể. Tác động của AI sẽ lớn hơn nhiều trong hệ thống mới này so với tác động của cùng một AI đó trong hệ thống hiện tại vì mỗi học sinh có thể nhận được sự giáo dục được cá nhân hóa theo nhu cầu và phong cách học tập của các em. Những học sinh học nhanh với một môn học nhưng chậm ở những môn học khác có thể được điều chỉnh. Những học sinh cần tập trung vào các kỹ năng cụ thể sẽ có giáo viên chuyên về kỹ năng đó. Giáo viên sẽ không cần phải chọn phong cách giúp ích cho nhiều học sinh nhất. Những giáo viên giỏi trong việc giúp đỡ những học sinh gặp khó khăn trong việc đọc và những giáo viên xuất sắc trong việc giúp học sinh tỏa sáng trong các cuộc thi toán học sẽ dành toàn bộ thời gian cho những gì họ làm tốt nhất.

Các quy tắc như lập trình 22 phút và chương trình giảng dạy theo độ tuổi đã được đưa ra để đối phó với sự không chắc chắn. Sau đó, nhiều dạng giàn giáo khác nhau đã được phát triển để tối ưu hóa hiệu suất của hệ thống. Mặc dù vô hình đối với người quan sát bình thường, nhưng các quy tắc đã trở thành chất keo kết dính hệ thống lại với nhau. Vì vậy, việc giới thiệu một AI cho phép chuyển một



quy tắc thành một quyết định thoát nhìn có vẻ hấp dẫn, nhưng tác động của nó có thể bị hạn chế vì quy tắc mà nó thay thế được kết hợp chặt chẽ với các yếu tố khác của hệ thống.

Việc đưa AI dự đoán nội dung tốt nhất tiếp theo vào hệ thống trường học hiện tại sẽ có tác động hạn chế vì quy tắc chương trình giảng dạy dựa trên độ tuổi với một giáo viên mỗi lớp là nền tảng của hệ thống giáo dục hiện tại, đặc biệt là ở trường tiểu học. Ngược lại, sử dụng chính xác cùng một AI, nhưng nhúng nó vào một hệ thống mới được thiết kế để tận dụng nội dung và tốc độ được cá nhân hóa của AI bằng cách kết hợp nó với thảo luận được cá nhân hóa, các dự án nhóm và hỗ trợ của giáo viên, điều này sẽ yêu cầu phân bổ và sửa đổi giáo viên và trợ giảng linh hoạt hơn nhiều đào tạo giáo viên, có thể sẽ dẫn đến tác động lớn hơn nhiều đối với giáo dục cũng như sự tăng trưởng và phát triển cá nhân.

Nói cách khác, quy tắc chương trình giảng dạy dựa trên độ tuổi là chất keo gắn kết phần lớn hệ thống giáo dục hiện đại và do đó, AI cá nhân hóa nội dung học tập chỉ có thể mang lại lợi ích hạn chế trong hệ thống đó. Thách thức chính để giải phóng tiềm năng của AI giáo dục được cá nhân hóa không phải là xây dựng mô hình dự đoán mà là tách giáo dục ra khỏi quy tắc chương trình giảng dạy dựa trên độ tuổi hiện đang gắn kết hệ thống lại với nhau.

NHỮNG ĐIỂM CHÍNH

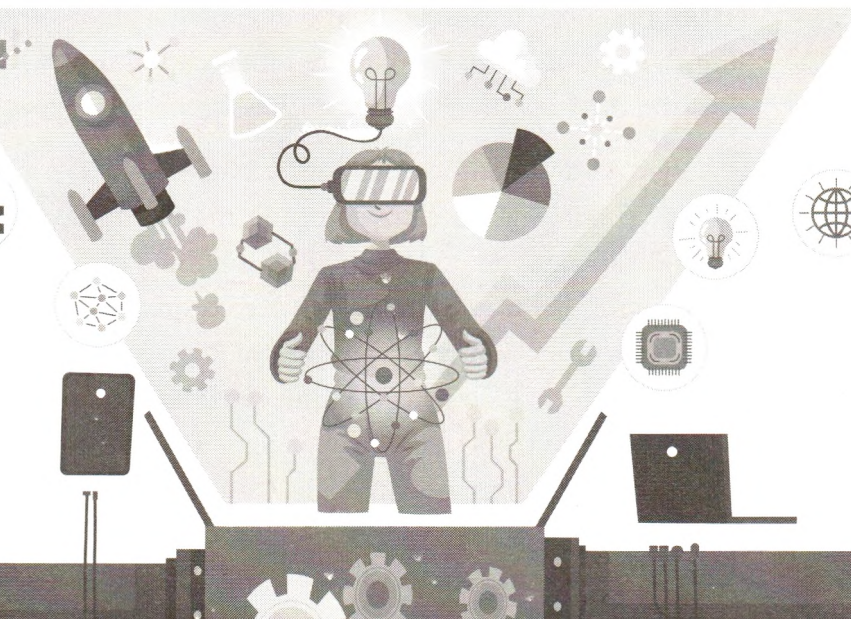
- Giống như SOP, danh sách kiểm tra là biểu hiện của các quy tắc và sự cần thiết phải tuân theo chúng. Họ ở đó để đảm bảo độ tin cậy và giảm lỗi. Cách khác là mọi người đưa ra quyết định dựa trên những quan sát của chính họ. Mặc dù việc chuyển từ một quy tắc sang một quyết định có thể cải thiện chất lượng của hành động cụ thể đó, nhưng nó cũng có thể tạo ra các vấn đề và sự không chắc chắn cho những người khác.
- Các quy tắc gắn kết với nhau trong một hệ thống. Đó là lý do tại sao khó có thể thay thế một quy tắc duy nhất bằng một quyết định do AI hỗ trợ. Do đó, thường xảy ra trường hợp một AI rất mạnh chỉ bổ sung giá trị cận biên vì nó được đưa vào một hệ thống có nhiều bộ phận được thiết kế để phù hợp với quy tắc và chống lại sự thay đổi. Chúng phụ thuộc lẫn nhau – được dán lại với nhau.
- Một ví dụ là AI giáo dục được cá nhân hóa dự đoán nội dung tốt nhất tiếp theo sẽ trình bày cho người học. Việc đưa AI này vào một hệ thống được thiết kế theo quy tắc chương trình giảng dạy dựa trên độ tuổi sẽ làm giảm lợi ích. Ngược lại, việc đưa chính AI đó vào một hệ



thống mới thúc đẩy thảo luận được cá nhân hóa (không dựa trên độ tuổi), các dự án nhóm và hỗ trợ của giáo viên có thể sẽ tạo ra tác động lớn hơn nhiều đối với giáo dục nói chung cũng như sự tăng trưởng và phát triển cá nhân. Thách thức chính để giải phóng tiềm năng của AI giáo dục được cá nhân hóa không phải là xây dựng mô hình dự đoán mà là giáo dục không gắn bó với quy tắc chương trình giảng dạy dựa trên độ tuổi hiện đang gắn kết hệ thống lại với nhau.

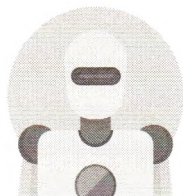
PHẦN 4

QUYỀN LỰC



Chương 7

DÍNH CHẶT VÀ BÔI TRƠN



Al không cứu chúng ta khỏi Covid-19, nhưng đáng ra nó đã có thể làm được. Điều đó không xảy ra bởi vì khi đối mặt với những điều không chắc chắn, nhiều quốc gia lựa chọn tuân theo các quy trình y tế công cộng đã được xây dựng dựa trên các quy tắc từ trước thay vì đưa ra quyết định. Trong những chương trước, chúng tôi đã lưu ý rằng dự đoán AI có khả năng thúc đẩy chuyển dịch các quy tắc sang quyết định. Do đó, đại dịch là cơ hội phù hợp để chúng ta bắt đầu cuộc thảo luận làm thế nào để AI có thể tạo điều kiện cho những thay đổi đó.

Việc AI đã không cứu chúng ta khỏi Covid-19 không có nghĩa là AI chưa sẵn sàng mà là chúng ta chưa sẵn sàng để nó làm điều đó. Ở nhiều quốc gia, các quy tắc truyền thống mà các cơ quan y tế công cộng trong khu vực công đã thiết lập không còn phù hợp với kiểu ra quyết định cần thiết để bảo vệ nền kinh tế trước đại dịch bất ngờ. Tuy nhiên, vẫn có một số trường hợp ngoại lệ. Chúng tôi lấy ví

dự về một nhóm nhỏ gồm các công ty lớn đã thiết lập một nền tảng đổi mới để vận hành hệ thống, cho phép đưa ra các quyết định phù hợp và nhanh chóng trong điều kiện không chắc chắn để ngăn chặn việc ngừng hoạt động do hệ thống dựa trên các quy tắc khắc nghiệt đã bỏ qua thông tin.

QUY TẮC ĐẮT ĐỎ NHẤT

Giờ đây, tất cả chúng ta đều đã quen thuộc với những rủi ro về y tế và sức khỏe tiềm ẩn có thể phát sinh trong đại dịch. Vào tháng 1 năm 2021, khoảng 9 triệu người Mỹ mắc Covid-19. Đối với họ, Covid-19 là vấn đề sức khỏe nghiêm trọng. Tuy nhiên, đối với hầu hết trong số 320 triệu người Mỹ khác, Covid-19 không phải là vấn đề sức khỏe. Họ không bị ốm cũng không bị lây nhiễm. Tuy nhiên, nhiều người vẫn chịu ảnh hưởng nghiêm trọng liên quan đến công việc, học tập và giải trí. Phần lớn mọi người bị ảnh hưởng bởi Covid-19 không phải vì vấn đề sức khỏe mà vì vấn đề dự đoán. Chúng ta thiếu thông tin để dự đoán ai là người bị nhiễm và có thể lây lan virus sang người khác.

Thông điệp từ các cơ quan y tế công cộng là bạn có thể an toàn bằng cách xem tất cả mọi người đều có khả năng bị nhiễm và lây bệnh như nhau. Với các bệnh truyền nhiễm có thể lây từ người này sang người khác, việc tiếp xúc với người khác trở



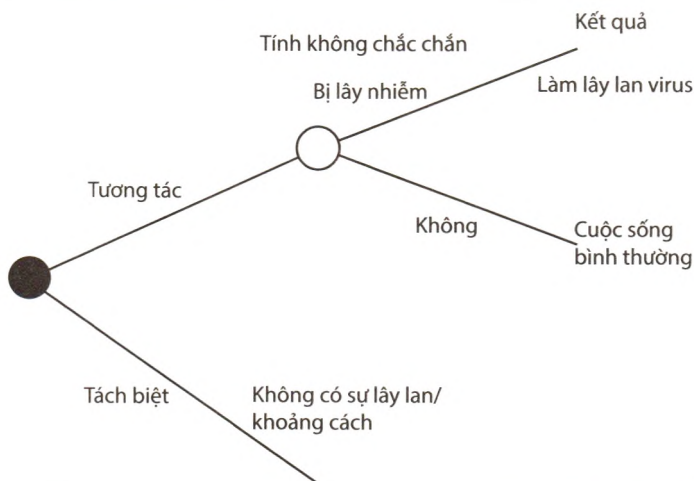
nên nguy hiểm hơn nếu bạn không biết ai là người lây bệnh. Đây là lý do chúng ta giữ khoảng cách với những người khác trong đại dịch. Đó là cách đơn giản nhất để bảo vệ chính chúng ta.

Hãy đặt điều này lên cây quyết định. Hành động được thực hiện là liệu sẽ tách biệt hay tương tác với những người khác (được mô tả trong Hình 7-1). Nếu giữ khoảng cách, bạn không lây bệnh, nhưng bạn phải giữ khoảng cách, điều này rất tốn kém cho cá nhân bạn. Nếu bạn tương tác với người khác, thì hậu quả phụ thuộc vào việc bạn có bị lây nhiễm hay không. Nếu bị nhiễm, bạn có thể lây lan virus. Nếu không, cuộc sống của bạn vẫn diễn ra bình thường.

Cây quyết định làm nổi bật vấn đề phát sinh vì hầu hết mọi người không bị lây nhiễm. Nếu bạn hiện đang bị nhiễm Covid-19, bạn sẽ nguy hiểm hơn nhiều so với nếu bạn không nhiễm. Nói cách khác, nếu biết ai bị nhiễm và ai không bị nhiễm, chúng ta có thể làm điều gì đó khác. Chúng ta có thể tránh xa những người bị nhiễm bệnh và cư xử bình thường hơn với những người không bị nhiễm. Đây là vấn đề dự đoán trọng tâm của đại dịch: chúng ta có thể tránh được nhiều tổn thất trong đại dịch khi biết ai là người lây nhiễm và giữ họ tránh xa những người khác. Những người không bị nhiễm có thể đi làm bình thường trong khi người nhiễm được cách ly. Khi đó, bạn không chỉ giữ cho mọi thứ bình

thường hơn mà còn có thể kiểm soát dịch bệnh khi các chuỗi lây truyền bị phá vỡ. Vấn đề là cần phải có thông tin để biến giãn cách xã hội từ một quy tắc thành một quyết định. Như vậy, nhu cầu về thông tin để giải quyết sự không chắc chắn đồng nghĩa với việc chúng ta gặp vấn đề về dự báo.

Hình 7-1: Tách biệt hoặc liên kết cây quyết định



COVID-19 - VẤN ĐỀ VỀ DỰ BÁO

Bước đầu tiên trong việc xác định vấn đề dự báo là đặt câu hỏi về sự không chắc chắn ở đâu. Xét ở khía cạnh này, đại dịch chứa đầy sự không chắc chắn. Một ẩn số lớn là khi nào mầm bệnh có khả năng gây đại dịch có thể tấn công. Thực tế đó



có thể là vấn đề mà AI có thể giải quyết. Tuy nhiên, điều chúng tôi muốn tập trung vào là một thứ gần gũi hơn: quản lý đại dịch. Đó là, khi một mầm bệnh đang trên bờ vực trở thành đại dịch hoặc một mầm bệnh đã biến chuyển thành đại dịch, thì đâu là điều không chắc chắn chính trong việc nhanh chóng dập tắt nó?

Đây có lẽ là một cách kỳ lạ để hình thành vấn đề quản lý đại dịch. Rốt cuộc, chúng ta đã quen nghĩ về nó như một thách thức đối với sức khỏe cộng đồng. Làm thế nào để chúng ta tìm ra vắc-xin để chấm dứt nó, các phương pháp điều trị có thể cứu sống hoặc giảm nhẹ để giảm thiểu sự lây lan? Nhưng khi chúng ta khám phá ra điều gì đang thực sự khiến dịch bệnh trở thành đại dịch, với tất cả những cái giá phải trả về con người không chỉ về sức khỏe mà còn cả sinh kế kinh tế và đời sống xã hội, chúng ta nhận ra rằng, chính những biện pháp giảm thiểu ngăn cản mọi người lây nhiễm cho nhau đang lấy đi cuộc sống bình thường của chúng ta với họ.

Trong những tháng đầu tiên của đại dịch, rất nhiều công cụ đã được phát triển để dự đoán ai có thể bị nhiễm bệnh. Một trong những cách lâu đời nhất mà các quan chức y tế công cộng phân biệt ai có hoặc không có khả năng bị nhiễm bệnh là theo dõi phơi nhiễm. Nếu ở gần một người truyền nhiễm, bạn có nhiều khả năng bị nhiễm bệnh hơn.

Việc theo dõi liên lạc như vậy có thể giúp các quan chức dự đoán ai có thể đã bị nhiễm bệnh gần đây. Ở nhiều quốc gia, đây là một quá trình tốn nhiều công sức và không chắc chắn, với việc các quan chức gọi điện cho những người bị nhiễm bệnh và hỏi họ đã ở đâu. Tại Hàn Quốc, các cơ quan y tế công cộng đã phát triển các quy trình mới bằng cách kết hợp dữ liệu từ camera mạch kín, giao dịch quét thẻ và điện thoại di động để hỗ trợ các nỗ lực truy tìm người tiếp xúc. Sự đổi mới không chỉ giới hạn trong việc theo dõi liên hệ. Các chuyên gia AI cũng đã phát triển các công cụ để dự đoán khả năng lây nhiễm. Một nhóm đã phát triển các công cụ để phát hiện các bệnh nhiễm trùng không có triệu chứng bằng cách cho mọi người ho vào điện thoại của họ. Tại biên giới Hy Lạp, công cụ AI tính đến cả các yếu tố như phương thức di chuyển, điểm khởi hành và thông tin nhân khẩu học và được cập nhật hằng tuần có thể phát hiện số người không có triệu chứng cao gấp 1,85 lần so với giám sát ngẫu nhiên, do đó giúp xác định khách du lịch nào có thể vào biên giới Hy Lạp mà không cần kiểm tra thêm. Các công cụ dự đoán không phải AI cũng được phát triển. Nhiều nơi dùng máy ảnh nhiệt và nhiệt kế để phát hiện sốt dựa trên nhận định rằng người có thân nhiệt tăng cao dễ mắc Covid-19. Ở Thái Lan, chó nghiệp vụ được huấn luyện để phát hiện bệnh ở người.



Đến mùa thu năm 2020, nhiều người đã thấy rõ rằng xét nghiệm kháng nguyên nhanh là công cụ hiệu quả nhất để dự đoán khả năng lây nhiễm. Mặc dù có thể phát hiện lượng virus rất thấp, xét nghiệm phản ứng chuỗi polymerase (PCR) cho kết quả chậm hơn và đắt hơn xét nghiệm kháng nguyên.

Dự đoán là quá trình lấp đầy phần thông tin còn thiếu và các xét nghiệm Covid-19 giúp điền thông tin liệu ai đó có bị lây nhiễm hay không. Giống như các dự đoán khác, xét nghiệm kháng nguyên nhanh không chính xác 100%. Tuy nhiên, rất hiếm khi có kết quả dương tính giả với các xét nghiệm kháng nguyên, nghĩa là không có khả năng ai đó cho kết quả dương tính và không bị lây nhiễm. Vì vậy, nếu bạn có thể kiểm tra mọi người và yêu cầu những người có xét nghiệm kháng nguyên dương tính ở nhà, thì sự lây lan của bệnh có thể được kiểm soát. Điều này cũng không đúng với các xét nghiệm PCR, khi có thể cho kết quả dương tính trong nhiều tuần hoặc nhiều tháng sau giai đoạn lây nhiễm. Nói cách khác, vào mùa thu năm 2020, chúng tôi đã có một công cụ dự đoán Covid-19 rẻ tiền có thể được sản xuất hàng loạt. Nó không phải là một công cụ AI mà là một loại thiết bị dự báo khác.

Được trang bị kiến thức này, chúng tôi đã hợp tác với nhà dịch tễ học Laura Rosella, nhà khoa

học chính trị Janice Stein và giám đốc điều hành Creative Destruction Lab Sonia Sennik để thiết kế và giúp các tập đoàn triển khai chương trình xét nghiệm nhanh để giúp các cơ sở làm việc mở cửa an toàn.

Ý tưởng là kiểm tra công nhân thường xuyên, cách ly những công nhân dương tính với Covid ở nhà và cho phép những người khác đi làm, được an ủi khi biết đồng nghiệp của họ gần đây đã có kết quả xét nghiệm âm tính. Công cụ dự báo đã có sẵn và kế hoạch có vẻ dễ thực hiện. Nó sẽ làm cho những nơi làm việc thiết yếu chưa đóng cửa an toàn hơn và theo thời gian, cho phép nền kinh tế bắt đầu mở cửa trở lại.

Tuy nhiên, chúng tôi sớm biết rằng công cụ dự báo là một phần dễ dàng. Hệ thống này được kết dính với nhau bằng nhiều quy tắc không phù hợp với việc ra quyết định dựa trên thông tin. Có các quy tắc về quyền riêng tư xung quanh việc thu thập thông tin sức khỏe của mọi người, các quy tắc của công đoàn về việc hạn chế quyền tiếp cận nơi làm việc, các quy tắc bảo mật dữ liệu xung quanh việc lưu trữ và xử lý thông tin cá nhân, các quy tắc xử lý chất thải nguy hại đối với việc loại bỏ dung dịch đệm sau khi các xét nghiệm được thực hiện, các quy tắc bồi thường cho người lao động về việc ai sẽ



chịu chi phí thời gian nghỉ khi ai đó có kết quả xét nghiệm dương tính và danh sách này vẫn tiếp tục.

Bất chấp sự cấp bách của việc giải quyết vấn đề thông tin đe dọa đóng cửa nền kinh tế, mức độ mà hệ thống được gắn chặt với các quy tắc khiến cho một giải pháp ra quyết định dựa trên thông tin gần như không thể thực hiện được. Chúng tôi cần tìm cách tra dầu hệ thống để giúp hệ thống phản ứng nhanh hơn với thông tin – cụ thể là với các dự báo về khả năng lây nhiễm. Chúng tôi đã thảo luận vấn đề này với một nhóm CEO và các nhà lãnh đạo tư tưởng, bao gồm Mark Carney, cựu thống đốc Ngân hàng Anh và Ngân hàng Canada; Brenda Fitzgerald, cựu giám đốc Trung tâm Kiểm soát và Phòng ngừa Dịch bệnh ở Atlanta; và tác giả Margaret Atwood. Đến tháng 10 năm 2020, 12 CEO nổi lên đồng ý cung cấp bối cảnh cho một hệ thống được bôi dầu. Mỗi người cam kết chỉ định một trong những cấp dưới trực tiếp của họ trực tiếp làm việc trong dự án này và loại bỏ các rào cản dựa trên quy tắc bất cứ khi nào có thể. Mục tiêu của chúng tôi là thiết kế một hệ thống trong môi trường này đủ hấp dẫn để truyền cảm hứng cho các công ty khác cũng như các quan chức y tế công cộng đánh đổi một số hệ thống dựa trên quy tắc, được dán chặt của họ để lấy các quyết định dựa trên thông tin, được bôi trơn.

12 công ty lớn này – các đối tác sáng lập của Hiệp hội sàng lọc nhanh CDL – đại diện cho sản xuất, vận chuyển, dịch vụ tài chính, tiện ích, giải trí và bán lẻ. Cùng nhau, họ sử dụng hơn nửa triệu công nhân. Các CEO rất háo hức bắt đầu chế độ thử nghiệm, mở cửa nơi làm việc và giữ an toàn cho nhân viên của họ. Ngay sau khi chúng tôi bắt đầu, những nhân viên được khảo sát cho biết họ cảm thấy an tâm rằng họ và đồng nghiệp sẽ được kiểm tra trước khi vào nơi làm việc. Một trong những thành viên sáng lập của chúng tôi đã triển khai xét nghiệm đầu tiên vào ngày 11 tháng 1 năm 2021, tại một địa điểm ở trung tâm thành phố Toronto. Trong vài tháng tới, hệ thống đã được chứng minh là hoạt động tốt. Một phần nhỏ những người bị nhiễm bệnh đã được xác định nên họ không bao giờ đến gần đồng nghiệp của mình tại nơi làm việc, công nhân cho biết họ cảm thấy an toàn hơn và các nhà quản lý có thể tiếp tục để các cơ sở sản xuất hoạt động. Sau đó, chúng tôi đã tạo một cẩm nang để chia sẻ với các công ty khác và cuối cùng là với các mô hình tổ chức khác, bao gồm tổ chức phi lợi nhuận, trại hè, nhà trẻ và trường học. Cẩm nang bao gồm hướng dẫn về cách thiết lập quy trình báo cáo dữ liệu, quản lý xét nghiệm kháng nguyên nhanh, thiết lập ki-ốt xét nghiệm vật lý, đào tạo nhân viên quản lý quy trình, truyền đạt chương trình tới nhân viên và công đoàn của họ, quản lý luồng dữ



liệu, loại bỏ đã sử dụng các xét nghiệm, xử lý hậu cần liên quan đến những nhân viên có kết quả xét nghiệm dương tính, đặt hàng các xét nghiệm kháng nguyên nhanh từ chính phủ,...

Theo thời gian, chất keo ở các công ty khác bắt đầu không còn dính nữa. Ở Canada, cả trên toàn quốc và theo từng tỉnh, ban đầu người lao động có thể tự kiểm tra tại nơi làm việc dưới sự giám sát, và cuối cùng là tại nhà mà không cần đến chuyên gia chăm sóc sức khỏe.

Nó cũng cần thiết để theo dõi các bài kiểm tra. Mọi người lo sợ rằng một số người tham gia đã tiếp xúc với Covid-19 sẽ hiểu sai kết quả xét nghiệm âm tính là tín hiệu cho thấy họ sẽ không phát bệnh trong vài tuần trong khi tốt nhất là vài ngày. Thử nghiệm thường xuyên đã giảm thiểu rủi ro này, vì vậy chúng tôi cần một hệ thống dữ liệu để theo dõi xem ai đã thử nghiệm và khi nào. Tuy nhiên, các công ty sẽ chỉ chấp nhận hệ thống dữ liệu nếu nó bảo vệ quyền riêng tư của người lao động. Chúng tôi đã phát triển một hệ thống theo dõi dữ liệu đáp ứng cả mối quan tâm của nhân viên về các yêu cầu về quyền riêng tư và sức khỏe cộng đồng để đảm bảo tuân thủ.

Các chính sách của công ty đã phát triển để hỗ trợ thử nghiệm nhanh. Trong trường hợp không có lương ốm đau và các biện pháp bảo vệ khác tại nơi làm việc, người lao động đã cảnh giác tham

gia, vì vậy người sử dụng lao động cần xác định ai sẽ kiểm tra và khi nào. Các công ty cần quyết định xem xét nghiệm có được tính là một phần của ngày làm việc hay không, xét nghiệm sẽ diễn ra ở đâu và phải làm gì khi ai đó có kết quả xét nghiệm dương tính. Họ cũng cần giao trách nhiệm cho các quyết định về sức khỏe và an toàn này. Ban đầu, không có quy trình hiện có để xác định trách nhiệm của người lao động, người quản lý và chuyên gia y tế. Chất keo bắt đầu mất kết dính khi các công ty đặt niềm tin vào cảm nang của chúng tôi, trong đó bao gồm các quy trình vận hành tiêu chuẩn được chia sẻ miễn phí và cập nhật liên tục. Cuối cùng, hệ thống xét nghiệm kháng nguyên nhanh tại nơi làm việc của chúng tôi đã được sử dụng tại hơn 2.000 tổ chức ở Canada, giúp hàng nghìn trường hợp nhiễm Covid-19 không thể đến nơi làm việc và trường học khi mọi người có kết quả xét nghiệm dương tính và ở nhà. Tuy nhiên, những thách thức là rất lớn. Phải mất 6 tháng để hầu hết những người tham gia ban đầu triển khai thử nghiệm trên quy mô lớn và 1 năm trước khi hàng chục nghìn công nhân được kiểm tra thường xuyên. Công cụ dự đoán là một phần dễ dàng. Tuy nhiên, đó chỉ là một phần nhỏ của sự thay đổi cần thiết để giúp giải quyết vấn đề thông tin về Covid-19 và cho phép mọi người quay trở lại làm việc và đi học.



CÁC HỆ THỐNG ĐƯỢC BÔI TRƠN

Thông điệp từ chương này là để tận dụng lợi thế của máy dự đoán, chúng ta muốn biến quy tắc thành quyết định. Tuy nhiên, hệ thống – tập hợp các quy trình mà theo đó một việc được thực hiện – phải có khả năng thích ứng với sự thay đổi đó. Nếu một quy tắc được gắn chặt với một quy tắc khác để hệ thống trở nên đáng tin cậy, thì việc đưa ra quyết định trong hệ thống đó có thể không có kết quả.

Ở đây chúng tôi đặc biệt nhắc lại một quy tắc mà nhiều người trong chúng ta đã tuân theo vào mùa xuân năm 2020: ở nhà. Chúng ta không biết ai là người lây nhiễm và sự không chắc chắn đó đồng nghĩa với quy tắc là tránh xa mọi người.

Đổi lại, quy tắc đó kéo theo đủ loại khó khăn. Đầu tiên, nhiều người làm việc bên ngoài nhà và khách hàng cho doanh nghiệp của họ cần có thể rời khỏi nhà. Các nhà hàng, cửa hàng bán lẻ và rạp chiếu phim không thể hoạt động trong thời gian phong tỏa. Nếu không được phép ra ngoài, thì nhiều người sẽ thất nghiệp. Các chính phủ trên toàn thế giới đã tiến hành trợ cấp tiền lương và hỗ trợ kinh doanh – một giải pháp tốn kém được phát triển để bù đắp cho những thách thức mà quy tắc này tạo ra.

Thứ hai, bản thân sự giãn cách cũng có những thách thức riêng. Tình trạng này ảnh hưởng đến

trạng thái tinh thần của mọi người. Giãn cách gây khó khăn cho việc kiểm tra xem trẻ em có an toàn hay không và liệu người già có những thứ họ cần hay không. Các chuyến thăm bác sĩ chuyển sang trực tuyến hoặc bị hủy bỏ hoàn toàn. Những vấn đề này tạo ra các quy tắc mới của riêng họ. Các thành viên trong gia đình kiểm tra lẫn nhau. Nhiều trường có chính sách gọi về nhà. Các bác sĩ được yêu cầu chủ động kiểm tra bệnh nhân của họ. Ở một số nơi trên thế giới, màn hình trong nhà đã được thêm vào để đảm bảo người già được an toàn.

Đại dịch nhắc nhở chúng ta rằng, chúng ta thường bị thu hút bởi các quy tắc và các quy tắc đó mang theo sự kém hiệu quả. Đối với Covid-19, việc không có giải pháp dự đoán cho vấn đề lây nhiễm đồng nghĩa với việc chúng ta phải nhanh chóng đóng cửa toàn bộ nền kinh tế, gây thất nghiệp hàng loạt và gián đoạn đời sống xã hội cũng như trường học. Dự đoán, nếu nó có sẵn và được tích hợp vào một hệ thống được bôi trơn, hoạt động tốt, sẽ cho phép các quyết định được sử dụng để quản lý đại dịch mà không ảnh hưởng đến kết quả sức khỏe và đồng thời giảm thiểu chi phí phát sinh cho toàn xã hội. Chúng ta sẽ thảo luận vấn đề này trong Chương 6. Các quy tắc có nghĩa là chúng ta cung cấp cho mọi người sản phẩm giống nhau hoặc nền giáo dục giống nhau sẽ hạn chế các quyết định mà chúng ta đưa ra và giá trị mà chúng ta tạo ra.



Các quy tắc là mục tiêu chính của chúng tôi khi tìm kiếm các cơ hội mới để ra quyết định mà khả năng dự đoán của AI có thể mở khóa. Đối với đại dịch, đã có những công cụ để tạo ra những dự đoán cần thiết. Các xét nghiệm kháng nguyên nhanh chóng đã giúp điền vào thông tin còn thiếu về việc liệu ai đó có bị lây nhiễm hay không. Cũng có sự đổi mới trong các thủ tục phụ thuộc, như trả lương khi ốm đau và cách ly. Khi các quyết định tương tác với nhau, việc chuyển từ các quy tắc sang các quyết định đòi hỏi một hệ thống phối hợp trơn tru. Những người ra quyết định cần biết những gì người khác đang làm, điều chỉnh các mục tiêu của họ và cho phép thay đổi. Tuy nhiên, một hệ thống mới có thể gây rối đến mức bạn có thể cần phải bắt đầu sử dụng nó trong một tổ chức mới, nơi nó có thể phát triển một cách tự nhiên, thay vì cố gắng thích ứng với nó trong các tổ chức hiện có.

Nói rộng hơn, phát hiện ra sự không chắc chắn cung cấp bước đầu tiên để mở ra các quyết định mới thông qua dự đoán. Làm như vậy một cách hiệu quả đòi hỏi phải thay đổi các thủ tục phụ thuộc, như đã lưu ý trong Chương 2, xác định một giải pháp hệ thống.

NHỮNG ĐIỂM CHÍNH

- Chúng tôi đã áp dụng một quy tắc – giãn cách xã hội – để kiểm soát đại dịch. Quy tắc này rất tốn kém. Yêu cầu giãn cách dẫn đến việc đóng cửa một phần đáng kể hệ thống giáo dục, hệ thống chăm sóc sức khỏe và nền kinh tế thế giới. Kết quả là sự cách ly có những tác động đến sức khỏe tâm thần mà sẽ mất nhiều thập kỷ để hiểu đầy đủ. Nhiều quy tắc khác được xây dựng dựa trên quy tắc giữ khoảng cách xã hội, chẳng hạn như giới hạn sức chứa của nhà hàng, quy trình giao thông công cộng, phương pháp giảng dạy ở trường học, hạn chế sự kiện thể thao, trợ cấp lương và quy trình chăm sóc khẩn cấp.
- Mặc dù hầu hết mọi người coi Covid-19 là một vấn đề sức khỏe, nhưng chúng tôi lại coi đó là một vấn đề về thông tin. Đối với những người bị nhiễm bệnh, Covid-19 thực sự là một vấn đề sức khỏe. Tuy nhiên, đối với đại đa số những người không bị nhiễm bệnh, Covid-19 không phải là vấn đề sức khỏe mà là vấn đề thông tin. Đó là bởi vì không có thông tin về người bị nhiễm bệnh, chúng tôi phải tuân theo quy tắc và đối xử với mọi người như thể họ có thể bị nhiễm bệnh. Điều đó dẫn đến đóng cửa nền kinh tế. Thay vào đó, nếu chúng ta có thể đưa

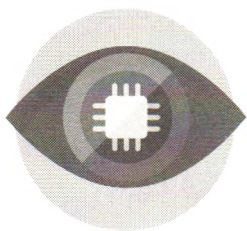


ra một dự đoán chính xác hợp lý, thì chúng ta đã có thể giải quyết vấn đề thông tin và chỉ cách ly những người có khả năng lây nhiễm cao. Các quy tắc là mục tiêu chính của chúng tôi khi tìm kiếm các cơ hội mới để ra quyết định mà khả năng dự đoán của AI có thể mở ra.

- Để tận dụng các máy dự đoán, chúng ta thường phải biến các quy tắc thành các quyết định. Tuy nhiên, hệ thống phải có khả năng thích ứng với sự thay đổi đó. Nếu một quy tắc được dán vào quy tắc khác để hệ thống trở nên đáng tin cậy, thì việc đưa ra quyết định trong hệ thống đó có thể không có kết quả. Chúng tôi mô tả một ví dụ liên quan đến Covid-19 nơi chúng tôi đã phát triển một hệ thống nhỏ nhưng có đầu, ban đầu bao gồm 12 công ty lớn, nơi các CEO chỉ đạo đội ngũ lãnh đạo cấp cao của họ đưa ra quyết định dựa trên thông tin dựa trên dự đoán về khả năng lây nhiễm của nhân viên từ xét nghiệm kháng nguyên nhanh. Điều này cho phép 12 công ty này duy trì hoạt động kinh doanh của họ trong một môi trường mà hệ thống hiện hành có thể buộc phải đóng cửa. Sự thể hiện thành công này sau đó đã thúc đẩy hơn 2.000 tổ chức khác áp dụng hệ thống này và chuyển từ các quy tắc sang các quyết định.

Chương 8

TƯ DUY HỆ THỐNG



Hằng năm, các thí sinh tập trung tại Công viên Bletchley – nơi Alan Turing từng tìm cách phá giải mật mã bằng tiếng Đức trong Thế chiến thứ hai – tham gia các thử thách với mục đích vượt qua các chương trình máy tính. Cuộc thi dựa trên trò chơi nổi tiếng (nay được gọi là bài kiểm tra Turing). Cụ thể, một người trò chuyện bằng tin nhắn qua máy tính với một thực thể vô hình. Thực thể đó có thể là một chương trình máy tính hoặc một người thật. Mỗi người đang cố gắng thuyết phục đối phương rằng họ thực sự là con người. Nếu bạn là một người cạnh tranh, thì thực tế, như tác giả Brian Christian đã nói, bạn đang cố gắng trở thành “con người nhất”. Thông thường, một con người sẽ chiến thắng, nhưng nhiều người gặp khó khăn trong việc thuyết phục đối tác của họ rằng họ là một con người.

Các cuộc đua như thế này, trong đó một người đọ sức với trí thông minh của máy móc, là trụ cột của nghiên cứu AI. Một thuật toán xác định những



gì trong ảnh tốt hơn bao nhiêu? Có phải xe tự lái ít gặp tai nạn hơn xe do người lái? AI có thể chọn ứng viên tốt hơn cho các cuộc phỏng vấn và tuyển dụng so với bộ phận nhân sự của bạn không? Máy tính có thể đánh bại nhà vô địch cờ vây thế giới không?

Các cuộc thi khơi dậy sự so sánh này là nguồn gốc của sự lo lắng về việc liệu máy móc có thay thế con người hay không. Điều thú vị là ô tô tốt hơn ngựa, nhưng ngựa vẫn đua. Và khi máy móc nhanh hơn con người ở bất kỳ khoảng cách nào, Thế vận hội vẫn tiếp tục diễn ra mà không gặp sự cố nào. Tại sao nó lại khác ngay cả khi một cỗ máy chơi cờ vây tốt hơn? Các số liệu nắm bắt được điều gì đó, nhưng sự thay thế không nhất thiết phải tuân theo.

Tuy nhiên, khi một người đang chọn người thực hiện một nhiệm vụ cụ thể, tình cảm hoặc thể thao sẽ không cắt đứt. Các số liệu được chọn để đánh giá hiệu suất dựa trên hiệu quả thuần túy và mời thay thế dựa trên chi phí. Nếu một chiếc máy có thể thực hiện nhiệm vụ đó và rẻ hơn, chắc chắn sẽ có sự thay thế. Những con ngựa vẫn có thể chạy đua, nhưng chúng không còn di chuyển mọi người xung quanh nữa. Giống như việc máy móc thay thế con người trong các công việc thể chất, có thể chúng sẽ làm điều tương tự đối với nhận thức.

Một ngành đã xuất hiện để cố gắng kiểm tra công việc của mọi người, từng nhiệm vụ một, để đánh giá xem liệu máy móc có thể thực hiện những nhiệm vụ đó trong thời đại AI hay không. Có 30 nhiệm vụ riêng biệt liên quan đến nghề nghiệp của bác sĩ X-quang (xem Hình 8-1). Chỉ một trong những nhiệm vụ này liên quan trực tiếp đến dự đoán của máy: nhiệm vụ thứ ba là giải thích kết quả của các quy trình chẩn đoán hình ảnh.

Mọi công việc đều có thể được giải nén theo cách này và được đánh giá mức độ dễ bị tổn thương đối với AI. Vào năm 2013, một nghiên cứu từ Trường Oxford Martin tuyên bố rằng, gần một nửa số việc làm ở Hoa Kỳ dễ bị ảnh hưởng bởi tự động hóa. Đây là nỗi sợ hãi hàng đầu về AI. Erik Brynjolfsson, Tom Mitchell và Daniel Rock đã đo lường “sự phù hợp với máy học” của 964 nghề nghiệp, 18.156 nhiệm vụ và 2.069 hoạt động công việc. Những công việc có rủi ro bao gồm những người lao động đưa ra nhiều dự đoán mà chúng tôi đã đánh dấu, bao gồm cả nhân viên hướng dẫn khách (người đưa ra khuyến nghị) và người cho phép tín dụng. Các công việc của nhà trị liệu xoa bóp, nhà khoa học động vật và nhà khảo cổ học vẫn an toàn. Không có gì ngạc nhiên khi các nhà kinh tế vĩ mô và lao động hàng đầu thế giới đã bày tỏ lo lắng rằng khi AI đảm nhận một số nhiệm vụ nhất định, người lao động có thể



để lại rất ít, đặc biệt là những người chưa đứng đầu trong phân phối thu nhập.

Một thập kỷ sau làn sóng AI hiện tại, máy móc đã thay thế con người trong rất ít nhiệm vụ. Chatbot đang đóng một vai trò lớn hơn trong dịch vụ khách hàng và dịch máy đang ngày càng chiếm được nhiều thị phần trong hoạt động đó. Nhưng tình trạng thất nghiệp do công nghệ vẫn chưa xuất hiện và có rất nhiều việc làm cho mọi người. Mặc dù có những AI có thể vượt trội hơn con người, nhưng trong nhiều trường hợp, những con người đó – mụn cóc và tất cả – vẫn rẻ hơn so với việc thay thế máy móc của họ. Vì vậy, trong khi các nhà kinh tế như Daron Acemoglu và Pascual Restrepo có thể lập luận rằng do chi phí vốn được trợ cấp so với lao động, đó chỉ là vấn đề thời gian, nhưng bây giờ tất cả chúng ta đều có thể thở phào.

Hình 8.1:

1. Có được lịch sử của bệnh nhân từ hồ sơ điện tử, phòng vấn bệnh nhân, báo cáo chính tả hoặc bằng cách liên lạc với các bác sĩ lâm sàng giới thiệu.
2. Chuẩn bị các báo cáo diễn giải toàn diện về các phát hiện.
3. Thực hiện hoặc giải thích kết quả của các quy trình chẩn đoán hình ảnh bao gồm chụp cộng hưởng từ (MRI), chụp cắt lớp vi tính (CT), chụp cắt lớp phát xạ positron (PET), nghiên cứu máy chạy bộ tim mạch hạt nhân, chụp nhũ ảnh hoặc siêu âm.

4. Xem lại hoặc truyền hình ảnh và thông tin bằng hệ thống lưu trữ hình ảnh hoặc hệ thống liên lạc.
5. Truyền đạt kết quả kiểm tra hoặc thông tin chẩn đoán cho bác sĩ giới thiệu, bệnh nhân hoặc gia đình.
6. Cung cấp tư vấn cho bệnh nhân chụp X-quang để giải thích các quy trình, rủi ro, lợi ích hoặc phương pháp điều trị thay thế.
7. Hướng dẫn nhân viên X-quang về các kỹ thuật, vị trí hoặc phép chiếu mong muốn.
8. Trao đổi với các chuyên gia y tế về chẩn đoán dựa trên hình ảnh.
9. Phối hợp các dịch vụ X-quang với các hoạt động y tế khác.
10. Ghi lại hiệu suất, diễn giải hoặc kết quả của tất cả các thủ tục được thực hiện.
11. Thiết lập hoặc thực thi các tiêu chuẩn để bảo vệ bệnh nhân hoặc nhân viên.
12. Phát triển hoặc giám sát các quy trình để đảm bảo kiểm soát chất lượng hình ảnh đầy đủ.
13. Nhận biết hoặc điều trị các biến chứng trong và sau các thủ thuật, bao gồm các vấn đề về huyết áp, đau, quá liều hoặc chảy máu.
14. Tham gia các hoạt động giáo dục thường xuyên để duy trì và phát triển chuyên môn.
15. Tham gia vào các hoạt động cải tiến chất lượng bao gồm các cuộc thảo luận về các lĩnh vực có nguy cơ mắc lỗi cao.
16. Thực hiện các thủ thuật can thiệp như sinh thiết dưới hướng dẫn hình ảnh, tạo hình mạch máu qua da, dẫn lưu đường mật xuyên gan hoặc đặt ống thông thận.
17. Xây dựng kế hoạch điều trị cho bệnh nhân X-quang.



18. Tiêm đồng vị phóng xạ cho bệnh nhân lâm sàng hoặc đối tượng nghiên cứu.
19. Tư vấn cho các bác sĩ khác về các chỉ định lâm sàng, hạn chế, đánh giá hoặc rủi ro của các ứng dụng chẩn đoán và điều trị của vật liệu phóng xạ.
20. Tính toán, đo lường hoặc chuẩn bị liều lượng đồng vị phóng xạ.
21. Kiểm tra và phê duyệt chất lượng chẩn đoán hình ảnh trước khi bệnh nhân xuất viện.
22. So sánh các quy trình y học hạt nhân với các loại quy trình khác, chẳng hạn như chụp cắt lớp vi tính, siêu âm, chụp cộng hưởng từ hạt nhân và chụp động mạch.
23. Chỉ đạo các kỹ thuật viên hoặc kỹ thuật viên y học hạt nhân về liều lượng, kỹ thuật, vị trí và dự báo mong muốn.
24. Thiết lập và thực thi các tiêu chuẩn bảo vệ bức xạ cho bệnh nhân và nhân viên.
25. Xây dựng kế hoạch và quy trình cho khoa y học hạt nhân.
26. Giám sát việc xử lý các vật liệu phóng xạ để đảm bảo tuân thủ các quy trình đã thiết lập.
27. Kê đơn hạt nhân phóng xạ và liều dùng cho từng bệnh nhân.
28. Xem xét các yêu cầu thủ tục và lịch sử y tế của bệnh nhân để xác định khả năng áp dụng các thủ tục và đồng vị phóng xạ sẽ được sử dụng.
29. Dạy y học hạt nhân, chẩn đoán hình ảnh hoặc các chuyên ngành khác ở trình độ sau đại học.
30. Kiểm tra các dụng cụ đánh giá liều lượng và đồng hồ khảo sát để đảm bảo chúng hoạt động bình thường.

Nguồn: O*NET, <https://www.onetonline.org/link/summary/29-1224.00>.

From "29-1224.00 – Radiologists", by the National Center for O*NET Development. Used under the CC BY 4.0 license.

Tuy nhiên, có một góc nhìn khác về cách AI có thể thay đổi công việc của chúng ta và cách mọi thứ được sản xuất. Giáo sư Tim Bresnahan của Stanford đã lập luận rằng, toàn bộ quá trình phân tích tiềm năng của AI thành các nhiệm vụ mà AI có thể thực hiện đã bỏ qua một yếu tố quan trọng. Đó là yếu tố thực sự thúc đẩy việc áp dụng triệt để các công nghệ mới trong quá khứ: sự thay đổi trên toàn hệ thống. Bresnahan lập luận rằng chúng ta đã thấy điều này ở những nơi tích cực triển khai AI, như Amazon, Google, Facebook và Netflix:

Thay thế cấp độ nhiệm vụ không đóng vai trò gì trong các ứng dụng công nghệ AI này. Những ứng dụng ban đầu rất có giá trị này không phải là những ứng dụng trong đó lao động đang thực hiện một nhiệm vụ và được thay thế bằng tư bản. Các nhà quan sát tập trung vào sự thay thế cấp độ nhiệm vụ, không phải vì nó xảy ra, mà vì định nghĩa về AI chung bao gồm “các nhiệm vụ thường do con người thực hiện”. Cho đến khi AI nói chung được thương mại hóa, điều này khó có thể xảy ra trong tương lai gần, phân tích nên tập trung vào khả năng và ứng dụng của các công nghệ AI thực tế. Mặc dù có thể có một số thay thế ở cấp độ nhiệm vụ trong



tương lai, nhưng nó không liên quan đến đề xuất giá trị của công nghệ AI.

AI tại các công ty công nghệ hàng đầu không phải chỉ là một dự án thử nghiệm. Nó bao gồm các hệ thống sản xuất quy mô lớn tạo ra doanh thu hàng tỉ đô la. Nó không được xây dựng theo từng nhiệm vụ, trong đó AI tham gia vào một số nhiệm vụ nhất định. Thay vào đó, các công ty công nghệ lớn đã xây dựng các hệ thống hoàn toàn mới. Việc áp dụng thành công AI thể hiện cái mà chúng ta sẽ gọi ở đây là tư duy hệ thống. Nó trái ngược với tư duy nhiệm vụ ở chỗ nó nhìn thấy tiềm năng lớn hơn của AI và nhận ra rằng để tạo ra giá trị thực, các hệ thống ra quyết định, bao gồm cả dự đoán của máy móc và con người, sẽ cần được tái tạo và xây dựng. Một số tổ chức đã nhận ra và thực hiện điều này, nhưng lịch sử cho chúng ta biết rằng so với những doanh nghiệp lâu năm, những người mới tham gia vào một ngành sẽ dễ dàng thực hiện thay đổi trên toàn hệ thống để tận dụng các công nghệ mới như dự đoán AI. Ô tô là phương tiện di chuyển tốt hơn ngựa, nhưng ô tô cần trạm xăng, hệ thống đường sá tốt hơn và một bộ luật hoàn toàn mới.

GIÁ TRỊ VÀ GIÁ (CHI PHÍ)

Các nhà kinh tế học có xu hướng tập trung vào chi phí. Chúng tôi, cũng với tư cách là các nhà kinh tế học, cảm thấy tội lỗi về điều đó như bất kỳ ai. Tiền đề của cuốn sách đầu tiên của chúng tôi, *AI trong cuộc cách mạng công nghệ 4.0*, là những tiến bộ của AI sẽ làm giảm đáng kể chi phí dự đoán, dẫn đến việc sử dụng nó được mở rộng. Ở cuốn sách đầu tiên đó, chúng tôi cho rằng những ứng dụng ban đầu của AI cũng chính là các dự báo, chẳng hạn như dự báo doanh số bán hàng hay dự báo thời tiết một cách rõ ràng, hoặc ngầm định trong việc phân loại ảnh và ngôn ngữ. Tuy nhiên, đến cuốn sách này chúng tôi nhận ra rằng cơ hội thực sự sẽ là những ứng dụng hay cách dùng mới cơ hội mới được triển khai khi chi phí dự báo giảm xuống mức thấp vừa đủ.

Đồng thời, qua quá trình làm việc với các công ty khởi nghiệp AI tại Creative Destruction Lab, chúng tôi nhận thấy rằng thông điệp chào hàng ban đầu của các doanh nghiệp này là cách một hệ thống AI như vậy sẽ có giá trị như thế nào đối với các tổ chức vì nó sẽ giúp họ tiết kiệm chi phí tuyển dụng. Khi định giá các sản phẩm AI đó, họ đã sử dụng tư duy chi phí và tính toán mức lương tiết kiệm được cũng như các chi phí khác và định giá máy thay thế của chính họ dựa trên điều đó.



Thông thường, đó là một vụ chào bán khó khăn. Nếu bạn đến một doanh nghiệp và nói với họ rằng bạn có thể giúp họ tiết kiệm 50.000 đô la mỗi năm cho chi phí lao động nếu họ loại bỏ một công việc này, khi đó sản phẩm AI của bạn sẽ hoàn thành việc loại bỏ phần việc đó một cách tốt hơn. Tuy nhiên, thực tế là sản phẩm của họ chỉ loại bỏ một nhiệm vụ trong công việc của một người và điều đó là không đủ để tiết kiệm cho khách hàng tiềm năng của họ bất kỳ chi phí lao động đáng kể nào.

Những quảng cáo bán hàng tốt hơn sẽ không tập trung vào sự thay thế mà tập trung vào *giá trị*. Những quảng cáo bán hàng này chứng minh làm thế nào một sản phẩm AI có thể cho phép các doanh nghiệp tạo ra nhiều lợi nhuận hơn bằng cách cung cấp các sản phẩm chất lượng cao hơn cho khách hàng của họ. Điều này có lợi là không phải chứng minh rằng AI của họ có thể thực hiện một nhiệm vụ cụ thể với chi phí thấp hơn con người. Và nếu điều đó cũng làm giảm sức đề kháng nội bộ đối với việc áp dụng AI, thì điều đó chỉ khiến nhiệm vụ bán hàng của họ trở nên dễ dàng hơn. Vấn đề ở đây là cách tiếp cận nâng cao giá trị cho AI, thay vì cách tiếp cận tiết kiệm chi phí, có nhiều khả năng tìm thấy lực kéo thực sự cho việc áp dụng AI.

Chúng ta đã thấy sự phân đôi tương tự trong các cuộc cách mạng công nghệ trước đây. Đối với điện, mà chúng ta đã thảo luận trong Chương 1, việc thay thế hơi nước trong sản xuất diễn ra chậm chạp và mất hàng thập kỷ. Các nhà máy hiện tại chỉ sử dụng điện nếu chi phí thấp hơn hơi nước. Đó là một sự khó bán đối với các nhà máy đã được thiết kế để chạy bằng hơi nước. Ngược lại, một khi các nhà sản xuất nhận ra rằng điện mang lại cho họ cơ hội thiết kế lại các nhà máy thành các cơ sở lắp đặt bằng phẳng lớn bên ngoài giá thuê thành phố đắt đỏ, thì sẽ có nhiều mối quan tâm hơn đến việc đầu tư vào các nhà máy mới hứa hẹn năng suất cao hơn đáng kể do thiết kế mới của họ. Thật vậy, ô tô điện từng được cho là một công nghệ hứa hẹn hơn ô tô chạy bằng xăng. Hóa ra, xăng cho phép ô tô di chuyển quãng đường dài hơn, và điều đó cuối cùng đã thắng thế, ít nhất là cho đến khi công nghệ pin phát triển vào đầu thế kỷ XXI. Trong trường hợp các nhà máy được thiết kế lại, giá trị điện được nâng cao, trong khi ở trường hợp khác, giao thông vận tải thì không. Giá trị đã thắng.

Quan trọng là, việc áp dụng một hệ thống mới đòi hỏi phải thay thế một hệ thống hiện có. Tính toán chi phí thuần túy sẽ hiếm khi thúc đẩy sự thay thế như vậy. Có những chi phí chuyển đổi trong việc xây dựng các hệ thống mới, và nếu điều tốt



nhất bạn làm là tiết kiệm một phần chi phí của hệ thống hiện có, thì điều đó khó có thể chứng minh được giá trị. Thay vào đó, nếu hệ thống mới thực hiện một điều gì đó mới – nghĩa là dẫn đến các cơ hội tạo ra giá trị mới – thì đó chính là điều sẽ thúc đẩy việc áp dụng.

THÁCH THỨC CỦA SỰ THAY ĐỔI TOÀN HỆ THỐNG

Đã có rất nhiều bài viết về tiềm năng của AI trong y học. Trong cuốn sách *Deep Medicine: How Artificial Intelligence Can Make Health Care Human Again*, Eric Topol giải thích cách AI có thể cải thiện chẩn đoán, giúp các bác sĩ có thời gian chăm sóc bệnh nhân và hiểu nhu cầu của họ. Các ứng dụng AI trong y học bao gồm chẩn đoán bệnh, phẫu thuật tự động, theo dõi bệnh nhân tại nhà, điều trị cá nhân hóa, khám phá và tái sử dụng thuốc.

Những cơ hội này đã tạo ra những lo lắng về “mặt tối của AI trong chăm sóc sức khỏe”, nơi AI cạnh tranh với MD để chẩn đoán. Có lẽ lý do *Deep Medicine* có sự ảnh hưởng như vậy là vì Topol hiểu hệ thống chăm sóc sức khỏe (ông là bác sĩ tim mạch và giáo sư y học phân tử tại Scripps Research), ông hiểu rõ AI (ông đã đầu tư đáng kể vào việc tìm hiểu các khả năng và hạn chế của công nghệ này

vì nó liên quan đến lĩnh vực chăm sóc sức khỏe), và ông cũng là bậc thầy giao tiếp và một dịch giả (ông là người sáng lập và giám đốc của Viện Dịch thuật Nghiên cứu Scripps). Chỉ có một vấn đề. Ông không phải là một nhà kinh tế học. Vì vậy, ông không viết về hành vi của con người, hoặc có lẽ ông cũng tin rằng bác sĩ là những bản năng nguyên thủy như vậy. Mối quan tâm của chúng tôi là nếu chúng ta chỉ đơn giản đưa các công nghệ AI mới vào hệ thống chăm sóc sức khỏe hiện có, các bác sĩ có thể không có động cơ sử dụng chúng, tùy thuộc vào việc họ sẽ tăng hay giảm mức thù lao mà mình được nhận (được quyết định bởi phí dịch vụ hoặc thù lao dựa trên khối lượng công việc).

Topol tin rằng nếu AI tiết kiệm thời gian cho bác sĩ, thì bác sĩ sẽ dành thêm thời gian đó để nói chuyện và kết nối với bệnh nhân của họ. Bằng chứng hoàn toàn không rõ ràng rằng các công cụ nâng cao năng suất trước đây dành cho bác sĩ đã làm tăng thời gian họ dành cho việc kết nối với bệnh nhân của mình. Nó có thể là ngược lại. Nếu AI tăng năng suất của các bác sĩ, họ có thể dành ít thời gian hơn cho từng bệnh nhân mà không làm giảm thu nhập của họ. Để đạt được những mục tiêu xứng đáng mà Topol mong muốn, chúng ta cần nhiều hơn những công nghệ AI mới. Chúng ta cần một hệ thống mới, bao gồm các biện pháp khuyến khích,



đào tạo, phương pháp và văn hóa mới để các bác sĩ sử dụng các công cụ công nghệ của họ theo cách mà cuốn sách của Topol mong muốn.

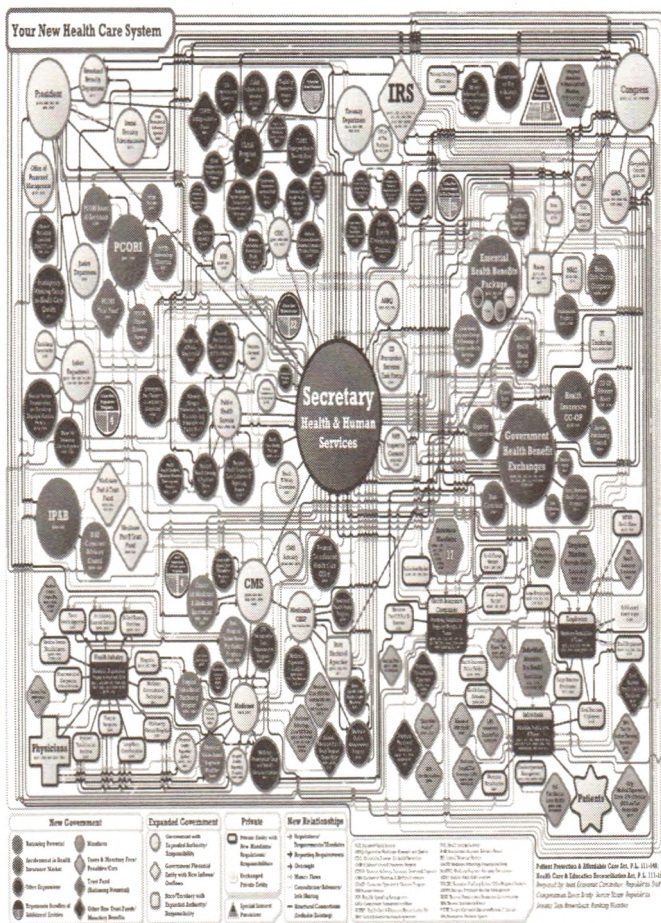
Vì vậy, không có gì đáng ngạc nhiên, mặc dù có nhiều cơ hội để tăng cường y học với AI, như được mô tả trong *Deep Medicine* và các nơi khác, y tế không đi đầu trong việc áp dụng AI. Trong một nghiên cứu về công việc AI và máy học trong các ngành, y tế xếp hạng gần cuối. Vào cuối năm 2019, tỷ lệ công việc liên quan đến AI trong ngành y tế nhỏ hơn mọi ngành khác ngoại trừ xây dựng, nghệ thuật và giải trí. Ngay cả các dịch vụ lưu trú và ăn uống, vận chuyển và kho bãi cũng thu hút nhiều lao động có kỹ năng liên quan đến AI hơn. Một lý do có thể là hệ thống y tế đặc biệt phức tạp. Hình 8-2 cho thấy hình ảnh hệ thống y tế Mỹ năm 2010 theo chương trình Obamacare của Tổng thống đương nhiệm.

Với rất nhiều quyết định phối hợp, giải pháp điểm và giải pháp ứng dụng sẽ vẫn hạn chế về giá trị nếu không kèm theo những thay đổi khác. Có thể dễ dàng hình dung cách AI thúc đẩy các phương pháp điều trị được cá nhân hóa, nhưng kèm theo đó là sẽ rất nhiều người ở các bộ phận khác nhau sẽ thay đổi công việc mà họ đang làm để thực hiện điều đó (ví dụ: thu thập thêm dữ liệu

cá nhân, cung cấp dịch vụ chăm sóc được cá nhân hóa hơn). Các giải pháp điểm AI trong lĩnh vực chăm sóc sức khỏe thường đưa ra những dự báo mà không ai có thể sử dụng (ví dụ: vì không có sẵn các lựa chọn điều trị). Các giải pháp ứng dụng AI thường cho phép các hành động mà không ai có thể thực hiện (ví dụ: vì các quy tắc trách nhiệm pháp lý gây khó khăn cho việc áp dụng) hoặc muốn thực hiện (ví dụ: vì chúng không phù hợp với hệ thống lương thưởng). Thách thức không lớn đến mức khiến giá trị của các dự báo không đủ lớn để đánh đổi hoặc các hành động trở nên vô ích; khó khăn là không dễ dàng để các bộ phận của hệ thống hoạt động đồng thời.



Hình 8-2: Sơ đồ hệ thống chăm sóc sức khỏe của Mỹ năm 2010



Nguồn: Joint Economic Committee, Republican Staff, "Understanding the Obamacare Chart," July 2010,
https://www.jec.senate.gov/public/_cache/files/96b779aa-6d2e-4c41-a719-24e865cac666/understanding-the-obamacare-chart.pdf.

Để thực hiện điều này, cần phải thay đổi hệ thống. Có nhiều nhận định về tiềm năng và sự cần thiết của một hệ thống chăm sóc sức khỏe mới được AI hỗ trợ. Nếu AI tham gia chẩn đoán, các quy tắc về việc ai được phép làm gì trong chăm sóc sức khỏe sẽ thay đổi. Với việc máy móc thực hiện chẩn đoán, vai trò chính của bác sĩ có thể là chăm sóc sức khỏe cho con người. Điều này sẽ yêu cầu các loại thay đổi đồng thời khác. Trong trường y, sinh viên không còn được yêu cầu ghi nhớ các dữ liệu thực tế và sẽ không còn được đánh giá dựa trên khả năng hiểu biết về sinh học, thể hiện qua điểm số cao trong các bài kiểm tra. Những kỹ năng này có thể không cải thiện nhiều sau một thập kỷ học tập, vì vậy, thay vào đó, các bác sĩ phải tiếp xúc và chữa trị cho bệnh nhân có thể chỉ cần một thứ gì đó như bằng đại học. Đổi lại, điều đó cũng sẽ đòi hỏi những thay đổi lớn về quy định đối với việc ai được phép và loại hình dịch vụ chăm sóc sức khỏe nào sẽ được cung cấp. Có thể chăm sóc bệnh nhân trở thành vai trò chính của dược sĩ. Có lẽ nhân viên xã hội chuyển sang đảm nhận công việc đã từng là của bác sĩ. Trong Chương 18, chúng ta sẽ tìm hiểu quy trình phát triển giải pháp hệ thống AI trong chăm sóc sức khỏe và tầm nhìn về những gì một hệ thống như vậy trong y học cấp cứu có thể bao gồm. Có lẽ không điều nào trong



số này sẽ xảy ra, bởi vì có quá nhiều thứ cần phải thay đổi.

AI cũng có thể thay đổi hệ thống y tế toàn cầu. Ngân hàng Thế giới đã nhấn mạnh cách các công nghệ như AI có thể tạo sân chơi bình đẳng giữa các quốc gia. Sự kết hợp giữa theo dõi bệnh nhân từ xa và chẩn đoán bằng máy có thể cải thiện việc chăm sóc sức khỏe ở những địa phương khó tiếp cận.

Các bác sĩ tim mạch của Cameroon làm việc tại các bệnh viện ở khu vực thành thị, tuy nhiên nhiều người trong số 25 triệu cư dân của nước này sống xa các bệnh viện này. Hầu hết mọi người mắc bệnh tim mạch không bao giờ được chẩn đoán. Nhà phát minh người Cameroon Arthur Zang đã phát triển Cardiopad để giải quyết vấn đề này. Cardiopad là một công cụ cho phép chụp tâm đồ từ xa. Không cần bác sĩ tim mạch địa phương thực hiện xét nghiệm. Công cụ này đã cho phép chẩn đoán từ xa cho hàng nghìn bệnh nhân, nhưng nó vẫn yêu cầu bác sĩ tim mạch đưa ra chẩn đoán cuối cùng. 20 bác sĩ tim mạch đã làm việc với Cardiopad vào năm 2020 đã bị choáng ngợp. Cả nước chỉ có 60 bác sĩ tim mạch. Cardiopad đã giải quyết được vấn đề truy cập vào điện tâm đồ, nhưng nó không giải quyết được vấn đề chẩn đoán theo tỷ lệ, đòi hỏi máy móc có thể tiến hành chẩn đoán và con người

sẽ chấp nhận chẩn đoán. Hơn nữa, một khi hàng nghìn người được chẩn đoán, thì cần phải có cơ sở hạ tầng để điều trị bệnh. Hiện tại, số lượng chẩn đoán và điều trị được gắn liền với nhau. Bằng cách phân bổ chẩn đoán, hệ thống hiện tại tạo ra ít sự không chắc chắn về số lượng bệnh nhân cần được điều trị. Cardiopad là một phần của hệ thống đó. Nó giải quyết vấn đề khoảng cách giữa bác sĩ tim mạch và bệnh nhân. Những cải tiến quy mô lớn trong chăm sóc tim mạch ở Cameroon đòi hỏi phải thay đổi cách hoạt động của chẩn đoán và phát triển các hướng điều trị mới, từ đó tận dụng lợi thế của những thay đổi đối với chẩn đoán. Những giải pháp hệ thống đó vẫn cần được phát triển. Cho đến lúc đó, Cardiopad sẽ tiếp tục cải thiện việc chăm sóc sức khỏe cho hàng nghìn bệnh nhân ở Cameroon, nhưng nó sẽ không cho phép cải thiện sức khỏe tim mạch ở cấp độ dân số.

IBM là một công ty lớn hơn nhiều so với Cardiopad. Watson ra đời với kỳ vọng sẽ trở thành mũi nhọn của công ty, hướng đến tác động to lớn đối với lĩnh vực chăm sóc sức khỏe. Thế nhưng, sản phẩm này đã không thành công như kỳ vọng. Đã có vấn đề về dữ liệu và rủi ro thực sự khi dự đoán sai. Một đối tác của IBM đã nói rằng, “Chúng tôi nghĩ rằng nó sẽ dễ dàng, nhưng hóa ra lại thực sự rất khó”.



Việc xác định các nhiệm vụ có thể áp dụng cho các giải pháp điểm AI và giải pháp ứng dụng AI là không khó. IBM đã tìm ra cách. Thế nhưng lại quá khó để đưa các giải pháp đó vào hệ thống hiện tại một cách hiệu quả và một hệ thống mới vẫn chưa xuất hiện.

XOAY LẬT

Khi bắt đầu Thời gian Chuyển đổi (The Between Times), các cơ hội áp dụng AI thường mang tính phản ứng. Một nhà cung cấp tiếp cận bạn bằng một AI mới với chức năng dự báo điều gì đó có liên quan đến tổ chức của bạn. Hoặc bạn có thể yêu cầu các nhóm nội bộ của mình tiến hành phân tích quy trình làm việc của họ để xem liệu có cơ hội nào sử dụng AI để hỗ trợ một hoặc nhiều nhiệm vụ hay không. Đây là một cách tiếp cận tốt, nhưng cũng là một cách tiếp cận tốt để thực hiện một điều: tìm giải pháp điểm cho AI.

Đến giờ, chúng tôi hy vọng đã thuyết phục được bạn rằng việc tìm kiếm cơ hội áp dụng AI có khả năng biến đổi tiềm ẩn những giá trị to lớn. Điều này đòi hỏi phải kiểm tra toàn bộ hệ thống và hiểu cách AI có thể tạo điều kiện thay đổi những hệ thống đó tốt hơn. Chính ở đó, những cơ hội quan trọng nhất dành cho AI nằm ở đó.

Nhưng thật dễ dàng để nói rằng bạn nên có tư duy hệ thống. Điều khó hơn là thực sự phát triển một cái. Theo định nghĩa, những điều không chắc chắn tiềm ẩn rất khó phát hiện. Các quy tắc kết dính hệ thống hiện tại lại với nhau rất khó loại bỏ. Bước đầu tiên là nhận ra rằng thay đổi hệ thống là cần thiết. Sự công nhận đó đã xảy ra trong một lĩnh vực của nền kinh tế. Chúng tôi chuyển sang nó tiếp theo.

NHỮNG ĐIỂM CHÍNH

- Tư duy ở cấp độ nhiệm vụ hiện là cách tiếp cận chủ đạo để lập kế hoạch đưa AI vào tất cả các lĩnh vực của nền kinh tế. Ý tưởng chính là xác định các nhiệm vụ cụ thể trong một nghề nghiệp dựa trên các dự đoán mà AI, chứ không phải con người, có thể tạo ra chính xác hơn, nhanh hơn hoặc rẻ hơn. Các nhà lãnh đạo doanh nghiệp, chuyên gia tư vấn quản lý và các học giả phần lớn đều hội tụ về phương pháp này.
- Sự thống trị của tư duy ở cấp độ nhiệm vụ là điều đáng ngạc nhiên vì việc triển khai AI ấn tượng nhất cho đến nay không phải là sự thay thế lao động của con người ở cấp độ nhiệm vụ, mà là các thiết kế ở cấp độ hệ thống mới chỉ có thể thực hiện được nhờ khả năng dự đoán mà AI

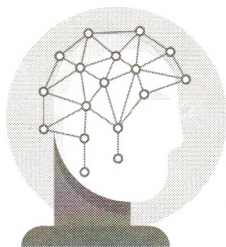


hiện có (ví dụ: Amazon, Google, Netflix, Meta, Apple). Suy nghĩ ở cấp độ nhiệm vụ dẫn đến các giải pháp điểm thường được thúc đẩy bởi tiết kiệm chi phí dựa trên thay thế lao động. Ngược lại, tư duy ở cấp độ hệ thống dẫn đến các giải pháp hệ thống thường được thúc đẩy bởi việc tạo ra giá trị chứ không phải tiết kiệm chi phí.

- Có nhiều ứng dụng cho AI trong chăm sóc sức khỏe: chẩn đoán bệnh, phẫu thuật tự động, theo dõi bệnh nhân tại nhà, điều trị cá nhân hóa, khám phá và tái sử dụng thuốc. Tuy nhiên, cho đến nay, hệ thống chăm sóc sức khỏe chỉ nhận được lợi ích cận biên từ AI. Một phần trong số đó là do thời gian cần thiết để được phê duyệt theo quy định, nhưng phần lớn là do những lợi ích bị tắt tiếng từ việc sử dụng các giải pháp điểm AI trong hệ thống chăm sóc sức khỏe hiện có. Cần có một cách tiếp cận giải pháp hệ thống để tận dụng tối đa sức mạnh của AI trong chăm sóc sức khỏe. Chúng ta phải bắt đầu với một phương tiện trống và tưởng tượng cách sức khỏe của mọi người có thể được phục vụ tốt nhất trong một hệ thống được thiết kế mới có quyền truy cập vào công nghệ dự đoán mạnh mẽ mới. Điều đó có nghĩa là suy nghĩ lại về đào tạo, thủ tục giao hàng, bồi thường, quyền riêng tư và trách nhiệm pháp lý. Điều đó có nghĩa là áp dụng tư duy hệ thống.

Chương 9

HỆ THỐNG VĨ ĐẠI NHẤT



“AlphaFold là thành tựu quan trọng nhất trong AI – Chưa từng có”, một bài báo trên *Forbes* tuyên bố. Phải thừa nhận rằng lối thoát đó được thúc đẩy để cường điệu hóa. Không phải đối với tạp chí học thuật nghiêm túc *Nature*, nhưng nó cũng tuyên bố tương tự rằng “sẽ không thay đổi mọi thứ”.

AlphaFold dự đoán cấu trúc protein. Protein là khối xây dựng của sự sống, chịu trách nhiệm cho hầu hết những gì xảy ra bên trong tế bào. Cách thức hoạt động của protein và tác dụng của nó được xác định bởi hình dạng 3 chiều của nó. Trong sinh học phân tử, “cấu trúc là chức năng”.

Các nhà khoa học từ lâu đã tự hỏi cấu trúc này đến từ đâu, làm thế nào các bộ phận cấu thành của protein ánh xạ ra khỏi nhiều vòng xoắn và nếp gấp của hình dạng cuối cùng của nó. Trong nhiều thập kỷ, các thí nghiệm trong phòng thí nghiệm là cách chính để có được cấu trúc protein tốt.

Với AlphaFold là công cụ dự đoán cấu trúc protein từ chuỗi axit amin của chúng, các nhà khoa



học có thể khám phá những sự thật mới về các khối xây dựng của sự sống. Các nhà nghiên cứu tại Đại học California, San Francisco, đã sử dụng AlphaFold để khám phá những chi tiết chưa biết trước đây về một loại protein quan trọng của SARS-CoV-2, giúp thúc đẩy sự phát triển của phương pháp điều trị Covid-19.

Các nhà khoa học tại Đại học Colorado, Boulder, đã dành nhiều năm cố gắng xác định cấu trúc của một loại protein vi khuẩn cụ thể để giúp chống lại tình trạng kháng kháng sinh. Họ đã sử dụng AlphaFold để tìm hiểu cấu trúc trong 15 phút. Một phòng thí nghiệm khác lưu ý rằng AlphaFold đã cung cấp cho nó cấu trúc protein trong 30 phút, sau khi thử các công cụ khác trong một thập kỷ. Người đứng đầu phòng thí nghiệm đó, Andrei Lupas, lưu ý: “Điều này sẽ thay đổi y học. Nó sẽ thay đổi nghiên cứu. Nó sẽ thay đổi kỹ thuật sinh học”.

SỰ SÁNG TẠO RA MỘT PHƯƠNG THỨC SÁNG TẠO

Nếu chúng ta phải chỉ ra một lĩnh vực mà AI có nhiều tiềm năng nhất để chuyển đổi nền kinh tế, thì đó là lĩnh vực ngược dòng so với hầu hết các hoạt động kinh doanh thông thường: trong hệ thống đổi mới và phát minh. Tin tốt là đây cũng là lúc những người thực hành AI dường như nhận ra sự cần thiết

của tư duy hệ thống trong việc áp dụng AI của họ. Nếu AlphaFold muốn thay đổi y học, nghiên cứu và kỹ thuật sinh học, thì đó không chỉ là một giải pháp điểm.

Tại Hội nghị kinh tế về trí tuệ nhân tạo NBER đầu tiên mà chúng tôi tổ chức vào năm 2017, các nhà kinh tế Iain Cockburn, Rebecca Henderson và Scott Stern đã lập luận rằng AI “có tiềm năng thay đổi chính quá trình đổi mới”. Khoa học dữ liệu đã là một thành phần của quy trình khoa học. AI sẽ làm cho khoa học dữ liệu đó tốt hơn, nhanh hơn và rẻ hơn. Nó sẽ cho phép các loại dự đoán mới. Điều này mở ra những hướng nghiên cứu mới và cải thiện năng suất trong phòng thí nghiệm. Là một cách mới để tạo ra sản phẩm, thay vì cải tiến một sản phẩm cụ thể, tác động kinh tế của các công cụ nghiên cứu không chỉ giới hạn ở khả năng giảm chi phí đổi mới. Thay vào đó, họ thay đổi hướng dẫn cho sự đổi mới.

Kính hiển vi cũng là một phương pháp phát minh mới. Từ kính hiển vi đã xuất hiện lý thuyết mầm bệnh. Từ lý thuyết này đã ra đời phần lớn y học hiện đại. Lý thuyết mầm bệnh khiến việc chiến đấu với virus và vi khuẩn trở nên khả thi. Nó cũng thay đổi các khía cạnh khác của y học. Phẫu thuật trở nên hữu ích về mặt y tế. Sinh con trở nên an toàn hơn. Bệnh viện trở thành nơi người ta khỏe hơn là nơi người ta chết. Tuy nhiên, AI không chỉ là

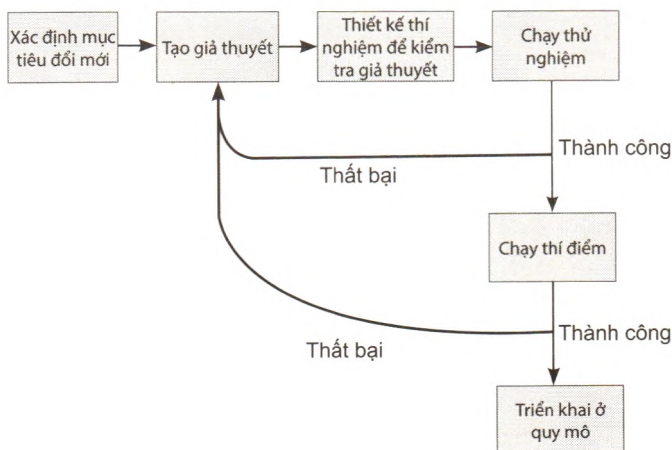


một phương pháp phát minh. Nó cũng có vẻ là một công nghệ có mục đích chung. Đây là lý do tại sao AI yêu cầu thay đổi hệ thống và tại sao tiềm năng của AI cho thấy một nghịch lý.

NHỮNG HỆ THỐNG SÁNG TẠO MỚI

Đổi mới liên quan đến một quá trình thử và sai có cấu trúc. Hình 9-1 cho thấy quá trình này hoạt động như thế nào trong nhiều ngữ cảnh. Tổ chức nghiên cứu xác định một mục tiêu và đưa ra các giả thuyết về cách đạt được mục tiêu đó. Sau đó, nó thiết kế và chạy một thử nghiệm để kiểm tra giả thuyết hàng đầu.

Hình 9-1: Quy trình sáng tạo



Thường thì thí nghiệm đó thất bại. Hy vọng rằng thất bại đó sẽ tạo ra những học hỏi và những giả thuyết mới, một trong số đó sẽ dẫn đến một thử nghiệm thành công. Sau đó, tổ chức sẽ chạy thử nghiệm và nếu thử nghiệm đó thành công, đôi khi có thể được triển khai trên quy mô lớn.

Quá trình này áp dụng cho các hệ thống đổi mới tương đối đơn giản như công cụ đề xuất nội dung và các hệ thống phức tạp hơn như hệ thống phát triển thuốc. Chúng tôi xem xét mỗi lượt.

Một hệ thống đề xuất, như xảy ra trên Amazon và Spotify, có thể có mục tiêu tối đa hóa mức độ tương tác của người dùng hoặc tăng doanh số bán hàng. Các giáo sư trường kinh doanh Dokyun Lee và Kartik Hosanagar đã làm việc với một nhà bán lẻ trực tuyến lớn để kiểm tra xem các chi tiết của công cụ giới thiệu ảnh hưởng đến doanh số bán hàng như thế nào. Họ so sánh các công cụ đề xuất “bộ lọc cộng tác” dựa trên đề xuất “những người đã mua cái này cũng đã mua cái này” với các công cụ dựa trên sự phù hợp với chuỗi từ khóa đã nhập. Công ty mà họ làm việc có một giả thuyết hoặc lý thuyết rằng công cụ giới thiệu sẽ tăng doanh số bán hàng. Sau đó, họ chạy thử nghiệm và hiển thị ngẫu nhiên cho một số người dùng công cụ đề xuất mới và những người khác hiển thị công cụ tìm kiếm dựa trên từ khóa cũ hơn. Trong hầu hết các danh mục



sản phẩm, số lượng mua hàng cao hơn nhờ công cụ mới. Sự can thiệp đã có hiệu quả, vì vậy công ty đã quyết định triển khai nó.

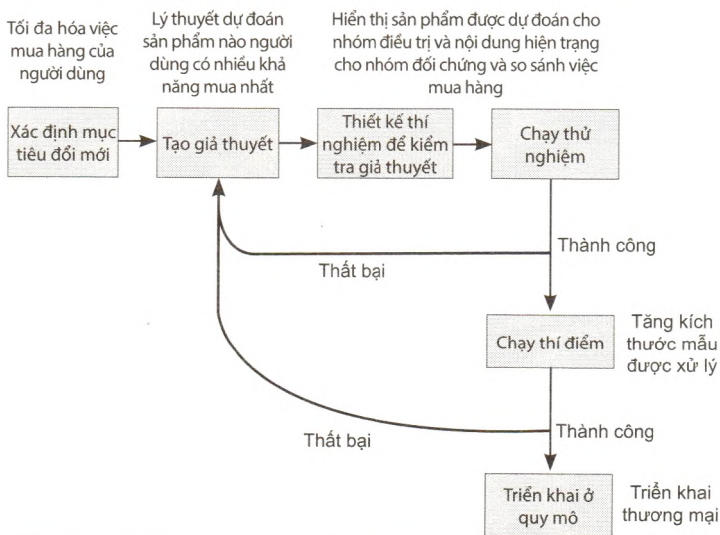
Công cụ giới thiệu cũng đề xuất một số cơ hội mới. Nó đã thay đổi cách phân phối sản phẩm, với nhiều người mua các sản phẩm phổ biến giống nhau và ít mua các sản phẩm ít phổ biến hơn hoặc sản phẩm có đuôi dài. Việc phân phối thay đổi vì khuyến nghị rằng “những người đã mua sản phẩm này cũng đã mua sản phẩm này” làm tăng doanh số bán các sản phẩm phổ biến nhất. Vào thời điểm đó, trong sách, mọi người cũng đã mua *The Da Vinci Code*, vì vậy nó sẽ được giới thiệu cho mọi người. Công ty đã quyết định chống lại những cải tiến sẽ làm tăng lợi nhuận trong suốt quá trình phân phối. Công cụ đề xuất là một giải pháp điểm AI phù hợp với quy trình làm việc hiện có. Việc đảo ngược tình trạng giảm mua hàng trong thời gian dài sẽ cần có sự phối hợp với các bộ phận khác. Các kỹ sư lo lắng rằng sẽ có rất nhiều công việc để thực hiện. Họ sợ rằng thuật toán mới sẽ gây ra những hậu quả không mong muốn. Và quan trọng nhất, “các kỹ sư không muốn phá vỡ hệ thống hiện có”. Sự thay đổi ở cấp độ hệ thống là quá nhiều so với những lợi ích dự kiến.

* Một AI tập trung vào đổi mới có thể thay đổi quá trình này. Thay vì đưa ra giả thuyết về loại công cụ đề xuất tốt nhất, AI có thể sử dụng dữ liệu hiện có để

tạo ra hàng nghìn công cụ đề xuất khả thi (xem Hình 9-2). Khi bước tạo giả thuyết này nhanh hơn, thì nó sẽ cho phép đổi mới các biện pháp có tác động mạnh hơn so với việc mua hàng ngắn hạn như bán hàng rời bỏ hoặc bán hàng dài hạn. Với các giả thuyết tốt hơn, có thể chạy nhiều thử nghiệm hơn với năng suất cao hơn và lợi tức đầu tư lớn hơn.

Hơn nữa, nếu dự đoán đủ tốt, có thể bỏ qua giai đoạn thử nghiệm. Một trong số chúng tôi (Ajay), cùng với các nhà kinh tế học John McHale và Alex Oettl, đã mô hình hóa ý tưởng này trong lĩnh vực khám phá vật liệu mới. Dự đoán tốt hơn ở giai đoạn phát triển giả thuyết có thể dẫn đến một hệ thống hoàn toàn mới.

Hình 9-2: Đơn giản: Thử nghiệm AB cho đề xuất máy





Sự thay đổi hệ thống tương tự có thể xảy ra đối với hệ thống R&D phức tạp hơn trong quá trình phát triển thuốc. Mục tiêu là thiết kế các loại thuốc mới. Quá trình này tương tự: đưa ra giả thuyết, thử nghiệm, thử nghiệm và triển khai.

Tất nhiên, những cơ hội này phụ thuộc vào sự tiến bộ đáng kể trong dự đoán AI. AlphaFold có thể chính xác là bước tiến đó. Chỉ riêng dự đoán của AlphaFold sẽ không thay đổi y học, nghiên cứu hoặc kỹ thuật sinh học. Các nhà nghiên cứu đang tìm kiếm một dự đoán giờ đây sẽ có thể hoạt động hiệu quả hơn.

Giống như công ty phát hiện gian lận tài chính Verafin đã thảo luận trước đó trong cuốn sách này, một dự đoán tốt hơn nhiều là một giải pháp điểm có thể đưa vào một hệ thống hiện có và làm cho nó hoạt động tốt hơn một chút.

Sự cường điệu xung quanh AlphaFold được thúc đẩy bởi tầm nhìn về một hệ thống nghiên cứu y học mới. Hệ thống đó sẽ “đòi hỏi phải suy nghĩ nhiều hơn và ít sử dụng pipet hơn”. Một khi có thể dễ dàng dự đoán cấu trúc của bất kỳ loại protein nào, thì cách tiếp cận nghiên cứu sẽ thay đổi. Những người làm việc để tận dụng cơ hội mà AlphaFold mang lại nhận ra rằng có nhiều cơ hội hơn để phát triển các phương pháp điều trị cho

các bệnh tương đối khó hiểu vì việc khám phá cấu trúc protein giờ đây rất đơn giản. Các phòng thí nghiệm chuyên xác định cấu trúc protein không còn phục vụ nhiều mục đích nữa. Tương lai đòi hỏi nhiều phòng thí nghiệm chuyển đổi các cấu trúc protein đã biết thành các phương pháp điều trị hữu ích.

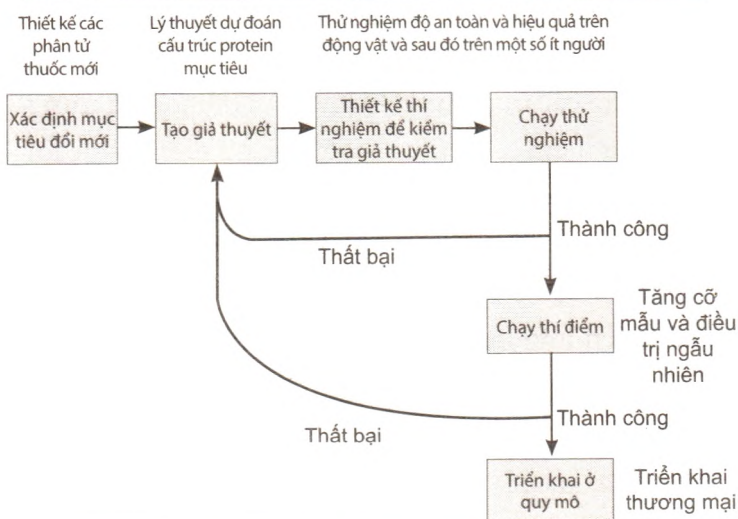
Dự đoán AlphaFold cung cấp những thay đổi về cách thức nghiên cứu có thể được thực hiện. Bằng cách thay đổi quá trình đổi mới, AlphaFold có thể thay đổi y học. Nhìn chung, tác động của AI đối với sự đổi mới cuối cùng có thể lớn hơn tác động của tất cả các ứng dụng AI khác. Vì đổi mới là trung tâm của năng suất, tăng trưởng kinh tế và phúc lợi của con người, nên thông qua tác động của nó đối với đổi mới, AI có thể có tác động lớn hơn so với các thể hệ công nghệ đa năng trước đây, từ động cơ hơi nước đến Internet.

Với AlphaFold, việc dự đoán cấu trúc protein mục tiêu không còn là một quá trình lặp đi lặp lại phức tạp giữa lý thuyết và thực nghiệm. Đó là giai đoạn bây giờ được đưa ra. Cùng với đó, các mục tiêu đổi mới có thể tham vọng hơn. Nhiều phản ứng protein thuốc có thể được kiểm tra. AI có thể ảnh hưởng đến năng suất của quy trình khám phá bằng cách cho phép cải thiện mức độ ưu tiên của



các đổi mới chảy qua quy trình đó. AI có thể tăng giá trị kỳ vọng của đổi mới và nó có thể tăng hoặc giảm thử nghiệm xuôi dòng, tùy thuộc vào đổi mới. AI có thể giảm chi phí liên quan đến các tắc nghẽn được xác định rõ trong quy trình khám phá (xem Hình 9-3). Mặc dù người ta thừa nhận rộng rãi rằng một hệ thống mới là cần thiết, nhưng một hệ thống như vậy không xuất hiện nếu không có nhiều thời gian, nỗ lực và nguồn lực.

Hình 9-3: Phức tạp: R&D cho phát triển thuốc



CÁC HỆ THỐNG NHẬN DIỆN

Quan điểm của chúng tôi là ngay cả việc nhận ra cơ hội thay đổi hệ thống cũng đã rất khó. Chúng ta đang bắt đầu nhận thấy sự tiến bộ về tác động của AI đối với hệ thống đổi mới. Một số hệ thống AI đang xuất hiện trong quá trình đổi mới. Ví dụ, giáo sư Alán Aspuru-Guzik của Đại học Toronto đang sử dụng AI cho hóa học. Trí tuệ nhân tạo AI dự đoán những giả thuyết cần thử nghiệm được tích hợp như một phần của giải pháp hệ thống liên quan đến cánh tay robot do trí tuệ nhân tạo điều khiển và phòng thí nghiệm di động được trang bị đầy đủ để chạy thử nghiệm tự động. Ông gọi hệ thống này là “phòng thí nghiệm hóa học tự lái”.

Tuy nhiên, nhiều bộ phận khác của nền kinh tế chưa nhận thấy sự cần thiết phải thay đổi. Và nhận ra nhu cầu thay đổi hệ thống chỉ là bước đầu tiên. Loại thay đổi phù hợp cần phải xảy ra, và điều đó cần rất nhiều sự đầu tư và một chút may mắn. Các công ty lớn hiếm khi thấy đáng để thay đổi cách thức hoạt động trong ngành của họ, đặc biệt nếu ngành của họ hiện đang có lợi nhuận. Rủi ro cho một sai sót là quá lớn.

Đây là lý do tại sao thay đổi công nghệ có thể dẫn đến sự gián đoạn. Công nghệ mở ra những cơ hội mới để xây dựng doanh nghiệp và phục vụ



khách hàng, nhưng không rõ chính xác bằng cách nào. Khi các công ty khởi nghiệp và các công ty nhỏ hơn có động cơ để đổi mới và các công ty lớn hơn thì không, thì sự đổi mới sẽ ươm mầm trong các thị trường nhỏ do các công ty nhỏ phục vụ cho đến khi sản phẩm của họ trưởng thành thành các lựa chọn thay thế khả thi cho các thị trường lớn, và cuối cùng, các công ty đương nhiệm sụp đổ và những cách thức kinh doanh mới nảy sinh từ những nơi đáng ngạc nhiên này. Chúng ta chuyển sang kiểu thay đổi này ở phần tiếp theo.

NHỮNG ĐIỂM CHÍNH

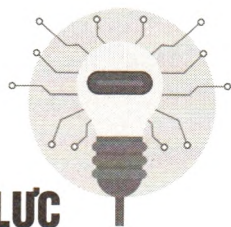
- Bản thân những đổi mới trong hệ thống đổi mới có thể có tác động theo tầng xuống nhiều hệ thống khác. Những tiến bộ trong công nghệ mài thấu kính đã dẫn đến những đổi mới trong thị trường quang học cá nhân (ví dụ: kính đeo mắt) mà còn trong thị trường công cụ nghiên cứu (ví dụ: kính hiển vi), cho phép đổi mới hơn nữa trong hệ thống đổi mới. Từ kính hiển vi, lý thuyết mầm bệnh đã ra đời, khiến cho việc chiến đấu với virus và vi khuẩn trở nên khả thi và thay đổi các khía cạnh khác của y học.
- Một trong những vai trò cốt lõi của AI trong hệ thống đổi mới là dự đoán hệ quả của các kết hợp mới. Trước đây chúng ta dựa vào lý thuyết khoa học hoặc phép thử và sai, giờ đây đôi khi chúng ta có thể (nếu có đủ dữ liệu để huấn luyện các mô hình) sử dụng dự đoán AI để tạo ra các giả thuyết.
- Việc tạo giả thuyết tự động có thể nâng cao năng suất đổi mới một cách đáng kể. Tuy nhiên, để hưởng lợi hoàn toàn từ công nghệ này, chúng ta phải xem xét lại toàn bộ hệ thống đổi mới, không chỉ một bước tạo ra giả thuyết. Ví dụ: việc tạo giả



thuyết nhanh hơn sẽ có ít tác động nếu bước tiếp theo trong quy trình, kiểm tra giả thuyết, không thay đổi và chỉ đơn giản là tạo ra một nút cổ chai ở hạ lưu.

Chương 10

SỰ GIÁN ĐOẠN VÀ QUYỀN LỰC



Hãy dành một chút thời gian để tóm tắt lại những phần trước. AI là một công nghệ có khả năng biến đổi. Lịch sử cho chúng ta biết rằng sự biến đổi không đến dễ dàng. Mặc dù thành quả ban đầu của một công nghệ mới là các giải pháp và ứng dụng điểm, nhưng việc áp dụng và chuyển đổi thực sự xảy ra khi các hệ thống mới được điều khiển bởi công nghệ này được tạo ra. AI cung cấp các dự đoán và do đó, giá trị được tạo ra bằng cách cải thiện quá trình ra quyết định. Các giải pháp điểm cho phép dự đoán để cải thiện các quyết định hiện có. Các ứng dụng có khả năng mở khóa các quyết định mới. Những quyết định mới không đến từ màu xanh. Thay vào đó, họ thay thế các quy tắc. Các quy tắc thừa nhận lỗi và thường có giàn giáo được xây dựng xung quanh chúng (ví dụ: nhà ga sân bay) để giảm thiểu hậu quả của chúng. Do đó, sự không chắc chắn có thể là mục tiêu cho dự đoán AI có thể bị chôn vùi. Kết quả là hệ thống được xây dựng dựa trên các quy tắc có thể rất mạnh mẽ – cái

mà chúng tôi gọi là “đính chặt”. Điều đó có nghĩa là thay đổi quy tắc thành quyết định và việc sử dụng AI có thể không đáng giá trừ khi chúng ta áp dụng tư duy hệ thống.

Có nhiều việc phải làm để đổi mới ở cấp độ hệ thống nếu muốn AI đạt được tiềm năng biến đổi của nó. Phần còn lại của cuốn sách này sẽ chỉ ra những cách làm hiệu quả và những thách thức mà nhiều tổ chức sẽ phải đối mặt. Điều quan trọng nhất cần phải hiểu rõ là quá trình này sẽ bị gián đoạn.

Gián đoạn. Đây chính là vấn đề. Là những nhà kinh tế học, chúng tôi tránh sử dụng các biệt ngữ và có lẽ không có từ nào “biệt lập” hơn trong giới công nghệ hơn là từ *gián đoạn*. Tuy nhiên, khi xem xét việc áp dụng AI trong bối cảnh thay đổi toàn hệ thống, nó phù hợp vì 3 lý do. Đầu tiên, như chúng ta đã thấy, các cơ hội ứng dụng AI có thể bị che khuất và do đó các ngành hiện tại dễ bị điếm mù. Thứ hai, những thách thức và sự đánh đổi trong việc gỡ bỏ các hệ thống hiện có và xây dựng những hệ thống mới là một phần của quá trình phá hủy sáng tạo đi kèm với sự thay đổi công nghệ mang tính chuyển đổi. Cuối cùng, khi các hệ thống cũ được thay thế bằng hệ thống mới, nhất thiết phải có sự thay đổi về quyền lực – cụ thể là quyền lực kinh tế – khiến cho



việc tích lũy quyền lực trở thành phần thưởng cho sự đổi mới hệ thống và sự gián đoạn tiềm ẩn là điều khiến người ta sợ hãi và phản đối. 3 điều này có liên quan đến điều mà nhiều người gọi là sự gián đoạn, vì vậy sẽ rất hợp lý khi sử dụng nó ở đây.

DỰ BÁO CÓ THỂ GÂY GIÁN ĐOẠN

Phần còn lại của cuốn sách này sẽ tập trung làm rõ dự báo bằng AI và việc áp dụng những dự báo đó có thể gây gián đoạn như thế nào. Tất nhiên, vì điều này chưa xảy ra, chúng tôi đã xây dựng suy đoán về những gì sẽ xảy ra. Tuy nhiên, suy đoán của chúng tôi dựa trên bởi hiểu biết về dự đoán bằng AI và bản chất kinh tế học của công nghệ. Rõ ràng chúng ta đã đạt được những bước tiến nhất định trong khả năng dự đoán với hỗ trợ từ thu thập dữ liệu và vi tính hóa.

Cho đến những năm 1990, hoạt động sản xuất điện vẫn được kiểm soát chặt chẽ. Ở hầu hết các quốc gia, toàn bộ quá trình sản xuất và cung cấp điện cho các doanh nghiệp và hộ gia đình thuộc quyền quản lý của các tổ chức được cơ cấu theo chiều dọc. Một lý do cho mô hình này là bởi chi phí cho sản xuất là rất lớn và hầu như không thể lưu trữ điện. Bạn chỉ có thể sản xuất lượng điện mà bạn xác định sẽ cần đến. Thế nhưng, vẫn có hàng triệu

quyết định về mục đích sử dụng điện được đưa ra mỗi phút trong ngày. Không đáp ứng được nhu cầu và toàn bộ hệ thống có thể bị sập.

Trên thực tế, bạn có thể sản xuất bao nhiêu điện tùy thích, nhưng bạn cũng phải quan tâm đến nhiều thứ khác. Bạn phải đảm bảo rằng các đường dây phân phối không đạt công suất ở các địa phương khác nhau. Bạn phải tính đến các nhà máy phát điện không sẵn sàng hoạt động. Và bạn phải tính đến biến động chi phí nhiên liệu được sử dụng ngay từ đầu. Kết quả cuối cùng là kiểm soát chặt chẽ dư thừa nguồn cung tại bất kỳ thời điểm nào. Sự thận trọng là nguyên tắc kiểm soát hàng đầu.

Nhưng theo thời gian, dự báo – dự báo thời tiết, dự báo kỹ thuật và dự báo khủng hoảng – trở nên tốt hơn. Xu hướng vi tính hóa giúp ích rất nhiều, cũng như tích lũy kinh nghiệm. Xét về mặt vận hành, hiệu quả dự báo tốt hơn cho phép các công ty điện lực bớt thận trọng hơn một chút. Tuy nhiên, những hệ thống được lên kế hoạch này đã không làm tốt công việc tiết kiệm chi phí nhiên liệu. Việc vận hành một nhà máy điện than hoặc điện hạt nhân lớn trở nên quá dễ dàng, so với việc khởi động và tạm dừng một nhà máy khí công suất thấp hơn.

Những tiến bộ trong dự báo này cuối cùng cũng tạo ra sự thay đổi trong cơ cấu tổ chức của



ngành: hướng tới tính mô-đun hóa lớn hơn và sự phối hợp ít tập trung hơn. Thay vì các máy phát điện được kiểm soát và lên kế hoạch chặt chẽ, dự báo nhu cầu và phân phối tốt hơn có nghĩa là các nhà máy riêng lẻ có thể đấu thầu giành quyền phân phối điện vào thị trường địa phương. Đến cuối những năm 1980, các nhà kinh tế học và chuyên gia về điện nhận ra rằng những thị trường điện như vậy có thể giảm chi phí mà không ảnh hưởng đến chất lượng.

Vấn đề cốt lõi đảm bảo cung vượt cầu không mất đi. Thay vào đó, dự đoán chất lượng cao về nhu cầu điện địa phương có nghĩa là các quyết định sản xuất của từng nhà máy có thể được phân cấp. Giá thầu có thể được nhận trước một ngày và số lượng cung đáp ứng số lượng cầu với ít rủi ro giá cao bất thường. Chắc chắn, một số đợt tăng giá đã xảy ra và chủ sở hữu của các nhà máy phát điện chắc chắn thích điều đó. Trên toàn hệ thống, dự đoán tốt hơn có nghĩa là ít lập kế hoạch tập trung hơn và cạnh tranh nhiều hơn. Tổng chi phí điện giảm mạnh. Không có gì đáng ngạc nhiên, hầu hết tất cả các nền kinh tế hàng đầu hiện nay đều có thị trường tập hợp điện không thể tưởng tượng được trước đây, tất cả đều được kích hoạt nhờ khả năng dự đoán tốt hơn.

Dự báo đã mở ra một phương thức tổ chức mới cho ngành điện, với việc phát điện phi tập trung. Khi dự báo tốt hơn, những sự kiện ngoài dự kiến như nhu cầu thực tế vượt xa kế hoạch sẽ trở nên hiếm hoi hơn. Do đó, thông tin từ dự báo sẽ được gửi trước đến các nhà máy phát điện, nhà điều hành đường dây truyền tải và nhà phân phối trong thời gian nhiều hơn một ngày và sau đó mỗi nhà máy có thể gửi thông báo về kế hoạch cung cấp của họ đến nhà điều hành hệ thống tập trung. Điều này có nghĩa là họ có nhiều quyền quyết định hơn, từ đó mở ra cơ hội đầu tư vào các nhà máy khác nhau và các phương án thay thế khác với hiệu suất tương đương. Không nhất thiết phải có một kế hoạch trung tâm cho mọi thứ.

Quá trình chuyển đổi của ngành điện là sự chuyển đổi từ tập trung sang phi tập trung. Nhưng điều quan trọng cần ghi nhớ là ý nghĩa thực sự của sự chuyển đổi đó, cụ thể ai là người nắm giữ quyền lực, không phải về điện năng mà về kinh tế.

Bạn có sức mạnh kinh tế nếu những gì bạn sở hữu hoặc kiểm soát vừa có giá trị vừa khan hiếm trên thị trường. Đây là lý do tại sao sở hữu bức chân dung người vợ của Picasso khác với việc sở hữu bức vẽ người bạn đời của bạn mà đứa con năm tuổi của



bạn là tác giả. Cả hai đều là duy nhất và do đó khan hiếm, và cả hai đều thể hiện một cách không chính xác diện mạo của một người, nhưng một cái có giá trị hơn. Đây cũng là lý do tại sao việc sở hữu một bức tranh gốc của Picasso khác với việc sở hữu một bản in của cùng một bức tranh. Cả hai đều tương đương về mặt thẩm mỹ, nhưng một cái khan hiếm trong khi cái kia thì không.

Sự khan hiếm, làm nền tảng cho quyền lực kinh tế, là thứ có thể được cải thiện nhờ cạnh tranh, vì vậy các nhà kinh tế học đôi khi coi sức mạnh kinh tế và sức mạnh độc quyền là tương đương nhau. Khi một thứ gì đó khan hiếm trước đây đối mặt với cạnh tranh, quyền lực sẽ thay đổi.

Những thay đổi về điện đã gây khó khăn lớn cho các nhà sản xuất vốn dựa vào cấu trúc quản lý theo chiều dọc của ngành để bảo vệ họ khỏi đe dọa cạnh tranh. Trên thực tế, đối với các nhà máy điện nói riêng, cạnh tranh có nghĩa là lợi nhuận thấp hơn, vì giờ đây họ phải đấu thầu để được phân phối thay vì dựa vào các hợp đồng dài hạn và các thỏa thuận khác. Tương tự như vậy, nhiều cơ hội mở ra cho nhiều nhà phân phối có thể kết nối các hệ thống địa phương với các thị trường khu vực rộng lớn hơn. Ở lục địa Hoa Kỳ, chỉ có 10 thị trường điện phân phối cho nhiều tiểu bang.

Chúng ta thường nhìn nhận điều này với sự gián đoạn. Một ngành mà các nhà cung cấp truyền thống có sức mạnh kinh tế đột nhiên trở thành đối tượng cạnh tranh và sức mạnh của họ bị suy giảm. Nhưng quyền lực không tự nhiên biến mất; nó thay đổi. Trong trường hợp điện, năng lượng đã chuyển từ các nhà cung cấp tích hợp theo chiều dọc sang một số nhà cung cấp khác, nhưng đáng kể nhất là người tiêu dùng điện.

Trong những trường hợp khác, sự gián đoạn dưới hình thức cạnh tranh có thể chuyển quyền lực từ nhà sản xuất truyền thống sang nhà sản xuất mới. Nói cách khác, sức mạnh độc quyền vẫn tiếp tục; chỉ tên của các nhà độc quyền thay đổi. Do đó, sức mạnh kinh tế không bị đe dọa bởi sự đổi mới; nó là giải thưởng từ sự đổi mới đó. Như chúng ta sẽ thấy, khi sự đột phá trở thành hình thức đổi mới hệ thống, thì việc chuyển giao ngọn đuốc giữa các chủ thể quyền lực sẽ xuất hiện. Tuy nhiên, những người không có cổ phần trong hệ thống hiện tại thường sẵn sàng tốt nhất để gạt hái phần thưởng từ việc tạo ra một hệ thống mới.

CÁC MỐI ĐE DỌA GIÁN ĐOẠN

Tại sao sự gián đoạn – đặc biệt là do đổi mới hệ thống – lại gây ra mối đe dọa như vậy đối với các



công ty đương nhiệm trong việc loại bỏ sức mạnh kinh tế của họ và ngược lại, mở ra một cơ hội như vậy cho những doanh nghiệp mới gia nhập ngành?

Thuật ngữ *gián đoạn* xuất hiện trong nghiên cứu của Clayton Christensen. Christensen lưu ý rằng các công ty đương nhiệm có thể thấy mình “đặt những câu hỏi sai” về công nghệ mới và giá trị của chúng đối với khách hàng. Vì vậy, họ né tránh một số công nghệ mang lại ít lợi thế cho khách hàng của chính họ. Ngược lại, chính những công nghệ đó lại thu hút những khách hàng không được các công ty dẫn đầu thị trường hiện tại phục vụ hoặc phục vụ không đầy đủ. Ví dụ, các nhà sản xuất ổ đĩa cứng hiện nay nhấn mạnh đến hiệu suất và khả năng lưu trữ, nhưng có những khách hàng sẵn sàng đánh đổi những điều này để lấy kích thước nhỏ hơn hoặc hiệu suất năng lượng. Những người tham gia có thể nắm bắt những cơ hội đó và nếu những công nghệ đó được cải thiện, cuối cùng họ sẽ trở thành những đối thủ cạnh tranh mạnh mẽ trong ngành.

Sự gián đoạn thực sự đầy thách thức phát sinh khi thay đổi công nghệ triệt để không cải thiện hiệu suất theo các chỉ số truyền thống, nhưng trong một số trường hợp, có thể cải thiện hiệu suất trên các chỉ số không phải là trọng tâm của ngành hiện tại.

Điều này có thể tạo ra điểm mù cho những người đương nhiệm. Như nhà sử học Jill Lepore mô tả lý thuyết:

Trong cuốn sách xuất bản năm 1997, *The Innovator's Dilemma* (tạm dịch: Tình huống khó xử của người thực hiện đổi mới), Christensen lập luận rằng, rất thường xuyên, không phải vì ban giám đốc đưa ra quyết định tồi tệ mà bởi nhờ họ đưa ra những quyết định đúng đắn, chính những quyết định đúng đắn đó đã giúp các công ty thành công trong nhiều năm. (“Thế tiến thoái lưỡng nan của người thực hiện đổi mới” là “làm điều đúng lại là điều sai”.) Như Christensen đã thấy, vấn đề là sự chuyển động của lịch sử, và đó không phải là vấn đề lớn như khi bỏ lỡ cơ hội, như khi nhờ máy bay, trừ phi bạn thậm chí không biết có chuyến bay, và đi lang thang trên sân bay, nơi mà bạn nghĩ là một cánh đồng, và máy bay chạy qua bạn khi cất cánh.

Chúng ta đã thấy rằng các thành phần tạo nên những điểm mù như vậy đối với AI nằm ở các quy tắc và sự không chắc chắn tiềm ẩn. Khi điều này xảy ra, quản lý sự gián đoạn không chỉ đơn giản là nhắm mục tiêu vào một nhóm khách hàng khác mà



thay vào đó là cải tổ tổ chức của các doanh nghiệp và các vấn đề cần được ưu tiên. Không có gì đáng ngạc nhiên, một con đường nhanh chóng dẫn đến sự gián đoạn là bỏ qua việc công nghệ đòi hỏi phải thay đổi tổ chức.

Một trường hợp điển hình là xe tăng – được giới thiệu bởi người Anh vào cuối Thế chiến thứ nhất. Xe tăng di chuyển khá nhanh và có thể gây hỗn loạn cho quân đội của kẻ thù. Ít nhất đó là ý tưởng đến từ chỉ huy xe tăng của Anh vào thời điểm đó, JFC Fuller. Tuy nhiên, trong thời kỳ giữa hai cuộc chiến, nước Anh phớt lờ kế hoạch của Fuller. Thay vào đó, xe tăng được đưa vào kỵ binh. Khi nước Đức tái quân sự hóa, “người đứng đầu trong quân đội Anh, Thống chế Archibald Montgomery Massingberd, đã đã có động thái đáp trả... bằng cách tăng số tiền chi cho thức ăn thô xanh cho ngựa lên gấp 10 lần. Các sĩ quan kỵ binh sẽ được cấp thêm một con ngựa; sĩ quan xe tăng cũng sẽ nhận được một con ngựa. Ngược lại, người Đức, trước tình trạng tổ chức quân đội đã bị suy giảm, không cố gắng đưa công nghệ mới vào cơ cấu hiện tại. Họ hiểu rằng công nghệ mới đồng nghĩa với tổ chức và chiến thuật quân đội mới. Họ gọi nó là “blitzkrieg”, và họ mời Fuller đến buổi ra mắt.

Những trò tai quái của quân đội Anh khiến người ta dễ dàng loại bỏ những câu chuyện như vậy

bởi xem đó là minh chứng cho sự ngạo mạn hoặc ngu xuẩn. Thế nhưng, những câu chuyện tương tự vẫn xuất hiện và thậm chí lặp đi lặp lại trong lịch sử kinh doanh. Điều này đã không được chú ý trong giới học thuật. Khi Clay Christensen đang phát triển các lý thuyết gián đoạn tại Harvard vào năm 1990, đồng thời Rebecca Henderson và Kim Clark cũng đang nghiên cứu cùng một hiện tượng. Thay vì tập trung vào phần cầu (hay thiếu các giá trị của khách hàng) như Christensen đã làm, Henderson và Clark xem xét phần cung (hay sự thiếu phù hợp của tổ chức). Họ xác định nhiều tình huống khác nhau, trong đó sự thay đổi công nghệ mang tính cấu trúc, làm thay đổi các ưu tiên của các tổ chức và vì các tổ chức khó thay đổi, điều đó tạo cơ hội cho các tổ chức mới có thể bắt đầu từ đầu.

Một ví dụ gần đây hơn là iPhone. Năm 2007, ngành công nghiệp điện thoại di động nằm trong sự thống trị của một doanh nghiệp Canada, Research In Motion (hay RIM hoặc BlackBerry), là công ty đã phát triển thiết bị liên lạc BlackBerry. Đây là một chiếc điện thoại di động, nhưng quan trọng hơn, với thiết kế kèm bàn phím, nó hoạt động như máy gửi email và nhắn tin. Các doanh nhân gần như dán mắt vào nó. Một ngoại trưởng Hoa Kỳ yêu thích nó đến nỗi bà đã thiết lập máy chủ riêng của mình để liên tục sử dụng nó khi ở văn phòng. Lý do thiết



bị này thành công đến vậy là bởi bàn phím được thiết kế hoàn hảo, mạng gửi tin nhắn hiệu quả và an toàn, đồng thời thiết bị được chế tạo đủ cứng để chịu được va đập.

Ngược lại, iPhone rất mong manh. Nó không có bàn phím mà người dùng BlackBerry yêu thích. Kết nối mạng của nó chậm hơn. Nó hao pin. Và nó là một chiếc điện thoại khủng khiếp. Không có gì ngạc nhiên khi toàn bộ ngành công nghiệp bao gồm RIM, Nokia và Microsoft đã bác bỏ nó. Họ bảo Apple hãy ra khỏi ngành và để nó cho các chuyên gia.

Có thể dễ dàng nhận thấy sự kiêu ngạo ở đây. Tuy nhiên, những người đương nhiệm cũng không hề sai trong tất cả các lời phê bình của họ. Điều họ không hiểu là Apple đã chọn một thiết kế mới cho điện thoại thông minh. Nó tích hợp phần cứng và phần mềm. Để có một thiết bị được kết hợp khác biệt, nó phải hy sinh hiệu suất trên tất cả các bộ phận cấu thành. Khi tách riêng và xem xét các thành phần, và tất cả trông thật tồi tệ. Nhưng nếu bạn nhìn và hiểu cả hệ thống, bức tranh sẽ khác. Không có gì ngạc nhiên khi tay chơi công nghệ nhìn thấy tiềm năng của iPhone lại là một tay chơi không có cơ cấu phát triển truyền thống trong ngành: Google.

Đây chính là thách thức của việc giải quyết sự thay đổi mang tính cấu trúc, hay như chúng tôi gọi nó ở đây là sự thay đổi toàn hệ thống. Đầu tiên, để thực hiện sự thay đổi đó, bạn cần những sản phẩm mà ban đầu không có tính cạnh tranh vì kèm theo những sản phẩm này là những lựa chọn hy sinh hiệu suất cho điều khách hàng có vẻ quan tâm. Thứ hai, các tổ chức được thành lập để tập trung vào hiệu suất đó không được trang bị để nhanh chóng hiểu được tất cả những đánh đổi mà công nghệ mới đang tạo ra. Nói cách khác, họ bỏ lỡ cả khu rừng vì những chiếc cây. Cuối cùng, họ không có phản hồi nhanh chóng liên quan đến sai lầm đó. iPhone mất 4 năm để tạo ra một vết lõm trong doanh số bán hàng của nhà sản xuất thiết bị cầm tay hiện tại. BlackBerry ghi nhận doanh số bán hàng lớn nhất sau năm 2007. Chỉ sau khi cả Apple và Google đều có thiết bị thì thiết kế điện thoại mới mới được ưa chuộng hơn. Đến thời điểm đó, đã quá muộn để các công ty đối thủ thực hiện cải cách thiết kế và bắt kịp, mặc dù họ đã cố gắng.

TÌNH THẾ KHÓ CỦA SỰ THAY ĐỔI HỆ THỐNG

Khi AI thúc đẩy các quyết định trở thành một phần của hệ thống, việc áp dụng AI có thể yêu cầu thiết kế lại tổ chức với một hệ thống mới. Như chúng ta vừa thảo luận, khó khăn mà các tổ chức



hiện tại gặp phải khi phát triển các hệ thống mới là chúng đã được tối ưu hóa để đạt được hiệu suất cao từ các công nghệ hiện có, trong khi việc áp dụng AI có thể đòi hỏi phải thay đổi trọng tâm. Trong một số trường hợp, AI thúc đẩy tổ chức trở nên mô-đun hơn, trong khi ở những trường hợp khác, nó có thể thúc đẩy tổ chức có sự phối hợp tốt hơn giữa các bộ phận. Thách thức là nhận ra rằng trọng tâm hiện tại là vấn đề và cần có sự thay đổi trên diện rộng.

Khi ban lãnh đạo cấp cao hiểu rằng cần có một thiết kế tổ chức mới để áp dụng và tích hợp dự đoán AI vào một hoặc nhiều lĩnh vực quyết định quan trọng, thì một thách thức nữa sẽ nảy sinh. Điều này là do thiết kế tổ chức luôn liên quan đến sự thay đổi về giá trị và do đó, quyền lực của các nhà cung cấp các nguồn lực khác nhau trong tổ chức. Những người cho rằng mình sẽ thua trong quá trình phân bổ lại quyền lực sẽ chống lại sự thay đổi. Các tổ chức hiếm khi hoạt động như một chế độ độc tài trong sách giáo khoa nơi CEO nói gì và thay đổi cứ thế xảy ra. Thay vào đó, những người mong đợi sức mạnh của họ bị suy giảm sẽ chống lại sự thay đổi. Trong quá trình này, họ có thể thực hiện các hành động nhằm giảm bớt sự dễ dàng trong việc thực hiện thay đổi. Tồi tệ nhất, việc lường trước những hành động đó có thể khiến việc thiết kế lại tổ chức bị cắt giảm hoàn toàn hoặc đảo ngược.

Có nhiều ví dụ về khả năng chống lại sự thay đổi để đáp ứng việc áp dụng các công nghệ đột phá. Ví dụ, hãy xem xét trải nghiệm của Blockbuster Video. Blockbuster là công ty dẫn đầu thị trường cho thuê băng video trong suốt những năm 1990 và 2000. Câu chuyện thường được biết đến về sự sụp đổ của Blockbuster là nó đã bị Netflix đánh sập và sự gia tăng của video theo yêu cầu vào cuối thập kỷ đầu tiên của thế kỷ XXI. Nhưng trên thực tế, Blockbuster không thụ động khuất phục trước những cách làm mới. Nó hiểu điều gì sắp xảy ra nhưng cuối cùng không thích nghi được.

Khi bắt đầu, Netflix hướng đến khai thác công nghệ DVD mới nhỏ gọn và mạnh mẽ hơn so với băng video trước đó mà Blockbuster cho thuê. Netflix thử nghiệm và cuối cùng thành công phát triển một doanh nghiệp từ đăng ký của khách hàng, cho phép khách hàng thuê 3 đĩa DVD cùng lúc mà không có bất kỳ giới hạn nào về thời gian họ giữ chúng. Các khách hàng của Netflix đặt mua đĩa DVD trực tuyến, sau đó chúng sẽ được gửi đến họ qua đường bưu điện. Mô hình này có hai ưu điểm. Đầu tiên, khách hàng không cần phải đến cửa hàng để lấy hoặc trả lại DVD. Thứ hai, không có khoản phí trả chậm nào trong khi khoản thu này có thể chiếm tới 40% doanh thu của một cửa hàng nhượng quyền của Blockbuster. Nhược điểm là Netflix không sở



hữu các bản phát hành mới nhất và nhu cầu xuất phát từ việc lên kế hoạch từ trước của khách hàng thay vì hành động bốc đồng.

Blockbuster nhận thấy rằng nguy cơ tranh giành khách hàng từ Netflix có và trong một số trường hợp, điều này ảnh hưởng đến doanh thu của chính họ. Vào đầu những năm 2000, Blockbuster nhận thấy một số nhược điểm của mô hình riêng mà Netflix đang khai thác và nó đã thử nghiệm với video theo yêu cầu. Blockbuster là người đầu tiên cung cấp dịch vụ video theo yêu cầu! Nhưng tốc độ băng thông khi đó không như ngày nay, vì vậy thử nghiệm đó đã không thành công. Nhưng Blockbuster đã nhận ra rằng nó cũng có thể có mô hình cho thuê DVD tương tự như của Netflix. Sự khác biệt là nó có thể cung cấp tùy chọn và trả lại DVD tại cửa hàng, không chỉ qua đường bưu điện.

Vấn đề phát sinh là loại dịch vụ đăng ký này sẽ cắt giảm 40% doanh thu phí trả chậm mà các cửa hàng của Blockbuster đang kiếm được. Hơn nữa, khách hàng đó không nhất thiết phải đến cửa hàng nữa và doanh thu từ những hàng hóa khác, chẳng hạn như bỏng ngô và kẹo sẽ bị ảnh hưởng. Vì vậy, trong khi công ty Blockbuster có thể được hưởng lợi từ việc làm theo mô hình Netflix, thì các cửa hàng nhượng quyền của nó bị thiệt hại. Vì vậy

xuất hiện sự phản đối, đặc biệt là khi mô hình mới tỏ ra thành công hơn. Cuối cùng, điều này khiến hội đồng quản trị Blockbuster quyết định thay đổi ban lãnh đạo cấp cao và trở lại mô hình ban đầu để hỗ trợ nhượng quyền. Để cạnh tranh với Netflix, nó đã cố gắng tăng cường những dịch vụ khác mà các cửa hàng nhượng quyền cung cấp ngoài cho thuê video. Cuối cùng, cách thức này vẫn không hiệu quả và trong vòng vài năm, Blockbuster ngừng hoạt động.

Tất nhiên, trường hợp của Blockbuster là một ví dụ ẩn tượng về sự thất bại trong việc thay đổi khi đối mặt với một công nghệ mới và cách mà các lực lượng bên trong đã ngăn chặn sự thay đổi đó trước khi quá muộn. Nhưng đó là một trường hợp được dự báo chính xác bởi vì xung đột giữa những người được hưởng lợi từ công nghệ mới và những người không được hưởng lợi quá gay gắt. Đơn giản là các cửa hàng bán lẻ không có vai trò gì trong thế giới mới, nhưng điều đó đủ để ngăn bản thân doanh nghiệp thích nghi, mặc dù lãnh đạo cấp cao của nó hiểu những gì tổ chức mới phải có.

Ở mức độ rộng hơn, như chúng tôi sẽ phác thảo trong các chương tiếp theo, AI có thể tạo ra sự thay đổi về tổ chức có thể có tác dụng phân cấp quyền lực hoặc sự phối hợp có thể tập trung hóa nó.



Dù bằng cách nào, những người thua cuộc từ những thay đổi đó có thể khá rõ ràng, và chính vì họ nắm giữ quyền lực dựa trên hệ thống tổ chức hiện tại nên họ sẽ có quyền lợi nhất định trong việc duy trì nó.

SỰ GIÁN ĐOẠN VÀ CƠ HỘI

Thay đổi cấp độ hệ thống là khó khăn, nhưng phần thưởng để thành công có thể rất lớn. Một câu hỏi liên tục xuất hiện liên quan đến AI là liệu các cỗ máy thể hiện khả năng dự đoán của AI hay không – dù là rô-bốt vật lý hay phần mềm các thuật toán – bản thân chúng có sức mạnh. Khi bạn biết AI là gì, mối lo ngại của bạn rằng máy móc sẽ có sức mạnh đã bị đặt nhầm chỗ. Vì sự kiên trì của những lập luận đó, chúng tôi giải quyết vấn đề tiếp theo.

NHỮNG ĐIỂM CHÍNH

- Người đương nhiệm thường có thể áp dụng các giải pháp điểm khá dễ dàng vì chúng cho phép cải thiện một quyết định hoặc nhiệm vụ cụ thể mà không yêu cầu thay đổi các quyết định hoặc nhiệm vụ liên quan khác. Tuy nhiên, những người đương nhiệm thường gặp khó

khẩn trong việc áp dụng các giải pháp cấp hệ thống vì những giải pháp đó yêu cầu thay đổi đối với các nhiệm vụ liên quan khác và tổ chức đã đầu tư vào việc tối ưu hóa các nhiệm vụ khác đó; hơn nữa, giải pháp hệ thống có thể kém hơn trong một số nhiệm vụ đó, đặc biệt là trong thời gian ngắn. Điều đó tạo tiền đề cho sự gián đoạn.

- Chúng tôi định nghĩa sức mạnh là sức mạnh kinh tế. Bạn có quyền lực nếu những gì bạn sở hữu hoặc kiểm soát khan hiếm so với nhu cầu. Sự khan hiếm, làm nền tảng cho sức mạnh kinh tế, là thứ có thể được cải thiện nhờ cạnh tranh, đó là lý do tại sao các nhà kinh tế đôi khi coi sức mạnh kinh tế và sức mạnh độc quyền là tương đương nhau. Khi một thứ gì đó khan hiếm trước đây bị cạnh tranh, quyền lực sẽ thay đổi.
- Đôi khi, cần phải có giải pháp cấp hệ thống để hưởng lợi đầy đủ từ AI. Việc thiết kế lại một hệ thống có thể dẫn đến sự thay đổi quyền lực ở cấp độ ngành (ví dụ: các ngành giàu dữ liệu trở nên mạnh mẽ hơn khi AI trở nên phổ biến hơn), cấp độ công ty (ví dụ: được thảo luận trong Chương 12), hoặc cấp độ công việc (ví dụ: nhượng quyền phim bom tấn bị mất

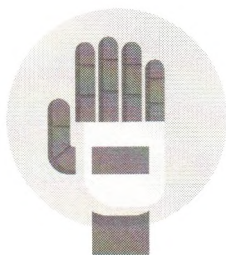


quyền lực trong quá trình chuyển sang cho thuê phim trực tuyến và gửi DVD qua thư). Những người đứng trước nguy cơ mất quyền lực sẽ chống lại sự thay đổi. Những người chống lại sự thay đổi hiện đang nắm giữ quyền lực (đó là lý do tại sao họ chống lại) và do đó có thể khá hiệu quả trong việc ngăn chặn sự thay đổi ở cấp độ hệ thống. Điều đó tạo ra bối cảnh cho sự gián đoạn.

Chương 11

MÁY MÓC

CÓ QUYỀN LỰC KHÔNG?



Các tiêu đề với nội dung như “Làm thế nào Amazon tự động theo dõi và sa thải công nhân kho theo “hiệu suất”, “Amazon đã sử dụng AI để tự động sa thải những công nhân năng suất thấp”, “Bạn có để robot sa thải nhân viên của mình không?”, “Bị Bot sa thải tại Amazon: ‘Bạn đang phải chống lại máy móc’” và “Đối với những người lao động được trả lương thấp hơn, các lãnh chúa robot đã đến”. Tiêu đề cuối cùng là của một bài báo trên *Wall Street Journal* vào tháng 5 năm 2019. Trong đó, Greg Ip đã tóm lược điểm chính sau: Đã đến lúc ngừng lo lắng robot sẽ thay thế công việc của chúng ta – và bắt đầu lo lắng rằng chúng sẽ quyết định ai sẽ có việc làm.

Rõ ràng điều đó đã thu hút sự chú ý của nhiều người. Nhận định đó chạm vào một nỗi sợ hãi nguyên thủy: Liệu máy móc có nắm quyền đối với con người không?



Những bài báo này sẽ khiến bạn tin rằng người lao động đang bước vào một căn phòng nhỏ khi họ kết thúc một ngày làm việc, được quét để đánh giá và sau đó có thể nhận được tấm biển màu xanh lá cây tươi sáng có nội dung “Hẹn gặp lại vào ngày mai” hoặc, thay vào đó, một tấm biển màu đỏ có nội dung “Bị sa thải” cùng với một tờ phiếu màu hồng được in tự động.

Cuối cùng, tình hình thực tế không đến mức cường điệu như thế. Đúng là Amazon đã sử dụng AI để dự đoán hiệu suất của công nhân. Và, vâng, điều đó có thể thúc đẩy việc đánh giá lại nhân viên. Và, vâng, sau khi đánh giá, một công nhân có thể bị chấm dứt hợp đồng. Thế nhưng, điều không đúng là nhân viên sẽ không bị sa thải mà không có sự can thiệp của con người. Thay vào đó, những gì Amazon đang làm là đo lường hiệu suất của công nhân, sử dụng AI để đánh giá xem hiệu suất đó có gây nên sự lo ngại nào hay không và sau đó để người quản lý quyết định phải làm gì. Nếu người quản lý đó chỉ mù quáng làm theo dự đoán của AI, thì có vẻ như AI đang kiểm soát quyết định. Như vậy, nó không khác với bất kỳ cơ chế đo lường hiệu suất, và rõ ràng là cơ chế này tệ hơn một số cơ chế chủ quan hơn mà hầu hết chúng ta phải đối phó.

Nhưng điều gì sẽ xảy ra nếu tất cả các tiêu đề đều đúng và bạn thực sự có thể được đánh giá và sau đó bị sa thải mà không có sự can thiệp của con người trong quá trình này không? Là máy móc quyết định ai làm gì? Bây giờ máy móc có phải là giai cấp tư sản đối với giai cấp vô sản của chúng ta không? Câu trả lời, như chúng tôi sẽ chứng minh trong chương này, là hoàn toàn không. Robot và máy móc nói chung không quyết định bất cứ điều gì và do đó không có quyền lực. Một người hoặc một nhóm người đang thực hiện các cuộc gọi làm cơ sở cho các quyết định. Chắc chắn là có thể tự động hóa mọi thứ và làm cho nó giống như một cái máy đang làm công việc bẩn thỉu. Nhưng đó là ảo tưởng. Ở cấp độ AI hiện tại, ai đó, tức con người sẽ đưa ra quyết định thực sự.

Chúng tôi không nói điều này để đưa ra một quan điểm triết học nào đó; chúng tôi để lại cuộc tranh luận này cho những người khác. Thay vào đó, việc chấp nhận rằng máy móc không quyết định là rất quan trọng nếu chúng ta định đánh giá đúng tiềm năng đột phá của AI. *Mặc dù AI không thể đưa ra quyết định cho máy móc, nhưng nó có thể thay đổi việc con người đưa ra quyết định.* Máy móc không có sức mạnh, nhưng khi được triển khai, chúng có thể thay đổi người làm.



Khi máy móc thay đổi người đưa ra quyết định, hệ thống cơ bản phải thay đổi. Các kỹ sư chế tạo máy cần hiểu hậu quả của phán đoán mà họ đưa vào sản phẩm của mình. Những người đã từng quyết định trong thời điểm này có thể không còn cần thiết nữa.

Quan điểm cho rằng máy móc không quyết định không còn xa lạ. Ada Lovelace, người viết chương trình máy tính đầu tiên vào năm 1842, đã nhìn thấy hạn chế này:

Ada cảnh báo người đọc về việc máy tính sẽ không thể làm gì nếu người dùng nhập thông tin “không đúng sự thật”. Ngày nay, chúng tôi gọi khái niệm này là “Rác vào, Rác ra”, tức là dữ liệu sai đưa ra quyết định sai. Chính xác thì cô ấy nói rằng: “Máy phân tích không có bất kỳ kỳ vọng nào để tạo ra bất cứ thứ gì. Nó có thể làm bất cứ điều gì khi chúng ta biết cách ra lệnh cho nó thực hiện. Nó có thể làm theo phân tích; nhưng nó không có khả năng dự báo bất kỳ mối quan hệ hay sự thật phân tích nào”.

Máy làm theo hướng dẫn và những hướng dẫn đó phải đến từ đâu đó. Hãy xem xét phiên bản giả định của thuật toán sa thải của Amazon, theo đó nó

thực sự đo lường hiệu suất và sa thải những người không có con người như một phần của quy trình. Khi lập trình thuật toán này, ai đó, tại ở một điểm nào đó, phải chỉ rõ các yếu tố đánh giá – bao gồm cách cân nhắc các yếu tố như mức lương hiện hành, sự sẵn có của nhân viên thay thế, yêu cầu đào tạo và những thứ khác như quy tắc pháp lý về thực hành tại nơi làm việc, cũng như cách cân nhắc các yếu tố xác suất chẳng hạn như dự đoán AI về các kỹ năng, khả năng và sự phù hợp về văn hóa. Một số kỹ sư ở đâu đó có thể đang cố gắng hoàn thành một chương trình, nhưng nhiều khả năng, nếu một hệ thống tự động như vậy được triển khai, những yếu tố phán đoán đó sẽ đến từ một quy trình có chủ ý hơn. Một hệ thống ra quyết định mới là cần thiết.

Thật hấp dẫn khi nghĩ rằng chúng ta có thể trao toàn bộ quyết định cho máy móc. Tuy nhiên, trong khi việc thực hiện một quyết định có thể trở nên hoàn toàn tự động, quyết định thực hiện hành động nào sau khi dự đoán vẫn phải đến từ một hoặc nhiều người.

TRỞ NÊN PHỔ BIẾN TOÀN CẦU

Máy móc dường như bắt đầu ra quyết định khi AI cho phép tự động hóa. Các máy dự báo có thể thay đổi thời gian và địa điểm của các quyết định,



cho phép con người tham gia vào quá trình thảo luận, đánh giá những việc cần làm khi một tình huống phát sinh và sau đó mã hóa điều đó vào máy.

Tự động hóa yêu cầu mã hóa phán đoán. Con người phải làm rõ phán đoán khi bắt đầu triển khai máy, thay vì chờ đến lúc nhận được dự báo. Điều đó có nghĩa là phán đoán đó phải hữu ích cho nhiều quyết định và nó phải được mô tả theo cách có thể được mã hóa. Điều đó không quá dễ dàng.

Công ty khởi nghiệp Ada ở Toronto sở hữu quy trình tự động hóa dịch vụ khách hàng rất đáng tham khảo. Những người sáng lập của Ada mô tả quy trình này là lớp tự động hóa hỗ trợ tương tác giữa các công ty và khách hàng của họ.

Ada cung cấp lớp tự động để hỗ trợ khách hàng của Zoom khi lượng người sử dụng nền tảng này tăng từ 10 triệu người mỗi ngày lên 300 triệu người trong nửa đầu năm 2020 – trước và sau khi Covid-19 xuất hiện. Nó đã tự động hóa 70% cuộc gọi bán hàng và 98% tương tác hỗ trợ khách hàng với người dùng miễn phí và 85% tương tác đó với người dùng trả phí. Có thể khi cần đặt lại mật khẩu Zoom hoặc khi máy ảnh của bạn không hoạt động, thì đối tượng tương tác của bạn chính là lớp tự động của Ada.

Trọng điểm là quá trình xây dựng phán đoán. Quy trình của Ada bắt đầu bằng việc dự đoán ý định của khách hàng khi họ bắt đầu tương tác với dịch vụ. Ý định đó có thể là thay đổi mật khẩu, cập nhật thông tin thẻ tín dụng hoặc nâng cấp lên một dịch vụ toàn diện hơn. Ada có thể bắt đầu chỉ với một dự đoán mà nó có thể thực hiện một cách tự tin: khách hàng muốn thay đổi mật khẩu.

Sau đó, Ada phát triển quy trình làm việc và phán đoán. Quy trình làm việc là tập hợp các hành động giúp khách hàng thay đổi mật khẩu. Nếu máy chắc chắn rằng khách hàng muốn đổi mật khẩu, thì Ada sẽ bắt đầu quy trình tự động thay đổi mật khẩu. Nếu không, Ada sẽ gửi nó cho nhân sự xử lý.

Đây là lúc phán đoán xuất hiện. Ada nên tự tin đến mức nào trước khi bắt đầu trình tự tự động? Điều này còn phụ thuộc. Có một lý do tại sao nó tự động hóa 98% tương tác từ khách hàng miễn phí và chỉ 85% từ khách hàng trả phí. Và không chỉ vì khách hàng miễn phí có các truy vấn dễ dàng hơn. Ngoài ra, nó đánh giá rằng mắc lỗi với khách hàng trả tiền là một vấn đề lớn hơn. Cổ phần cao hơn, vì vậy ngưỡng cho phép tự động hóa cao hơn.

Khi Ada thu thập dữ liệu về các yêu cầu gửi đến và ý định của khách hàng, nó sẽ xây dựng các



quy trình công việc tự động hơn. Ngoài mật khẩu, nó có thể thêm các bản cập nhật thẻ tín dụng và các vấn đề kỹ thuật khác nhau. Nó cũng có thể xác định những yêu cầu nào là cuộc gọi bán hàng, để mua dịch vụ trả phí hoặc để nâng cấp lên cấp cao hơn.

Lúc này phán đoán trở nên đặc biệt quan trọng. Hậu quả của việc nhầm lẫn một cuộc gọi bán hàng lớn hơn hậu quả của việc nhầm lẫn đổi mật khẩu. Để làm tốt điều này, Ada cần quyền truy cập vào dữ liệu và quy trình quyết định để xác định những yêu cầu gửi đến nào sẽ được tự động hóa. Một số quyền ra quyết định chuyển từ đại diện dịch vụ khách hàng sang quản lý công ty và các kỹ sư tại Ada. Dự đoán tốt hơn về ý định của khách hàng tạo cơ hội cho tự động hóa. Nhưng liệu cơ hội đó có đáng giá hay không tùy thuộc vào sự đánh giá của con người để cân nhắc lợi ích của tự động hóa so với chi phí sai lầm. Làm tốt tất cả những điều này đòi hỏi những thay đổi ở cấp hệ thống đối với việc thu thập dữ liệu, ra quyết định và phân bổ trách nhiệm.

BẠN CÓ CẢM THẤY MAY MẮN KHÔNG?

Một nỗi sợ hãi khác nảy sinh đối với những cỗ máy có sức mạnh là những cỗ máy dự báo hiện nay thường chịu trách nhiệm về thông tin mà chúng ta

nhìn thấy để giúp chúng ta hiểu về thế giới và đưa ra quyết định, từ mua sắm đến bỏ phiếu cho ai. Nếu máy móc cung cấp thông tin cho chúng ta, liệu sức mạnh của chúng ta có bị giảm đi một cách tinh vi không? Như chúng ta sẽ thấy ở đây, mối quan hệ của chúng ta với cỗ máy dự đoán không phải là một chiều. Vâng, nó cung cấp cho chúng ta thông tin và điều đó ảnh hưởng đến chúng ta, nhưng chúng ta cũng cung cấp cho máy thông tin được sử dụng để thay đổi dự đoán của máy. Nói cách khác, từ góc độ kinh tế, máy móc (và chủ sở hữu của chúng) không nắm giữ tất cả các thẻ. Họ cần chúng tôi để duy trì chất lượng của họ. Vì vậy, mặc dù có thể cảm thấy như bạn không có quyền kiểm soát, nhưng bạn có nhiều hơn những gì bạn nghĩ.

Hãy xem xét thứ gần nhất mà chúng ta có với một siêu AI ngày nay: Google Search, một cỗ máy dự đoán. Bạn hỏi nó một câu hỏi hoặc thậm chí chỉ một vài thuật ngữ mà bạn muốn biết thêm thông tin, nó sẽ kiểm tra các trang web và sau đó cung cấp cho bạn một danh sách (đôi khi trong hàng chục nghìn) được xếp hạng theo thứ tự mà Google coi là có khả năng nhất đối với bạn. Trước đây, những thứ hạng này phần lớn được xác định bởi PageRank, một hệ thống tính điểm do Larry Page tạo ra để giả định rằng bạn có nhiều khả năng muốn có một



trang web mà các trang web khác liên kết đến. Giờ đây, dựa trên hàng nghìn tỉ lượt tìm kiếm và nhấp chuột, xếp hạng của Google là một dự đoán dựa trên học sâu, không chỉ tính đến những gì người khác đã làm trong quá khứ mà còn tiếp tục cập nhật và cũng sử dụng những gì nó biết về bạn để đưa ra xếp hạng được cá nhân hóa chỉ dành cho bạn. Bạn may mắn như thế nào?

Hóa ra, không hẳn là may mắn như vậy. Bạn có thể không nhận thấy điều đó, bởi vì hầu hết các tìm kiếm của Google không được thực hiện từ trang chủ của Google (tại www.google.com), nhưng trang chủ đó có hai nút (xem Hình 11-1). Sau khi nhấp cụm từ tìm kiếm, bạn có thể nhấp vào “Tìm kiếm của Google”, sau đó trả về danh sách các trang web được xếp hạng quen thuộc đó, cùng với các quảng cáo tài trợ cho toàn bộ hoạt động. Nhưng bên cạnh đó là một nút khác, “Tôi cảm thấy may mắn”. Thay vào đó, hãy nhấp vào nút đó và nó sẽ đưa bạn đến ngay trang web được xếp hạng đầu tiên. Thực tế là chúng ta hiếm khi nhấp vào nút đó và điều đó cho thấy chúng tôi không cảm thấy may mắn như vậy. Các dự báo không đủ tốt.

Hình 11-1

Trang chủ Google, 2021



Nguồn: Google và biểu tượng Google là thương hiệu của Google LLC.

Đây không phải là một nút mới. Nó xuất hiện ngay trên trang chủ đầu tiên của Google (xem Hình 11-2), và mặc dù Google mong muốn bảo vệ diện tích trang chủ của mình để có giao diện tối giản, nó vẫn ở đó. “Tôi cảm thấy may mắn” ban đầu được đặt ở đó bởi người đồng sáng lập Google, ông Serge Brin, người đã coi đó là một cách để làm nổi bật Google Search khi nó được ra mắt. Năm 2007, Brin tuyên bố rằng tỷ lệ người dùng Google cảm thấy may mắn chỉ khoảng 1%, vì đó là số lượng tìm kiếm sử dụng nút “Tôi cảm thấy may mắn”. Đồng thời, các ước tính chỉ ra rằng việc duy trì thêm nút lựa chọn đó đã khiến Google mất 100 triệu đô la doanh thu quảng cáo. Google giữ nút đó vì những lý do xây dựng thương hiệu thuần

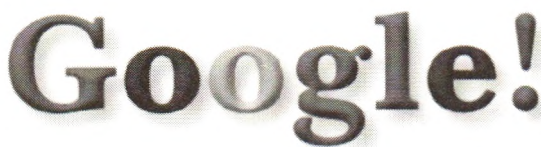


túy, giữ khuôn mặt người, nếu không sẽ là hoạt động nhân tạo hoàn toàn.

Tại sao chúng ta không cảm thấy may mắn? Câu trả lời đơn giản là bởi kết quả tìm kiếm đầu tiên thường không như chúng ta mong muốn. Chúng ta muốn duyệt trang đầu tiên đó và chọn một cái gì đó. Sau đó, chúng ta có thể nhận ra nó không đúng. Vì vậy, chúng ta quay lại và chọn một liên kết khác. Từ quan điểm của Google, nó không thể làm tốt hơn nữa. Nếu ai đó vào Google tìm kiếm “prediction machines” và cảm thấy may mắn, họ sẽ được đưa đến trang web của cuốn sách đầu tiên của chúng tôi (predictionmachines.ai). Nhưng nếu họ muốn mua nó từ Amazon thì sao? Trong trường hợp đó, đây sẽ không phải là đường dẫn hiệu quả nhất đến thông tin đó. Điều gì sẽ xảy ra nếu họ không muốn cuốn sách, mà là một bài viết tóm tắt? Bạn có thể biết, nhưng Google không biết bạn có thể đưa ra quyết định gì. Không có điều đó, nó phục vụ bạn dự đoán tốt nhất của nó, nhưng vẫn để ngỏ cho sự phán đoán của chính bạn để kết thúc quá trình. Google có thể khao khát trở thành một cỗ máy ra quyết định, nhưng nếu không có phán đoán, nó không thể là cỗ máy quyết định của bạn. Vì vậy, nó chỉ còn lại một vai trò trong dự báo và phải để phần còn lại cho bạn. Tất nhiên, điều đó cũng tốt thôi, vì bạn có thể chọn nhấp vào một quảng cáo.

Hình 11-2

Trang chủ Google, 1998



Search the web using Google!

10 results ▾

Google Search

I'm feeling lucky

Index contains ~25 million pages (soon to be much bigger)

Nguồn: Google và biểu tượng Google là thương hiệu của Google LLC.

Ví dụ về công cụ tìm kiếm của Google – Google Search cho thấy việc tự động hóa một quyết định khó đến mức nào. Khó, nhưng không phải là không thể. Ở thời điểm viết cuốn sách này, trên thực tế, Google đã cảm thấy may mắn hơn. Với sự ra đời của tìm kiếm hỗ trợ bằng giọng nói, người dùng đưa ra các truy vấn đầy đủ hơn và nhờ vậy Google có thể đưa ra câu trả lời rõ ràng hơn và tự tin hơn. Do đó, đối với nhiều truy vấn được hỏi thường xuyên hơn hoặc được hỏi với cách diễn đạt và mục đích rõ ràng hơn, Google sẽ, bằng cả tìm kiếm bằng giọng nói và tìm kiếm khác, đưa ra câu trả lời rõ ràng thậm chí không yêu cầu truy cập trang web khác để biết thêm thông tin. Đối với các tình huống



khác, ngay cả với tìm kiếm bằng giọng nói, kết quả sẽ hiện thị trên màn hình với nhiều lựa chọn và để người dùng quyết định. Điều này cũng đúng với các tìm kiếm bằng giọng nói khác như Alexa hoặc Siri. Quá trình này tương tự như của Ada. Khi điều đó xảy ra, máy của Google sẽ quan sát các lựa chọn được đưa ra và sử dụng thông tin đó để cập nhật các dự đoán của nó. Con người là một phần quan trọng của hệ thống.

Khi dự đoán đủ tốt, phán đoán và hành động rõ ràng, tự động hóa có thể được thực hiện. Nếu không, hãy để con người quyết định. Quá trình này được gọi là phán đoán theo ngoại lệ. Như tác giả Janelle Shane nhận xét, AI hoạt động tốt hơn khi có bối cảnh và mục tiêu được thu hẹp. Giải pháp thường là xây dựng khả năng phán đoán cho các tình huống mới, nhưng với một tập hợp đủ phức tạp các tình huống có thể xảy ra, sẽ luôn có một số tình huống trong đó các đề xuất của AI có vấn đề.

Từ đó, có thể kết luận rằng, chúng ta có khả năng đã hình thành phán đoán phù hợp và cũng có thể mô tả phán đoán đó cho các tình huống phổ biến hơn. Đối với những tình huống đó, có thể mã hóa đánh giá thành một quy trình tự động và thu được kết quả tốt. Ngoài những tình huống

đó – trong những tình huống ít phổ biến hơn – việc mã hóa này là không thể. Nếu chúng ta nhận ra thời điểm những tình huống đó phát sinh, thay vì mong đợi máy móc được tự động hóa để xử lý mọi tình huống, thì một giải pháp kết hợp sẽ phù hợp hơn. Điều quan trọng là khi một tình huống nằm ngoài khả năng phán đoán đã được mã hóa vào máy, thì tình huống đó sẽ được biết và truyền đạt cho con người, người sau đó sẽ quyết định phải làm gì.

Để xem lại, dự đoán AI là không hoàn hảo. Để giảm thiểu rủi ro mắc sai lầm, chúng tôi bắt tay vào hai hướng tấn công. Đầu tiên, trước khi thực tế xảy ra, chúng tôi nghiên cứu các trường hợp ngẫu nhiên và đưa ra kết luận về việc máy nên chọn gì cho từng trường hợp ngẫu nhiên đó. Thứ hai, sau khi thực tế xảy ra, chúng tôi thừa nhận rằng không phải tất cả các trường hợp ngẫu nhiên đều sẽ được giải quyết, vì vậy chúng ta sẽ dựa vào con người để can thiệp và thực hiện cuộc gọi. Khi khả năng dự đoán của AI được cải thiện, chúng ta sẽ cần phân bổ nhiều nguồn nhân lực hơn cho cả hai chức năng phán đoán này. Nói cách khác, các trường hợp ngoại lệ yêu cầu một thiết kế hệ thống bao gồm sự cộng tác của con người với máy móc.



TRÁCH NHIỆM ĐỐI VỚI PHÁN QUYẾT TRÊN DIỆN RỘNG

Đôi khi, có thể tự động hóa hoàn toàn quy trình quyết định bằng cách chỉ định trước phán quyết. Khi quẹt thẻ tín dụng, bạn đã khởi động một bộ thuật toán xác định xem giao dịch sẽ được xử lý hay từ chối. Quyết định xử lý giao dịch nào diễn ra rất lâu trước khi bạn quẹt thẻ.

Không AI nào có thể đưa ra những đánh giá đó; những phán đoán đó cũng không thể được thực hiện một cách khả thi hoặc hợp lý theo một cách phi tập trung nào đó. Thay vào đó, phán đoán được áp dụng trước hàng triệu quyết định chấp nhận hoặc từ chối và sau đó được mã hóa để sử dụng trên quy mô lớn. Máy móc không đưa ra quyết định. Thay vào đó, máy móc có thể thay đổi những người đưa ra quyết định, từ những cá nhân quyết định tại thời điểm quyết định được đưa ra, sang những cá nhân phán xét điều gì quan trọng trước khi đưa ra quyết định cụ thể.

Điều này đưa chúng ta trở lại lý do tại sao việc máy móc có thể đưa ra quyết định sa thải nhân viên lại có tác động như vậy. Máy không thể sa thải (hoặc thuê). Thay vì, một người (hoặc một nhóm) đã thực hiện phán đoán trong việc lựa chọn tiêu chí về cách máy có thể sử dụng dự đoán để đưa ra lựa

chọn đưa ra quyết định. Điều này tổng quát hơn: chưa từng có ai bị mất việc vào tay người máy. Họ mất việc vì cách ai đó quyết định lập trình robot.

Làm thế nào chúng ta đến vào thời điểm mà chúng ta có thể dễ dàng đổ lỗi cho máy móc về những gì cuối cùng là hành động của con người là một điều thú vị. Trớ trêu thay, một trong những đặc điểm chính của chủ nghĩa tư bản – được Friedrich von Hayek nhấn mạnh – là nó cho phép những người ra quyết định cá nhân hành động và, theo thuật ngữ của chúng tôi, áp dụng phán đoán cá nhân của họ cho các lựa chọn. Như nhà sử học Lewis Mumford nhận xét: “Chính vì những đặc điểm nhất định trong chủ nghĩa tư bản mà máy móc – vốn là một tác nhân trung lập – thường có vẻ là một phần tử độc hại trong xã hội, coi thường mạng sống con người, thờ ơ với lợi ích của con người. Cổ máy đã phải gánh chịu tội lỗi của chủ nghĩa tư bản”. Thật vậy, toàn bộ thuật ngữ – chủ nghĩa tư bản – dường như gọi lên sức mạnh máy móc. Trên thực tế, chính con người áp dụng phán đoán được mã hóa trong máy móc mới có sức mạnh đó. Những con người đó phải chịu trách nhiệm và hệ thống pháp luật và quản lý cần phải hiểu điều đó.

Vấn đề với tự động hóa máy móc là nó che khuất người cuối cùng chịu trách nhiệm về một



quyết định. Mã hóa phán quyết có nghĩa là các quyết định của một người có thể có quy mô phi thường. Vì nhiều lý do, chúng tôi muốn biết người đó là ai. Rốt cuộc, nếu không có trách nhiệm và sự xác định của nó, làm sao ai đó có thể chịu trách nhiệm về một quyết định? Ý nghĩa của nhu cầu thiết kế hệ thống mới là hiển nhiên khi chúng ta xem xét khả năng chuyển từ phán đoán được áp dụng tại thời điểm và địa điểm ra quyết định sang phán đoán được hệ thống hóa trước quyết định và có thể ở xa địa điểm ra quyết định. Chúng tôi sẽ phân tích điều này một cách chi tiết trong Chương 13.

Sau khi bác bỏ lập luận rằng máy móc có sức mạnh, chúng ta chuyển sang một khía cạnh khác của AI có xu hướng chạy song song với nỗi sợ hãi về sự kiểm soát của máy móc, và đó là phản hồi. Máy dự đoán là máy học. Trong một số cấu hình, chúng có thể được lập trình để tiếp tục tìm hiểu và cập nhật tự động. Đó là một phần quan trọng của giá trị của họ. Chúng có thể phát triển khi hoàn cảnh thay đổi. Nhưng đồng thời, khi có quyền lực, một cỗ máy dẫn đầu có thể dẫn đầu. Trong quá trình này, nó khó khăn hơn để cạnh tranh. Tiềm năng tích lũy quyền lực cho những người áp dụng AI là điều chúng tôi xem xét tiếp theo.

NHỮNG ĐIỂM CHÍNH

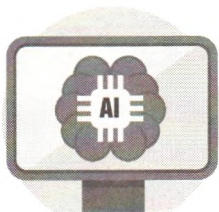
- Máy móc không thể đưa ra quyết định. Tuy nhiên, AI có thể đánh lừa mọi người nghĩ rằng máy móc đưa ra quyết định. Máy móc có thể xuất hiện để quyết định khi nào chúng ta có thể hệ thống hóa phán quyết. AI tạo ra một dự đoán, sau đó máy dựa trên phán đoán đã được hệ thống hóa của con người để thực hiện một hành động (quyết định).
- AI dự đoán là không hoàn hảo. Để giảm thiểu rủi ro mắc sai lầm, chúng tôi bắt tay vào hai hướng tấn công. Đầu tiên, trước khi triển khai AI, chúng tôi xem xét các tình huống dự phòng và đưa ra kết luận về hành động mà máy nên thực hiện cho từng tình huống. Thứ hai, sau khi triển khai AI, chúng tôi dựa vào con người để can thiệp khi AI không thể dự đoán với độ tin cậy đủ cao hoặc khi AI dự đoán một kịch bản mà chúng tôi chưa hệ thống hóa phán quyết (con người trong vòng lặp).
- Mặc dù máy móc không có sức mạnh, nhưng chúng có thể tạo ra sức mạnh thông qua quy mô và chúng có thể phân bổ lại sức mạnh bằng cách thay đổi khả năng phán đoán của chúng được sử dụng ở đâu và khi nào để ra



quyết định. Các hệ thống dựa trên AI có thể tách phán quyết ra khỏi quyết định sao cho có thể đưa ra phán quyết vào một thời điểm và địa điểm khác. Nếu sự phán đoán thay đổi từ việc mọi người triển khai riêng lẻ cho từng quyết định sang thay vào đó được hệ thống hóa thành phần mềm, thì điều này có thể dẫn đến (1) mở rộng quy mô và do đó thay đổi quyền lực do thay đổi thị phần và (2) thay đổi người thực hiện quyết định, do đó là sự thay đổi quyền lực từ bất kỳ ai từng áp dụng phán quyết sang bất kỳ ai cung cấp phán quyết đó để mã hóa hoặc sở hữu hệ thống chứa phán quyết đó.

Chương 12

TÍCH LŨY QUYỀN LỰC



Đối mới ở cấp độ hệ thống là điều không hề dễ dàng. Vậy tại sao không để các đối thủ cạnh tranh trải qua tất cả khó khăn và đánh đổi để tìm ra rồi bạn chỉ cần sao chép? Bởi AI mang lại lợi thế rất lớn cho những người đi trước. AI học hỏi, tích lũy và khi được triển khai càng sớm thì nó càng có thể bắt đầu học hỏi sớm hơn. Càng học nhiều, nó càng cải thiện độ chính xác dự báo tốt hơn. Dự báo càng tốt thì hệ thống mới càng hiệu quả. Bánh đà bắt đầu quay. Điều này giải thích tại sao một bộ phận trong cộng đồng đầu tư mạo hiểm đang đầu tư rất mạnh vào các dự án AI dường như còn non trẻ. Việc học đến từ dữ liệu và do đó, lợi thế của người đi trước đến từ vòng phản hồi trong dữ liệu đó.

BenchSci là một công ty ứng dụng AI trong y học. Mục tiêu của nó là độ dài của quá trình phát triển thuốc. Thách thức của nó là giúp các nhà khoa học tìm kim đáy bể dễ dàng hơn: thông tin cụ thể được nhúng trong vô số nghiên cứu khoa học đã



công bố và cơ sở dữ liệu nội bộ của các công ty được phẩm. Để đưa một ứng cử viên thuốc mới vào thử nghiệm lâm sàng, các nhà khoa học phải chạy thử nghiệm. BenchSci nhận ra rằng các nhà khoa học có thể thực hiện ít thí nghiệm hơn và thành công hơn nếu họ áp dụng những hiểu biết sâu sắc hơn từ vô số thí nghiệm trước đó.

Bằng cách sử dụng máy học để đọc, phân loại và sau đó trình bày những hiểu biết sâu sắc từ nghiên cứu khoa học, BenchSci nhận thấy rằng các nhà khoa học có thể thực hiện một nửa số thí nghiệm so với mức bình thường cần thiết để đưa một loại thuốc vào thử nghiệm lâm sàng. Bằng cách tìm ra các công cụ phù hợp (trong trường hợp này là thuốc thử sinh học – công cụ cần thiết để tác động và đo lường biểu hiện protein) trong tài liệu đã xuất bản thay vì khám phá lại chúng từ đầu, thời gian sản xuất các loại thuốc mới có thể giảm đáng kể. Tất cả những điều đó được thêm vào khoản tiết kiệm tiềm năng hơn 17 tỉ đô la mỗi năm. Trong một ngành mà lợi nhuận từ R&D trở nên mỏng như dao cạo, điều đó có thể biến đổi thị trường. Ngoài ra, nhiều mạng sống có thể được cứu bằng cách đưa các loại thuốc mới ra thị trường nhanh hơn.

Điều đáng chú ý nhất ở đây là BenchSci đang làm điều mà Google đã và đang làm cho toàn bộ

Internet: tìm kiếm. Nó chỉ đơn giản là làm điều đó trong một lĩnh vực chuyên biệt. Nếu không có máy học, BenchSci sẽ không thể xử lý nghiên cứu y sinh đã công bố và diễn giải nghiên cứu đó theo cách có thể chuyển thành tiết kiệm chi phí thực sự cho khách hàng của mình. Cũng giống như cách Google có thể giúp bạn tìm ra cách sửa máy rửa bát mà không cần mất nhiều thời gian đến thư viện, BenchSci giúp các nhà khoa học xác định thuốc thử phù hợp mà không cần thực hiện hàng loạt thử nghiệm. Trước BenchSci, các nhà khoa học thường sử dụng Google hoặc PubMed để tìm kiếm tài liệu (trong nhiều ngày), sau đó đọc tài liệu (trong nhiều ngày), sau đó đặt hàng và thử nghiệm từ 3-6 thuốc thử trước khi chọn một (trong nhiều tuần). Bây giờ, họ tìm kiếm BenchSci (trong vài phút), sau đó đặt hàng và thử nghiệm từ 1-3 thuốc thử trước khi chọn một (nghĩa là ít thử nghiệm hơn và ít tuần hơn).

BenchSci có nên lo lắng về sự cạnh tranh từ Google? Điều đó phụ thuộc vào khả năng tạo ra một con hào có thể phòng thủ được xung quanh một doanh nghiệp, điều này lại phụ thuộc vào bản chất của dữ liệu bên dưới AI.



DỮ LIỆU VÀ NGÀNH KINH DOANH DỰ BÁO

Để hiểu những gì cần thiết để cạnh tranh trong môi trường kinh doanh có sự tham gia của AI, trước tiên chúng ta cần xem xét những gì cần thiết để đưa ra những dự đoán tốt hơn và rẻ hơn. Không có cây đũa thần nào mà bạn có thể vẫy tay và đột nhiên có AI. Thay vào đó, nó yêu cầu xác định và quản lý các yếu tố tạo dự đoán và dữ liệu cần thiết để liên kết các yếu tố đó với nhau.

Do đó, công việc dự đoán là công việc thu thập các thuật toán và dữ liệu tốt hơn và rẻ hơn. Hãy xem xét các thuật toán. Để xây dựng thuật toán dự đoán, bạn cần đào tạo mô hình với đầu vào (ví dụ: hình ảnh) và đầu ra (ví dụ: mô tả văn bản về nội dung trong những hình ảnh đó). Điều này yêu cầu dữ liệu đào tạo. Dữ liệu đào tạo của bạn càng tốt thì các thuật toán dự đoán của bạn càng tốt. Thách thức chính mà nhiều doanh nghiệp phải đối mặt là họ phải tạo dữ liệu đào tạo mà họ cần (bằng cách thuê chuyên gia để phân loại mọi thứ) hoặc mua nó từ các nguồn khác (chẳng hạn như từ hồ sơ sức khỏe).

Dữ liệu đào tạo chỉ là phần đầu của câu chuyện. AI khác với các công cụ khác vì nó là công cụ học hỏi. Bạn càng sử dụng nó, nó càng trở nên tốt hơn. AI học hỏi từ phản hồi. Nó nhập dữ liệu và đưa ra

dự đoán. Sau đó, AI có thể quan sát xem dự đoán có thành hiện thực hay không. Nếu đúng như vậy, AI sẽ tự tin hơn vào thuật toán của mình. Nếu dự đoán không xảy ra như mong đợi, AI sẽ học cách cải thiện các dự đoán trong tương lai của nó.

Các mô hình AI thường cần được đào tạo lại trên dữ liệu mới và cập nhật do môi trường cơ bản thay đổi. Điều này có thể xảy ra đối với các ứng dụng điều hướng khi đường thay đổi và dân số di chuyển xung quanh một vị trí. Nó cũng có thể xảy ra đối với quảng cáo hướng tới mục tiêu xác định khi thói quen của người tiêu dùng thay đổi. Do đó, các mô hình AI trở nên lỗi thời và các dự đoán xấu đi theo thời gian.

Mặc dù dữ liệu đào tạo mới có thể giảm bớt vấn đề này, nhưng trong một số trường hợp, bằng cách thu thập dữ liệu mới một cách có mục đích để giải thích cho từng tình huống mới, có lẽ cách tiếp cận phù hợp nhất về mặt cạnh tranh để duy trì tính chính xác của các dự đoán trong môi trường động là liên tục cập nhật mô hình theo những gì chúng tôi gọi là *dữ liệu phản hồi*. Dữ liệu phản hồi được tạo bằng cách liên tục đo hiệu suất của các dự đoán. Để làm điều này, bạn thu thập thông tin một cách độc lập về độ chính xác của các dự đoán và ánh xạ thông tin đó tới dữ liệu đầu vào tạo ra các dự đoán



đó. Kết hợp dữ liệu này, bạn có dữ liệu phản hồi mà bạn có thể sử dụng để cập nhật thuật toán.

Chẳng hạn, khi điện thoại của bạn sử dụng hình ảnh của bạn để bảo mật, ban đầu bạn huấn luyện điện thoại nhận ra bạn. Nhưng sau đó bạn không cần phải làm lại, mặc dù khuôn mặt của bạn có thể thay đổi. Bạn có thể hoặc không thể đeo kính, hay đang trang điểm. Với những trường hợp này, dự đoán bạn là bạn có thể bắt đầu trở nên kém tin cậy hơn. Vì vậy, điện thoại của bạn sẽ cập nhật thuật toán bằng cách sử dụng hình ảnh mà bạn cung cấp cho điện thoại mỗi khi bạn mở khóa. Tất cả điều này có thể xảy ra trên điện thoại vì tất cả dữ liệu đào tạo là về bạn. Trong các tình huống khác, dữ liệu đào tạo cần được cập nhật với dữ liệu đầu vào và kết quả dự đoán từ nhiều người dùng. Trong trường hợp đó, các vấn đề về quyền riêng tư có thể trở nên nghiêm trọng và gây ra những thách thức từ việc phối hợp thông tin trên nhiều nguồn.

Tóm lại, để cạnh tranh với dự đoán, bạn cần bắt đầu với các thuật toán tốt và quyền truy cập vào dữ liệu đầu vào. Nhưng trong nhiều tình huống, bạn cũng sẽ cần quyền truy cập vào dữ liệu phản hồi. Không có gì ngạc nhiên khi chiến lược dữ liệu của bạn sẽ quyết định liệu bạn có thể cạnh tranh một cách bền vững hay không. Trong một số trường

hợp, có thể đó là lợi thế đáng kể của người đi đầu vì dự đoán chất lượng cao hơn sẽ thu hút nhiều người dùng hơn, những người này sẽ tạo ra nhiều dữ liệu phản hồi hơn, giúp cải thiện dự đoán của bạn và thu hút nhiều người dùng hơn nữa. Trong tình huống đó, các đối thủ cạnh tranh, những người không xây dựng bộ sưu tập dữ liệu phản hồi trong thiết kế của họ, có thể không bắt kịp. Các vòng phản hồi có thể tạo ra lợi thế của người đi đầu.

DỰ ĐOÁN KHẢ THI TỐI THIỂU

Những lợi thế của người đi đầu này phụ thuộc vào mức độ dự đoán cần phải tốt như thế nào để tham gia thị trường. Trong nền kinh tế công nghiệp, các nhà máy thường được xây dựng ở quy mô tối thiểu chỉ để có đủ chi phí cạnh tranh để tham gia thị trường. Đó là bởi vì, trong sản xuất, chi phí đơn vị trung bình thường giảm theo quy mô nhà máy – đến một điểm nhất định. Điểm đó được gọi là “quy mô hiệu quả tối thiểu”.

Nhiều AI cũng phải đối mặt với quy mô hiệu quả tối thiểu. Tuy nhiên, thang đo không dựa trên thông lượng của nhà máy mà dựa trên dữ liệu đào tạo và thước đo ngưỡng không phải là chi phí đơn vị mà là độ chính xác của dự đoán. Sự thành công trên thị trường của AI phụ thuộc vào độ chính xác



của các dự đoán của nó. Để trở nên hữu ích, các dự đoán phải đủ tốt để có thể thương mại hóa được. Độ chính xác của dự đoán ngưỡng có thể được đặt theo quy định (ví dụ: độ chính xác dự đoán tối thiểu cần thiết để đưa ra quyết định y tế dựa trên AI để chẩn đoán), khả năng sử dụng (ví dụ: độ chính xác dự đoán tối thiểu được yêu cầu từ dịch vụ trả lời email tự động để đảm bảo chi phí của màn hình bất động sản), hoặc cạnh tranh (ví dụ: độ chính xác dự đoán tối thiểu cần thiết để tham gia vào một thị trường hiện tại, chẳng hạn như tìm kiếm, để cạnh tranh với Google và Bing).

Viết một chương trình AI chính xác không cần đầu tư lớn vào tài sản vật chất; phần mềm không phải là một ngành kinh doanh thâm dụng vốn. Rào cản chính là dữ liệu. Để AI đủ chính xác, nó cần có đủ dữ liệu. Việc thu thập dữ liệu để đạt được quy mô hiệu quả tối thiểu này cần có thời gian và công sức. Lợi thế của việc khởi chạy trước phụ thuộc vào mức độ nỗ lực cần thiết để có một dự đoán khả thi về mặt thương mại.

Đôi khi không cần nhiều nỗ lực. Trong những năm đầu tìm kiếm trên Internet, chúng tôi có khả năng chịu lỗi cao. Các công cụ tìm kiếm cung cấp nhiều liên kết và người dùng có thể quét các liên kết đó và chọn liên kết tốt nhất. Nếu công cụ tìm

kiếm hiển thị một liên kết không liên quan, thì ít có hại. Trong những ngày đầu của Internet thương mại, điều này đã dẫn đến hàng tá tìm kiếm khác nhau công cụ, mỗi công cụ có phương pháp riêng để xác định kết quả tìm kiếm tốt nhất. Cuộc cạnh tranh rất khốc liệt.

Ngược lại, trong các phương tiện tự trị, khả năng chịu lỗi của chúng ta thấp. AI cần phải tốt hơn con người một cách có thể đo lường được để có thể tin tưởng giao phó cuộc sống của con người. Công ty đầu tiên xây dựng AI như vậy sẽ có ít sự cạnh tranh ban đầu do quy mô dữ liệu cần thiết để xây dựng AI đó. Có một số cấp bách vì AI càng sớm có thể đạt đến quy mô hiệu quả tối thiểu, thì AI càng sớm có thể bắt đầu tạo ra lợi nhuận từ các dự đoán.

Tuy nhiên, lợi thế của việc sớm đạt được quy mô hiệu quả tối thiểu này sẽ chỉ tồn tại trong thời gian ngắn nếu thị trường đang phát triển nhanh chóng. Tất cả những gì cần thiết là các công ty khác có đủ dữ liệu để xây dựng các dự đoán vượt qua ngưỡng tối thiểu để có thể thực hiện được. Quy mô hiệu quả tối thiểu không đủ để tạo ra lợi thế bền vững từ việc trở thành người đi đầu.

Lý do là, theo nghĩa kỹ thuật, có sự giảm dần lợi nhuận theo quy mô trong dữ liệu. Bạn nhận



được nhiều thông tin hơn từ lần quan sát thứ 10 so với lần quan sát thứ 100 và nhiều thông tin hơn từ lần quan sát thứ một trăm so với lần quan sát thứ 1 triệu. Khi bạn thêm các quan sát, mỗi quan sát mới có tác động nhỏ hơn đến chất lượng dự đoán của bạn.

Để dữ liệu tạo ra lợi thế lâu dài, những người đi đầu cần khai thác một lực lượng kinh tế quan trọng hơn để làm việc có lợi cho họ: dữ liệu phản hồi. Hoạt động trong lĩnh vực này, họ có thể thu thập dữ liệu phản hồi mà sau đó họ có thể sử dụng để trực tiếp cải thiện các dự đoán của mình, khiến những người khác khó bắt kịp hơn. Lợi thế không phải là tung ra khi những người khác không thể. Ưu điểm là việc khởi chạy cho phép thu thập dữ liệu phản hồi.

Việc ra mắt cũng thúc đẩy đầu tư vào phần cứng máy tính và tài năng cần thiết để tận dụng tối đa dữ liệu. Sự cạnh tranh giữa những người tham gia sớm thúc đẩy đầu tư này, nâng cao chất lượng và khiến những người khác khó cạnh tranh hơn. Hiện tượng các nhà lãnh đạo ban đầu có thể thống trị một ngành cũng tồn tại đối với nhiều ngành sử dụng nhiều công nghệ. Khi các công ty lâu đời cải thiện sản phẩm của họ, việc đầu tư vào học tập và R&D cần thiết để cạnh tranh với các công ty lâu đời

trở nên khó khăn. Ví dụ, đã từng có nhiều nhà sản xuất máy bay thương mại. Ngày nay, việc thành lập một nhà sản xuất máy bay có thể cạnh tranh với Boeing và Airbus về hiệu suất, độ an toàn và hiệu quả chi phí có thể tiêu tốn hàng chục hoặc hàng trăm tỉ đô la.

Trong cuốn sách của mình, *Technology and Market Structure* (tạm dịch: Công nghệ và Cấu trúc Thị trường), giáo sư John Sutton của Trường Kinh tế Luân Đôn đã xác định hàng chục ví dụ như vậy, từ dược phẩm đến chất bán dẫn đến sắc ký lỏng. Sự cải tiến đều đặn trong công nghệ có nghĩa là, trên thực tế, phạm vi hiệu quả tối thiểu tăng theo thời gian. Sự gia tăng này (mà Sutton gọi là “chi phí chìm nội sinh”) có thể dẫn đến sức mạnh thị trường dài hạn và do đó là một phần thưởng rất lớn cho những người đi đầu.

Điều này đã xảy ra trong quảng cáo trực tuyến và trong tìm kiếm. So với những Trang vàng hoặc một tờ báo, Google đưa ra những dự đoán tuyệt vời về việc ai muốn cái gì khi nào, điều này cho phép quảng cáo được nhắm mục tiêu. Bằng cách liên kết quảng cáo với mua hàng, Google có thể hưởng lợi từ vòng phản hồi để hệ thống tìm hiểu xem từng dự đoán có đúng hay không – và sau đó cập nhật mô hình cho lần tiếp theo – làm cho bất kỳ người chơi



mới nào cũng khó bắt kịp. Mặc dù quy mô hiệu quả tối thiểu tương đối thấp để ra mắt công cụ tìm kiếm vào những năm 1990, những cải tiến liên tục của Google – thông qua đầu tư vào phần cứng, tài năng và dữ liệu – có nghĩa là rất khó để bất kỳ công cụ tìm kiếm mới nào gia nhập thị trường ngày nay.

VÒNG PHẢN HỒI NHANH

Nếu bạn có thể đưa AI của mình vào lĩnh vực này sớm, thì AI có thể thu thập dữ liệu từ khách hàng. Dữ liệu đó sẽ làm cho các dự đoán thậm chí còn tốt hơn, tạo ra một vòng phản hồi tích cực và một rào cản gia nhập cho bất kỳ ai khác muốn cạnh tranh. Một khách hàng tiềm năng sớm có thể tăng tốc nếu vòng phản hồi đủ nhanh và dữ liệu này tiếp tục tạo ra các dự đoán tốt hơn.

Theo cách này, các máy dự đoán bổ sung những gì mà theo truyền thống là lợi thế của con người – chúng có thể học hỏi từ các kết quả. Mức độ lợi thế của AI từ việc học có liên quan đến độ trễ trong phản hồi. Khi dự đoán tỷ lệ tử vong cho bảo hiểm nhân thọ, phản hồi có thể bị trì hoãn trong nhiều thập kỷ. Trong trường hợp này, với vòng phản hồi chậm như vậy, khả năng duy trì vị trí dẫn đầu của một công ty có vị trí dẫn đầu sớm trong việc dự đoán tỷ lệ tử vong sẽ bị hạn chế. Nhưng nếu dữ liệu

phản hồi có thể được tạo nhanh chóng sau khi tạo dự đoán, thì một khách hàng tiềm năng sớm có thể chuyển thành một khách hàng tiềm năng thậm chí còn lớn hơn theo thời gian và do đó, mang lại lợi thế cạnh tranh bền vững.

Khi cho ra mắt công cụ tìm kiếm Bing vào năm 2009, Microsoft được được hậu thuẫn tuyệt đối của toàn thể công ty. Tập đoàn đã đầu tư hàng tỉ đô la vào sản phẩm mới này. Tuy nhiên, sau một thập kỷ, thị phần của Bing vẫn kém xa Google, cả về lượng tìm kiếm và doanh thu quảng cáo. Một lý do khiến Bing rất khó bắt kịp là do vòng phản hồi của nó. Trong tìm kiếm, thời gian giữa dự đoán (cung cấp một trang có một số liên kết được đề xuất để phản hồi truy vấn) và đưa ra phản hồi (người dùng nhấp vào một trong các liên kết) phải ngắn, thường là vài giây. Trong trường hợp đó, nó là vòng phản hồi mạnh. Google đã vận hành một công cụ tìm kiếm dựa trên AI trong nhiều năm, với hàng triệu người dùng và hàng tỉ lượt tìm kiếm hằng ngày. Nó thu thập nhiều dữ liệu hơn và tìm hiểu các sở thích nhanh hơn. Nội dung mới liên tục được tải lên trực tuyến, vì vậy không gian tìm kiếm không ngừng mở rộng. Mỗi khi người dùng thực hiện một truy vấn, Google sẽ đưa ra dự đoán về 10 liên kết hàng đầu và sau đó người dùng chọn liên kết tốt nhất trong số đó. Điều này cho



phép Google cập nhật mô hình dự đoán của mình, cho phép học hỏi liên tục dựa trên không gian tìm kiếm không ngừng mở rộng. Với rất nhiều dữ liệu đào tạo dựa trên nhiều người dùng hơn, Google có thể xác định các sự kiện mới và xu hướng mới nhanh hơn Bing. Cuối cùng, vòng phản hồi nhanh, kết hợp với việc tiếp tục đầu tư vào các tài sản bổ sung, chẳng hạn như các cơ sở xử lý dữ liệu khổng lồ, đã giúp Google duy trì vị trí dẫn đầu và Bing không bao giờ bắt kịp. Điều đó cũng có nghĩa là bất kỳ công cụ tìm kiếm nào khác cố gắng cạnh tranh với Google và Bing thậm chí còn chưa bao giờ bắt đầu. Các công cụ tìm kiếm như DuckDuckGo, loại bỏ cá nhân hóa để đảm bảo quyền riêng tư, phục vụ các thị trường ngách quan trọng.

Một vòng phản hồi nhanh tạo ra một cuộc đua, bởi vì nếu đối thủ cạnh tranh của bạn đã được hưởng lợi từ một vòng phản hồi như vậy, dự đoán của nó sẽ được cải thiện nhanh chóng. Vòng phản hồi nhanh khuếch đại chi phí chìm nội sinh của Sutton. Nếu tụt lại phía sau quá xa, bạn sẽ không thể bắt kịp. Hãy tưởng tượng AI đầu tiên có thể điều hướng một chiếc ô tô qua thành phố New York một cách an toàn. Sau khi đạt được sự chấp thuận theo quy định, AI đó sẽ tiếp tục thu thập dữ liệu và ngày càng tốt hơn. Khi AI thứ hai được phê duyệt, nó sẽ không có cùng số lượng và chất lượng dữ liệu.

Nó không chắc là tốt. Không có lợi thế về chi phí thực và chất lượng dự đoán thấp hơn, giá trị của người tiêu dùng đối với AI tốt thứ hai sẽ thấp hơn.

Do đó, các vòng phản hồi nhanh dẫn đến đua xe. Có thể có những lợi thế lớn khi đến sớm, vì vậy các công ty có thể đầu tư mạnh mẽ để bánh đà chạy. Theo nghĩa này, các khoản đầu tư lớn vào AI dường như chưa sẵn sàng sẽ có ý nghĩa hơn. General Motors đã trả khoảng 1 tỉ đô la cho công ty khởi nghiệp lái xe tự trị Cruise, công ty dường như có rất ít hơn một nguyên mẫu và vài chục người. Tại sao GM lại trả nhiều như vậy? Một khi bánh đà bắt đầu quay, nơi có các vòng phản hồi nhanh và chi phí chìm nội sinh, thì bất kỳ người tham gia muộn nào cũng khó có thể bắt kịp.

DỰ ĐOÁN KHÁC BIỆT

Các sản phẩm cạnh tranh thường được phân biệt. Thông thường, họ thu hút các nhóm khách hàng khác nhau. Ví dụ, Coca-Cola và PepsiCo bán các loại cola cạnh tranh có hương vị khác nhau và hình ảnh thương hiệu khác nhau. Tương tự, BMW và Mercedes-Benz bán các loại xe hạng sang cạnh tranh có phong cách khác nhau và tính năng khác nhau. Những hình ảnh và tính năng thương hiệu này thu hút những người khác nhau. Trong những



tình huống này, rất khó để định nghĩa “tốt hơn”. Coke vốn dĩ không tốt hơn Pepsi. Chỉ khác một chút thôi. Khi các sản phẩm được phân biệt, sẽ có chỗ cho các đối thủ cạnh tranh hơn là một nhà cung cấp thống trị duy nhất. Kể từ khi Coke và Pepsi tung ra thị trường hơn một thế kỷ trước, nhiều loại nước giải khát mới thành công, chẳng hạn như Red Bull và Honest Tea, đã tìm được những thị trường ngách riêng biệt và phát triển mạnh.

Tương tự, một số AI sẽ thu hút các nhóm khác nhau. Hãy xem xét một công ty đang tìm cách thay thế trung tâm cuộc gọi của mình bằng một chatbot. Khi một chatbot đủ tốt để trở nên hữu ích, có nhiều cách để xác định chatbot tốt nhất. Các công ty khác nhau sẽ có những nhu cầu khác nhau. Một công ty có thể muốn chatbot hoạt động hiệu quả, trả lời các câu hỏi của khách hàng một cách nhanh chóng. Một người khác có thể muốn tập trung vào bán hàng, chuyển đổi các truy vấn đến thành doanh thu mới. 1/3 có thể muốn chatbot trở nên thoải mái, khiến mọi người thư giãn và xua tan sự tức giận. Có lẽ vì những cách khác nhau để định nghĩa “tốt hơn”, nên có hàng chục công ty chatbot, bao gồm nhiều công ty nhỏ hơn đang tìm kiếm những góc ngách riêng biệt và có khả năng sinh lời.

Một ví dụ liên quan là phát hiện khối u ác tính. AI được xây dựng ở châu Âu dựa trên dữ liệu từ

những người có làn da sáng hơn một cách không cân xứng. Ngược lại, các AI được xây dựng ở châu Á sử dụng cơ sở dữ liệu của các bệnh nhân châu Á. Những AI này được phân biệt. Những người châu Âu tốt hơn cho người da trắng, và những người châu Á tốt hơn cho người châu Á. Mặc dù “tốt hơn” thường có nghĩa là “chính xác hơn”, nhưng các AI được phân biệt bởi vì độ chính xác trong một ngữ cảnh có thể không có nghĩa là chính xác trong một ngữ cảnh khác.

Không giống như nước giải khát, chatbot và phát hiện khối u ác tính, nhiều AI chỉ được phân biệt bằng chất lượng. Rõ ràng là một dự đoán tốt hơn có nghĩa là gì, và nó có thể được đo. Khi chất lượng được xác định rõ ràng, cũng như trong các ngành công nghiệp khác, các sản phẩm chất lượng cao nhất sẽ được hưởng lợi từ nhu cầu cao hơn. Tuy nhiên, AI khác với các ngành khác bởi vì trong hầu hết các ngành khác, chất lượng tốt hơn có giá cao hơn. Những người bán giày kém chất lượng tồn tại bằng cách tính giá thấp hơn. Điều này sẽ khó khăn trong bối cảnh của AI. AI dựa trên phần mềm. Điều này có nghĩa là, một khi mô hình được xây dựng, chi phí để tạo ra một dự đoán chất lượng cao khác cũng giống như chi phí để tạo ra một dự đoán chất lượng thấp. Nếu dự đoán tốt hơn có giá bằng



dự đoán xấu hơn, thì không có lý do gì để mua dự đoán chất lượng thấp hơn.

Như chúng tôi đã lưu ý trước đó, Google có nhiều dữ liệu hơn và hưởng lợi từ vòng phản hồi nhanh. Điều đó không đủ để tạo ra lợi thế. Nó cũng phải rõ ràng cho khách hàng thấy một tìm kiếm tốt hơn trông như thế nào. Google và Bing cung cấp kết quả tương tự cho các tìm kiếm phổ biến. Nhập từ *thời tiết* vào Google hoặc Bing và kết quả có thể giống nhau. Trường hợp Bing thất bại là trong các tìm kiếm ít phổ biến hơn. Nhập từ khóa *gián đoạn* và (tại thời điểm chúng tôi viết cuốn sách này) trang kết quả đầu tiên của Bing chỉ cung cấp các định nghĩa từ điển, trong khi Google cung cấp cả định nghĩa và liên kết đến nghiên cứu về những đổi mới mang tính đột phá. Mặc dù bắt kịp ở một số khía cạnh, Bing lại vẫn thua kém ở những khía cạnh khác. Và không có danh mục nào mà Bing được coi là vượt trội. Trong tìm kiếm, “tốt hơn” có nghĩa là tìm liên kết mà người dùng có nhiều khả năng nhấp vào và ở lại. Điều này đúng với tất cả người dùng, mặc dù các liên kết tốt nhất có thể khác nhau đối với mỗi người. Với định nghĩa rõ ràng về dự đoán và vòng phản hồi nhanh, Bing không đủ khác biệt để giành được thị phần đáng kể.

HỆ THỐNG PHẢN HỒI

Các vòng phản hồi được xây dựng có chủ ý. AI dự đoán giá trị của phản hồi đảm bảo rằng dữ liệu kết quả có thể được thu thập. Trong Chương 6, chúng ta đã thảo luận về giải pháp hệ thống của AI dự đoán nội dung học tập tốt nhất cho một cá nhân vào một ngày cụ thể. Điều này sẽ cá nhân hóa giáo dục, cho phép học sinh di chuyển với tốc độ phù hợp với họ và mọi người đều học được nhiều hơn. Chúng tôi đã thảo luận về những thách thức cấp hệ thống về phân bổ giáo viên và phát triển xã hội. Các vòng phản hồi gợi ý nhu cầu thay đổi cấp độ hệ thống hơn nữa. AI yêu cầu dữ liệu về việc liệu nội dung có cải thiện thành tích của học sinh hay không. AI nhận được dữ liệu đó càng sớm thì càng tốt. Thách thức là làm thế nào để thiết kế một chương trình giảng dạy học thuật để đảm bảo rằng học sinh hiểu sâu và ghi nhớ các khái niệm, đồng thời duy trì các vòng phản hồi đủ nhanh để cải thiện AI. Điều này sẽ yêu cầu vượt qua các rào cản pháp lý để truy cập dữ liệu của học sinh kết hợp với những tiến bộ công nghệ trong việc bảo vệ quyền riêng tư của học sinh. Giống như các phần khác của hệ thống này, phần phản hồi của giải pháp hệ thống AI cho giáo dục cá nhân hóa chưa sẵn sàng.



Mặc dù AI của giải pháp điểm tạo ra dự đoán, nhưng sức mạnh đến từ việc sớm sử dụng AI trong một ngành là kết quả của phản hồi. AI phải có quyền truy cập vào dữ liệu kết quả để học hỏi. AI lái xe tự động cần quyền truy cập vào các vụ tai nạn. Mọi hệ thống lái tự động sẽ đảm bảo loại phản hồi đó. Tai nạn, may mắn thay, rất hiếm. Để hoạt động tốt, một hệ thống lái xe tự trị sẽ cần tiếp cận với các vụ tai nạn cận kề. Càng nhiều tai nạn gần như vậy, nó có thể học càng nhanh. Điều này đòi hỏi một hệ thống xác định thời điểm một tai nạn cận kề xảy ra và sau đó xây dựng một quy trình học tập để tránh những tai nạn cận kề như vậy trong tương lai. Tránh tai nạn là không đủ. Sự thoải mái của hành khách cũng rất quan trọng, do đó, một giải pháp hệ thống AI tạo lợi thế cho những người di chuyển sớm cũng sẽ được hưởng lợi từ cách đo lường sự thoải mái. Do đó, AI có thể cần được thiết kế để học hỏi và cân nhắc nhiều biện pháp đo lường kết quả.

NGƯỜI CHIẾN THẮNG ĐƯỢC NHIỀU NHẤT

Tiềm năng của máy dự đoán là vô cùng lớn. Vòng phản hồi có nghĩa là những người tham gia sớm có lợi thế thực sự. Nhập cảnh sớm có nghĩa là có nhiều dữ liệu hơn. Nhiều dữ liệu hơn có nghĩa là dự đoán tốt hơn. Dự đoán tốt hơn có nghĩa là nhiều khách hàng hơn, từ đó dẫn đến nhiều dữ liệu hơn.

Các vòng phản hồi tạo ra một cuộc đua để triển khai AI trên quy mô lớn.

Nhưng hãy nhớ rằng các dự đoán giống như các sản phẩm được thiết kế chính xác, thích ứng cao cho các mục đích và bối cảnh cụ thể. Các công ty phân biệt được bối cảnh và mục đích, dù chỉ một chút, có thể tạo ra một không gian có thể phòng thủ được. Ấc quý, hoặc có lẽ là một thiên thần, sống trong các chi tiết của hệ thống mà dữ liệu được thu thập và sử dụng.

NHỮNG ĐIỂM CHÍNH

- Bất chấp những thách thức khi thực hiện đổi mới cấp hệ thống với AI, có một lý do chính đáng để bắt đầu nó sớm hơn là muộn: AI mang lại một lợi thế cho những người đi trước vì AI học hỏi. Nó được triển khai càng sớm, nó càng có thể bắt đầu học sớm hơn. Càng học nhiều, nó càng có độ chính xác dự đoán tốt hơn. Nó càng tốt thì hệ thống mới càng hiệu quả.
- AI là phần mềm. Vì vậy, một khi mô hình AI được xây dựng, chi phí cận biên để tạo thêm một dự đoán gần bằng không. Do đó, nếu một AI trở nên tốt hơn một chút so với các AI khác trong giai đoạn đầu phát triển thị trường, thì



sẽ có nhiều người dùng chuyển sang hệ thống có AI đó. Với nhiều người dùng hơn, AI được hưởng lợi từ nhiều dữ liệu phản hồi hơn; với nhiều dữ liệu phản hồi hơn, AI tạo ra các dự đoán thậm chí còn tốt hơn. Dự đoán tốt hơn thu hút nhiều người dùng hơn. Và như thế. Khi bánh đà bắt đầu quay, AI chỉ có lợi thế nhỏ lúc đầu có thể phát triển lợi thế lớn theo thời gian. Lợi thế đáng kể được trao cho người đi trước dẫn đến cuộc đua. Các công ty sẽ đầu tư mạnh mẽ hơn so với cái nhìn đầu tiên có vẻ hợp lý bởi vì phần thưởng cho người đầu tiên là quá lớn.

- Các vòng phản hồi có thể có ý nghĩa quan trọng đối với thiết kế hệ thống. Để AI học hỏi, nó phải có quyền truy cập vào dữ liệu kết quả. Ví dụ: một hệ thống AI giáo dục sử dụng dự đoán về nội dung tốt nhất tiếp theo để hiển thị cho người học phải được thiết kế để thu thập phản hồi thường xuyên nhất có thể, để xác định xem học viên đã học tài liệu hay chưa và để đánh giá mức độ tương tác của họ. Vì vậy, điều này sẽ không làm giảm dự đoán nội dung tốt nhất tiếp theo (giải pháp điểm) vào hệ thống giáo dục hiện có. Thay vào đó, thiết kế lại hệ thống giáo dục sẽ tạo và thu thập dữ liệu phản hồi tần số cao được đo bằng phút thay vì giữa kỳ.

PHẦN 5

AI GÂY GIÁN ĐOẠN NHƯ THẾ NÀO



Chương 13

MỘT SỰ TÁCH RỜI TUYỆT VỜI



Hỏi: “AI của bạn sẽ làm gì cho khách hàng của bạn?”

Đáp: “Nó sẽ cung cấp cho họ những hiểu biết sâu sắc”.

Nếu chúng tôi có 1 đô la cho mỗi lần một nhà sáng lập khởi nghiệp đưa ra câu trả lời đó cho các cố vấn tại Creative Destruction Lab, chúng tôi sẽ giàu có.

Thông tin chi tiết là một từ kích hoạt đối với chúng tôi vì nó thể hiện chính xác lối suy nghĩ sai lầm về cách một tiến bộ AI sẽ tạo ra giá trị. Đối với dự đoán AI mới, “thông tin chi tiết” là mã cho “chúng tôi không biết phải làm gì với dự đoán đó”.

Câu trả lời đúng sẽ là phác thảo quyết định rằng dự đoán sẽ được cải thiện. AI chỉ có giá trị nếu nó dẫn đến việc ra quyết định tốt hơn. Và điều này có nghĩa là những cơ hội mới để tạo ra giá trị

từ AI đều liên quan đến cách chúng cải thiện các quyết định.

Tin tốt là quyết định hiện diện ở khắp mọi nơi. Các quyết định là thứ đưa “cái chung” vào AI như một công nghệ có mục đích chung. Và nhu cầu ra quyết định tốt ngày càng tăng. Các ước tính cho thấy, vào năm 1960, chỉ có 5% công việc yêu cầu kỹ năng ra quyết định. Đến năm 2015, con số đó là hơn 30%. Và những công việc đó được trả lương cao hơn và có yêu cầu tuyển dụng khắt khe hơn về trình độ học vấn, kỹ năng và kinh nghiệm.

Khả năng dự đoán của AI có khả năng nâng cao giá trị của các kỹ năng ra quyết định – tức là khả năng tiếp nhận cái mà một người có thể gọi là “cái nhìn sâu sắc” và biến nó thành các quyết định tốt hơn. Tuy nhiên, như chúng tôi sẽ chứng minh trong chương này, câu hỏi quan trọng không phải là liệu ai sẽ nắm bắt những cơ hội mới để ra quyết định hay không mà là ai.

CHÌA KHÓA ĐỂ RA QUYẾT ĐỊNH LÀ SỰ PHÁN XÉT

“Nếu bạn bị đau đầu khủng khiếp, tôi đưa cho bạn một lọ thuốc và 9 viên sẽ chữa khỏi cho bạn còn một viên sẽ giết bạn, bạn có uống một viên không?”



Chủ sở hữu của Chicago Bulls, Jerry Reinsdorf, đã đặt câu hỏi giả định này cho huyền thoại bóng rổ Michael Jordan. Hầu hết chúng ta sẽ trả lời là không. Quyết định thực sự là liệu Jordan có nên trở lại thi đấu hay không khi anh ấy đang hồi phục sau chấn thương bàn chân. Đó là mùa giải thứ hai của anh ấy trong Hiệp hội Bóng rổ Quốc gia và Jordan muốn quay trở lại đó. Nhưng các bác sĩ đã nói với anh ấy rằng nếu anh ấy thi đấu, có 10% khả năng chấn thương tái phát kết thúc sự nghiệp. Jordan lập luận rằng 90% khả năng mọi thứ sẽ ổn có thể là đủ tốt. Do đó, câu hỏi liên quan đến thuốc đau đầu.

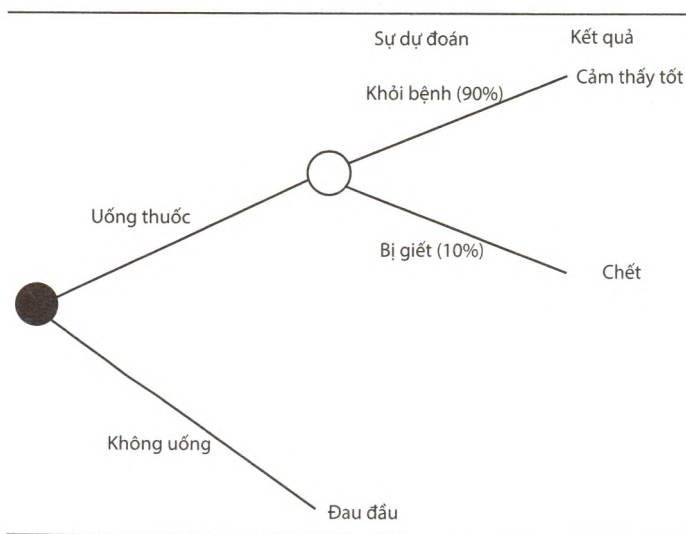
Phản ứng của Jordan với Reinsdorf về việc uống thuốc: “Điều đó phụ thuộc vào mức độ đau đầu tồi tệ của vua”.

Khi đưa ra tuyên bố này, Jordan đã lập luận rằng không chỉ xác suất – tức là dự đoán – mới quan trọng. Phần thưởng cũng quan trọng. Trong ví dụ này, phần thưởng đề cập đến đánh giá của người đó về mức độ đau liên quan đến cơn đau đầu so với việc được chữa khỏi hoặc tử vong. Phần thưởng là những gì chúng ta gọi là sự phán xét.

Để làm rõ sự khác biệt giữa dự đoán và phán đoán, như Michael Jordan đã làm, chúng tôi đưa ra một cây quyết định cho quyết định uống thuốc

trong Hình 13-1. Ở gốc cây có hai nhánh – trong một nhánh, Jordan uống viên thuốc; mặt khác, anh ta thì không. Nếu anh ta chọn uống viên thuốc, hai nhánh đại diện cho hai kết quả không chắc chắn – anh ta khỏi đau đầu hay anh ta bị giết bởi viên thuốc? Ở đầu những nhánh đó là kết quả của cảm giác tốt hay chết. Ngược lại, nếu anh ta không uống thuốc, thì thực sự không có gì là không chắc chắn. Anh ta sẽ bị đau đầu và không có nguy cơ tử vong hoặc không bị đau đầu. Như vậy, đầu cành không thuốc cũng là đầu đuôi câu chuyện, với kết cục chắc chắn Jordan phải đau đầu.

Hình 13-1: Cây quyết định uống thuốc của Michael Jordan

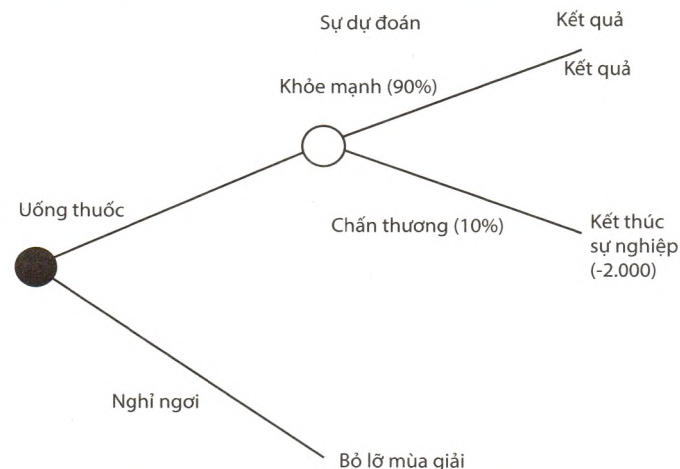




Thật dễ dàng để xếp hạng các kết quả. Cảm thấy tốt hơn là đau đầu, tốt hơn là chết. Nhưng liệu 10% khả năng tử vong có đủ để loại trừ cảm giác dễ chịu? Một kết quả mô tả là không đủ. Như Jordan đã lưu ý, cần phải có một số thước đo về cường độ để so sánh những gì bạn nhận được từ việc tránh được cơn đau đầu. Khả năng đó để quyết định mức độ quan trọng của một cái gì đó là sự phán xét.

Tất nhiên, câu hỏi giả định được thiết kế để đưa ra quyết định rõ ràng, vì khó có thể tưởng tượng được cơn đau đầu khiến bạn có cơ hội tử vong 1/10. Vì vậy, hãy xem xét quyết định thực sự của Jordan và Reinsdorf (được mô tả trong Hình 13-2). Ở đó, chúng tôi đã thêm các con số vào các kết quả cụ thể để phản ánh cường độ của chúng - nghĩa là chúng tôi đã đưa vào một biểu diễn của phán đoán. Ngoài các nhãn, cây quyết định trông giống như trong Hình 13-1. Nhưng với những con số đến từ phán quyết được thêm vào, giờ đây chúng tôi có đủ thông tin để đưa ra quyết định.. Nghi ngại chắc chắn mang lại cho Jordan -10, trong khi thi đấu mang lại 90% cơ hội là 100 và 10% cơ hội là -2.000. Do đó, bằng cách chơi, Jordan sẽ nhận được số tiền là -110 ($= 0,9 (100) + (0,1) (-2,000)$). Jordan không nên thi đấu, vì -10 tốt hơn -110.

Hình 13-2: Cây quyết định thi đấu và nghỉ ngơi của Michael Jordan



Nhưng Jordan và Reinsdorf đang tranh cãi về phán quyết. Jordan tin rằng anh ấy nên được phép thi đấu và lập luận rằng điều đó phải đáng giá 200 đô la cho anh ấy và đội để điều đó xảy ra, và cái giá phải trả khi kết thúc sự nghiệp là -1.000 đô la. Nếu đó là phán đoán chính xác, thì phần thưởng từ việc chơi sẽ trở thành $80 (= 0,9 (200) + (0,1) (-1.000))$. Sự bất đồng của họ không phải về dự đoán – mà đến từ các chuyên gia y tế. Sự bất đồng của họ là về sự phán xét.

Cuối cùng, Jordan “uống thuốc” và quay trở lại trò chơi, mặc dù có những hạn chế đáng kể về thời gian mà Reinsdorf áp đặt. Chicago Bulls cuối cùng



đã lọt vào vòng loại trực tiếp năm đó, mặc dù thua nhiều trận sớm trong mùa giải mà Jordan vắng mặt. Họ có thành tích tệ thứ hai (30-52) đối với một đội vượt qua vòng loại trực tiếp trong lịch sử. Họ đã được khen thưởng khi đối đầu với Boston Celtics hùng mạnh của Larry Bird, đội sẽ giành chức vô địch NBA năm đó. Tuy nhiên, trong Ván 2 của Vòng 1 Hội nghị Miền Đông đó, Jordan đã ghi được 63 điểm, tính đến thời điểm viết bài này, vẫn là kỷ lục ghi điểm trong một trận NBA Playoff.

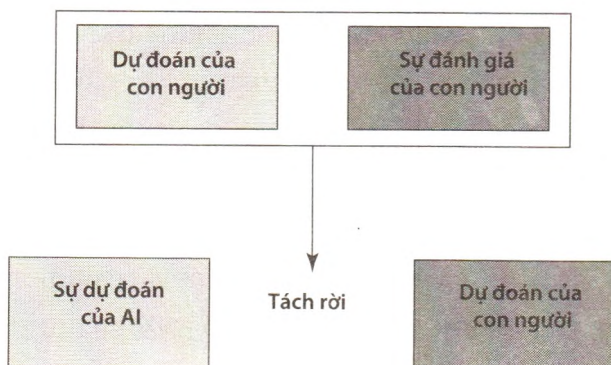
DỰ BÁO AI THÚC ĐẨY PHÁN ĐOÁN RÕ RÀNG

Sự bất đồng giữa Jordan và Reinsdorf nảy sinh vì họ đã nhận được chẩn đoán – thực tế là một dự đoán – từ một chuyên gia y tế, và cả hai đều không ở vị trí để thách thức điều đó. Nhưng hãy xem có bao nhiêu quyết định được đưa ra mà không có dự đoán rõ ràng. Chuyện gì xảy ra sau đó? Khi một người lính cứu hỏa phải lựa chọn vào lúc cứu người này hay người khác trong trường hợp khẩn cấp, họ đang xem xét các khả năng tương đối rằng họ có thể giải cứu thành công người này chứ không phải người khác hay những người đó là ai – ví dụ, một người lớn tuổi so với một đứa trẻ. Lính cứu hỏa sẽ thực hiện một cuộc gọi, nhưng mức độ chính xác mà họ đặt trọng số cho các kết quả khác nhau có

thể là ẩn ý hơn là rõ ràng. Đánh giá của chúng tôi về hiệu quả của quyết định của họ dựa trên sự kết hợp của các yếu tố.

Tuy nhiên, dự đoán của AI có khả năng đưa phần quyết định đó ra khỏi tầm tay của người ra quyết định. Dự đoán gây ra sự tách rời (xem Hình 13-3). Việc tách rời dự đoán và phán đoán không phải là một khái niệm giả định chỉ áp dụng trong lớp học mà không áp dụng trong thế giới thực. Sự tách biệt giữa dự đoán và phán đoán này làm cơ sở cho một bài báo gần đây của McKinsey về tương lai của bảo hiểm. Nó vạch ra tầm nhìn cho bảo hiểm xe hơi vào năm 2030. Một khách hàng lên xe của họ. Trợ lý kỹ thuật số cá nhân của khách hàng vạch ra các lộ trình tiềm năng. AI bên dưới trợ lý dự đoán khả năng xảy ra tai nạn và khách hàng sử dụng phán đoán của họ để quyết định.

Hình 13-3 Dự báo AI gây ra sự tách rời





Điều này có thể hoạt động như sau. Bạn đến Vancouver để làm việc và bạn đã thuê một chiếc ô tô. Bạn đang ở tại khách sạn Sutton trên phố Burrard ở trung tâm thành phố và có một cuộc họp tại Đại học British Columbia. Bạn có thể đi theo con đường ngắm cảnh dọc theo mặt nước hoặc con đường buồn tẻ dọc theo West 4th. Tuyến đường tuyệt đẹp sẽ đưa bạn qua Bãi biển Kitsilano, Công viên Jericho và Ngân hàng Tây Ban Nha. Nó thật đẹp. Tuyến đường ngắm cảnh chậm hơn một chút nhưng không nhiều. Bằng mọi cách, bạn sẽ đến cuộc họp đúng giờ.

Vấn đề thực sự là trên tuyến đường ngắm cảnh, rất nhiều người khác đang ngắm cảnh. Họ bị phân tâm. Giả sử chiếc xe cho thuê của bạn được trang bị AI cho bạn biết chính xác khả năng xảy ra tai nạn trên tuyến đường ngắm cảnh là bao nhiêu. Sau đó, bạn cần áp dụng phán đoán và đánh giá xem khung cảnh có đáng để mạo hiểm hay không. Để làm điều đó, bạn vẽ một cây quyết định và thêm các khoản hoàn trả cho các kết quả khác nhau, giống như Michael Jordan đã làm trong ví dụ trước. Bạn tính toán phần thưởng kỳ vọng của mình và quyết định đi theo con đường đẹp đẽ.

Ai làm điều đó? Không ai. Nó quá phức tạp. Đó là học thuật, theo nghĩa tồi tệ nhất của từ này. Tuyệt vời trong lý thuyết và vô dụng trong thực tế.

Nó không phải được. Bạn có thể sắp xếp lại quyết định thành một quyết định mà bạn quen thuộc và thực hiện thường xuyên. Sau khi trợ lý AI dự đoán xác suất, nó sẽ cho bạn biết giá. Nó nói rằng phí bảo hiểm của bạn sẽ tăng thêm một đô la nếu bạn chọn tuyến đường ngắm cảnh.

Điều này có vẻ như là một điều nhỏ. Người lái xe có thể quyết định tuyến đường dựa trên giá cả. Đến lượt mình, giá được xác định bởi khả năng xảy ra tai nạn và chi phí sửa chữa, những chi phí này không được khách hàng biết đến. AI đang tính toán xác suất xảy ra tai nạn và ấn định chi phí. Khách hàng chỉ nhìn thấy một mức giá.

Máy thực hiện dự đoán. Khách hàng thực hiện phán quyết. Tất cả những gì khách hàng phải làm là đánh giá xem mức giá đó có xứng đáng với lợi ích hay không. Dễ. Điều này đã và đang xảy ra. Các công ty đang định giá bảo hiểm dựa trên quyết định lái xe từng phút. Nhiều người giảm giá cho những người cài đặt ứng dụng viễn thông trên điện thoại của họ, giả định rằng khách hàng lái xe tốt. Ví dụ, Tesla không cần dựa vào dữ liệu điện thoại. Nó có thể sử dụng dữ liệu từ chính chiếc xe để đo điểm an toàn dựa trên phanh gấp, khoảng cách không an toàn sau đó và các yếu tố khác. Với việc định giá bảo hiểm dựa trên hành vi của người



lái xe, khách hàng sẽ nhận được phí bảo hiểm thấp hơn và tất cả chúng ta có thể có được những con đường an toàn hơn.

Dự đoán và phán đoán không được kết hợp. Công ty bảo hiểm định giá hành vi rủi ro. Khách hàng đánh giá liệu hành vi đó có xứng đáng hay không. Điều mà ví dụ bảo hiểm cho thấy là phán đoán có thể được tách rời khỏi dự đoán và con người chúng ta đã quen với việc đưa ra phán đoán. Đây là bản chất của cái mà các nhà kinh tế học gọi là “sự ưa thích bộc lộ”. Chúng ta có thể hiểu sở thích của ai đó từ quyết định của họ. Các nhà tiếp thị đã làm điều này trong nhiều thập kỷ.

Năm 1971, Paul Green và Vithala Rao đã xuất bản một bài báo mô tả một công cụ hoàn toàn mới để đánh giá những gì người tiêu dùng muốn. Bài báo có tiêu đề “Đo lường kết hợp để định lượng dữ liệu phán đoán”, bắt đầu bằng cách lưu ý rằng “việc định lượng phán đoán của người quản lý hoặc người tiêu dùng từ lâu đã đặt ra vấn đề cho các nhà nghiên cứu tiếp thị”. Họ nhấn mạnh rằng “nghiên cứu về quá trình ra quyết định của người tiêu dùng đòi hỏi phải xác định cách thức người mua đánh đổi các tiêu chí xung đột trong việc đưa ra quyết định”. Phương pháp này yêu cầu người tiêu dùng xếp hạng các lựa chọn khác nhau. Các lựa

chọn chỉ là giả thuyết, nhưng đủ dễ dàng vì chúng đã quen thuộc.

Green và Rao đã sử dụng ví dụ về thẻ giảm giá. Một thẻ có thể giảm giá 5% trong 10 cửa hàng và có giá 14 đô la. Một thẻ khác có thể giảm giá 10% tại năm cửa hàng và có giá 7 đô la. Thẻ thứ ba có thể được giảm giá 15% trong 10 cửa hàng và có giá 21 đô la. Bằng cách để người tiêu dùng xếp hạng sở thích của họ, một nhà thống kê có thể tìm ra giá trị mà người tiêu dùng đặt trên mỗi thẻ. Lựa chọn tiết lộ sự phán xét.

Theo thời gian, phương pháp tiên tiến. Nó được sử dụng để đánh giá giá trị của Pepperoni hoặc bánh pizza Hawaiian, xe tải Ford hoặc ô tô Toyota, và thậm chí cả sở thích của nghiên cứu sinh tiến sĩ Trung Quốc tại các trường đại học Hoa Kỳ ở lại Hoa Kỳ hoặc trở về Trung Quốc. Bằng cách hỏi các sinh viên xem họ thích vị trí nhà khoa học nghiên cứu khu vực tư nhân ở Boston với mức lương 70.000 đô la hay vị trí quản lý khu vực công ở Bắc Kinh với mức giá 50.000 đô la, các nhà nghiên cứu đã biết được đánh giá của sinh viên về giá trị tương đối của cuộc sống ở Hoa Kỳ hoặc Trung Quốc.

Khung sở thích được tiết lộ tương tự này có một luồng nghiên cứu song song trong kinh tế học, bắt đầu với công trình đoạt giải Nobel của Daniel



McFadden vào đầu những năm 1970. Nó là cơ sở cho các công cụ hiện đại để đo lường nhu cầu bằng cách sử dụng dữ liệu của máy quét tạp hóa và luồng nhập chuột trực tuyến.

Mười lăm năm trước, có lẽ nhà kinh tế học hàng đầu trong khu vực là Pat Bajari. Bajari hiện là phó chủ tịch phụ trách AI cốt lõi của Amazon và là giám đốc nhà kinh tế học. Trước khi gia nhập Amazon, Bajari là giáo sư tại các trường đại học Harvard, Stanford, Duke, Michigan và Minnesota. Ông ấy là thành viên của Hiệp hội Kinh tế lượng và đã viết các bài báo có tiêu đề khó hiểu như “Ước tính nhu cầu với người tiêu dùng không đồng nhất và Đặc điểm sản phẩm không được quan sát: Phương pháp tiếp cận khoái lạc” và “Công cụ ước tính đơn giản cho phân phối các hệ số ngẫu nhiên”. (Gợi ý: nó không đơn giản như vậy.) Bajari là một trong những nhà kinh tế lượng hàng đầu trong thế hệ của ông. Giấy tờ của ông là trù tượng. Đây các biểu tượng và phương trình. Chúng tôi không thể ngờ rằng ông ấy sẽ phát triển Amazon thành một trong những nơi tuyển dụng các tiến sĩ kinh tế lớn nhất thế giới.

Tuy nhiên, đó chính xác là những gì ông ấy đã làm. Nó liên quan nhiều đến kỹ năng của ông ấy với tư cách là một người cố vấn và một nhà lãnh đạo. Nó cũng liên quan trực tiếp đến giấy tờ của ông ấy. Ước tính nhu cầu là trọng tâm trong hoạt

động kinh doanh của Amazon. Nó cần biết những gì người tiêu dùng đánh giá cao và họ sẵn sàng trả bao nhiêu. Nếu Amazon biết đánh giá của người tiêu dùng về giá trị của một sản phẩm, thì Amazon có thể cung cấp cho họ sản phẩm phù hợp vào đúng thời điểm với mức giá hợp lý. Các công cụ để ước tính phán đoán của người tiêu dùng tồn tại trong nghiên cứu tiếp thị và kinh tế lượng. Nhóm kinh tế của Amazon, dưới sự lãnh đạo của Bajari, đã tìm ra cách xác định phán đoán đó trên quy mô lớn.

Một khi chúng ta nhận ra rằng chúng ta có thể hiểu được sự phán xét từ các quyết định, thì rõ ràng là con người chúng ta luôn đưa ra những phán đoán. Chúng ta giỏi phán đoán. Phán đoán chỉ xa lạ khi tách rời khỏi một dự đoán.

CƠ HỘI PHÁN ĐOÁN

Tách dự đoán và phán đoán tạo ra cơ hội. Điều đó có nghĩa là ai đưa ra quyết định không phải do ai dự đoán và phán đoán tốt nhất theo nhóm mà là ai đưa ra phán đoán tốt nhất bằng cách sử dụng dự đoán AI.

Một khi AI đưa ra dự đoán, thì những người có phán đoán tốt nhất có thể tỏa sáng. Như chúng tôi đã lưu ý, về mặt khái niệm và ngày càng tăng trong thực tiễn, AI có thể tiến hành dự đoán với độ chính xác cao hơn nhiều bác sĩ X-quang. Mặc dù



nó phụ thuộc vào những gì chính xác được dự đoán, nhưng trên thực tế, AI có thể được đào tạo không phải bằng cách quan sát dự đoán của bác sĩ X-quang mà bằng cách đối chiếu hình ảnh với kết quả quan sát được, đáng tin cậy – ví dụ, giải phẫu bệnh có tìm thấy khối u ác tính không? Do đó, dự đoán AI có tiềm năng trở nên vượt trội hơn so với dự đoán của con người rất nhiều, đến nỗi Vinod Khosla, nhà tiên phong công nghệ và nhà đầu tư nổi tiếng vào AI, gợi ý rằng trong tương lai, việc các bác sĩ X-quang không dựa vào dự đoán của AI có thể là một sai lầm.

Vấn đề nằm ở đây – dự đoán AI sẽ làm gì với giá trị phán đoán của bác sĩ X-quang? Với cách thức hoạt động của các bác sĩ X-quang (ít nhất là ở Hoa Kỳ), họ phần lớn không quan tâm đến các thông tin khác về bệnh nhân. Do đó, nếu AI dự đoán rằng một bệnh nhân cụ thể có khối u ác tính với xác suất 30%, thì trong những trường hợp có thể tưởng tượng được, hệ thống y tế có thể chấp nhận phán đoán của một bác sĩ X-quang rằng bệnh nhân đó nên được chẩn đoán và điều trị khối u so với của người khác mà họ không nên? Quả thực khó có thể tưởng tượng được. Thay vào đó, người ta nghi ngờ rằng, một số ủy ban gồm các chuyên gia y tế sẽ cân nhắc và tranh luận về các quy tắc chẩn đoán trước bất kỳ dự đoán nào của máy, và sau đó, phán

quyết của ủy ban đó sẽ được áp dụng trên quy mô lớn. Quyết định của bác sĩ X-quang trở nên tách rời thành dự đoán của máy và phán quyết của ủy ban.

Sau khi AI đưa ra dự đoán, các hệ thống mới có thể phát sinh để tận dụng các dự đoán tốt hơn, nhanh hơn và rẻ hơn cũng như phán đoán phù hợp hơn. Trong *AI trong cuộc cách mạng công nghệ 4.0*, chúng tôi đã nhấn mạnh cơ hội để Amazon thay đổi mô hình kinh doanh của mình để hãng vận chuyển các mặt hàng đến tận nhà trước khi bạn đặt hàng. Mô hình kinh doanh đó hiện đang tồn tại. Stitch Fix làm điều đó cho quần áo. Như Giám đốc điều hành Katrina Lake đã nói, “Chúng tôi đưa ra các lựa chọn cá nhân và độc đáo bằng cách kết hợp dữ liệu và học máy với sự đánh giá của chuyên gia con người”. Nó không dừng lại ở đó. Hàng tồn kho đắt đỏ trong ngành thời trang. Nhóm khoa học dữ liệu đã phát triển các thuật toán tích hợp các quyết định mua lại những gì sẽ có trong kho với dự đoán về những thay đổi trong nhu cầu dự kiến.

Trong Chương 11, chúng tôi đã chỉ ra rằng máy móc không có quyền lực bởi vì phán quyết đưa ra các quyết định luôn đến từ con người, ngay cả khi máy móc có thể sẽ thực hiện một quyết định. Trong chương tiếp theo, chúng ta sẽ thảo luận về các kỹ năng liên quan đến phán đoán khi nó tách rời khỏi dự đoán. Việc hiểu những kỹ năng này giúp giải



thích cách thức việc tách rời dự đoán và phán đoán thay đổi ai là người phù hợp để đưa ra quyết định. Việc tách rời tạo ra một cơ hội mới cho việc áp dụng AI tập trung vào việc cải thiện các kỹ năng liên quan đến khả năng phán đoán.

NHỮNG ĐIỂM CHÍNH

- Dự đoán và phán đoán là hai yếu tố cơ bản để ra quyết định. Trong cây quyết định, dự đoán tạo ra xác suất mà mỗi nhánh trong cây sẽ xảy ra. Sự phán xét tạo ra phần thưởng liên quan đến kết quả ở cuối mỗi nhánh. Thông thường, chúng ta đưa ra quyết định mà không nhận ra rằng dự đoán và phán đoán là hai yếu tố đầu vào riêng biệt vì cả hai đều nằm trong suy nghĩ của cùng một người (người ra quyết định). Khi chúng tôi giới thiệu AI, chúng tôi chuyển dự đoán từ con người sang máy móc và do đó chúng tôi tách dự đoán khỏi phán đoán. Điều đó có thể thay đổi người đưa ra phán quyết.
- Chúng ta luôn đưa ra quyết định và không bao giờ nghĩ đến những dự đoán hay phán đoán. Chúng ta chỉ quyết định. Mặc dù chúng ta không nghĩ rõ ràng về dự đoán và đánh giá mỗi khi đưa ra quyết định, nhưng vẫn có thể suy ra phán đoán thông

qua các kỹ thuật phân tích sau khi đưa ra quyết định (chúng tôi gọi đây là “ưu tiên được tiết lộ”). Các nhà kinh tế và tiếp thị từ lâu đã sử dụng các công cụ thống kê để đo lường sự phân xét dựa trên các lựa chọn.

- Các quyết định là các khối xây dựng chính của một hệ thống. Trước AI, sự khác biệt giữa dự đoán và phán đoán là không liên quan từ góc độ thiết kế hệ thống vì cả hai chức năng đều xảy ra bên trong một tâm trí của một người. Tuy nhiên, AI thay đổi điều đó. Khi chuyển dự đoán từ con người sang AI, chúng ta có thể suy nghĩ lại về thiết kế hệ thống. Nếu AI nhanh hơn và rẻ hơn nhiều, liệu chúng ta có thể dự đoán thường xuyên hơn không? Chúng ta có thể làm điều đó cho những quyết định ít quan trọng hơn không? Chúng ta có thể hệ thống hóa bản án và do đó tự động hóa và mở rộng quy mô quyết định không? Liệu chúng ta có thể giao vai trò phán đoán cho một người hoặc một nhóm người khác, những người có khả năng phán đoán tốt hơn đáng kể so với hệ thống trước đó, nơi phán đoán bị hạn chế đối với cùng một bộ óc đã tạo ra dự đoán không? Cơ hội cho thiết kế hệ thống mới là rất lớn vì AI tạo ra những cơ hội mới ngay từ cấp độ cơ bản nhất: thành phần quyết định.

Chương 14

TƯ DUY XÁC SUẤT



Vào năm 2018, một chiếc xe tự lái của Uber đã đâm chết một người đi bộ ở Tempe, Arizona. Đó là vụ tai nạn chết người đầu tiên của một chiếc ô tô được thiết kế để lái tự động. Chiếc xe được cho là đã nhìn thấy người đi bộ nhưng không phanh. Các bản tin dẫn lời một giáo sư Đại học Princeton nói rằng đó “một lời cảnh báo cho tất cả các công ty đang thử nghiệm xe tự hành kiểm tra hệ thống của họ để đảm bảo chúng tự động dừng khi cần thiết”. Vào ngày xảy ra tai nạn, cảnh sát trưởng Tempe có một lời giải thích khác: “Rõ ràng là rất khó tránh khỏi vụ va chạm này dựa trên cách cô ấy bước ra từ trong bóng tối”.

Uber có thực sự lập trình phương tiện của mình để giết người không? Dĩ nhiên là không. Nhưng nói xe không thấy người cũng không đúng. Thay vào đó, 6 giây trước khi va chạm, chiếc xe đã dự đoán sự hiện diện của một vật thể không xác định. Vào thời điểm chiếc xe dự đoán rằng vật thể

không xác định có khả năng là một người, thì đã quá muộn để phanh khẩn cấp tạo ra sự khác biệt.

Nói cách khác, cả hai cách giải thích đều sai vì chúng mang tính quyết định. Chiếc xe đã xác định được một vật thể và có một xác suất nhỏ rằng vật thể đó có thể là một người. Nếu sớm hơn phương tiện đã dự đoán đối tượng có khả năng là người thì nó đã phanh gấp, và thảm kịch đã tránh được. Xem xét báo cáo tai nạn cho thấy rằng chiếc xe đã cảm nhận được một người với xác suất rất thấp. Không phải bằng không, nhưng cực kỳ nhỏ. Hơn nữa, chiếc xe đã được lập trình để tiếp tục miễn là khả năng thứ gì đó là con người không quá cao. Đó có thể là 0,01%, 0,0001% hoặc 0,000000001%, nhưng không bao giờ bằng không. Đó không phải là cách máy móc hoạt động.

Đây là một kết quả khủng khiếp. Xe tự lái không nhìn thấy người đi bộ có xác suất đủ cao để phanh. Cần nhiều nỗ lực hơn để thấy rằng đây là một quyết định tồi. Sau vụ tai nạn, Uber đã đóng băng chương trình xe tự hành của mình. Khi nó tiếp tục lái xe tự động vào tháng 12 năm 2018, chương trình đã khác. Những chiếc xe này bị giới hạn ở tốc độ 25 dặm 1 giờ hoặc ít hơn, và lúc nào chúng cũng có hai tài xế đảm bảo an toàn. Nó đã thực hiện một số thay đổi khác, từ giám sát của bên thứ ba đối với



người lái xe đến các quy trình phanh tự động khác nhau. Một quyết định dựa trên ngưỡng sẽ không còn cắt giảm nó nữa.

SUY NGHĨ KHI ĐẶT CƯỢC

Sự khác biệt giữa một quyết định tồi tệ và một kết quả tồi tệ là rất quan trọng. Đôi khi quyết định tốt dẫn đến kết quả xấu. Đó là một trong những thông điệp chính của cuốn sách của Annie Duke, một người chơi poker chuyên nghiệp *Thinking in Bets* (tạm dịch: Tư duy trong cá cược). Tính đến thời điểm viết bài này, Duke là người phụ nữ duy nhất chiến thắng Giải vô địch Poker Quốc gia NBC. Poker là một trò chơi của cả may mắn và kỹ năng. Có thể chơi bài của bạn một cách hoàn hảo và thua cuộc. Cũng có thể đặt cược lớn vào một ván bài xấu và gặp may mắn.

Khi mọi thứ trở nên tồi tệ, Duke lập luận rằng điều quan trọng là phải suy nghĩ xem đó là một quyết định tồi tệ hay xui xẻo. Nếu đó chỉ là sự xui xẻo, thì hãy phân loại nó là một kết quả tồi tệ và tiếp tục. Nếu đó là một quyết định tồi, hãy rút kinh nghiệm và làm tốt hơn vào lần sau.

Quá nhiều người chơi poker nghiệp dư nhận kết quả tồi tệ và thay đổi chiến lược của họ. Tương tự như vậy, quá nhiều người thực hiện một vụ cá

cược ngu ngốc lớn và giành chiến thắng. Sau đó, họ đưa ra quyết định tiếp theo từ kết quả trong quá khứ. Thói quen mà cô ấy gọi là “kết quả” này khiến những người chơi này trở nên tồi tệ hơn theo thời gian. Không có khả năng nhận ra liệu một kết quả có phải là kết quả của sự may mắn hay không, sự không chắc chắn khiến việc học trở nên khó khăn.

Michael Jordan đã thi đấu vào cuối mùa giải 1985-1986 đó và anh ấy không bị chấn thương. Anh ấy đã đạt được một kết quả tốt. Đó là điều tốt nhất, bất kể xác suất và hậu quả tương đối của chấn thương kết thúc sự nghiệp và ngoài cả mùa giải. Sau mùa giải đó, anh ấy đã giành được 6 chức vô địch và 5 Giải thưởng Cầu thủ giá trị nhất, đồng thời trở thành vận động viên kiếm được nhiều tiền nhất mọi thời đại, ở mức 2,6 tỉ đô la. Có vẻ như chơi những trò chơi đó là đúng phán quyết. Có lẽ anh ấy nên gia nhập lại đội sớm hơn. Jordan không dính chấn thương, nhưng điều đó không có nghĩa là anh ấy đã quyết định đúng.

Suy nghĩ khi đặt cược đòi hỏi bạn phải nhận ra rằng các dự đoán là không chắc chắn và hiểu rằng kết quả mà bạn trải nghiệm một phần được quyết định bởi sự may mắn. Và điều đó không dễ dàng. Đối với ô tô, trước khi tự lái, dự đoán và phán đoán thuộc về người lái. Nếu một người lái xe tông



phải người đi bộ, chúng ta sẽ không bao giờ biết liệu họ đã dự đoán sai (họ nghĩ rằng khả năng đâm phải người đó thực sự bằng 0, vì vậy họ đã không phanh) hay do phán đoán sai (họ đã vội vàng và đánh lái cao hơn), đến đích nhanh hơn là tránh tai nạn). Nếu họ gặp tai nạn, chúng tôi cho rằng phán đoán của họ là đúng, nhưng họ đã mắc lỗi cơ học khi đưa ra dự đoán về vụ va chạm. Hiện tại, xã hội có vẻ ổn với điều này.

Khi thiết kế một chiếc ô tô tự hành, bạn có thể tính toán đến lỗi dự đoán. Nhưng sau đó, bạn phải định lượng được khả năng phán đoán, bao gồm những việc không dễ dàng và dễ chịu như tính toán giá của sự sống và so sánh điều đó với trải nghiệm của một hành khách trên ô tô (việc dừng lại quá thường xuyên trong quá trình di chuyển vì quá thận trọng là điều gây khó chịu). Mọi người đã luôn ngầm đánh đổi điều này nhưng lại từ chối khi được yêu cầu rõ ràng. Và điều đó cũng không hề bớt khó chịu hơn đối với nhóm nhân sự kỹ thuật và đạo đức khi xử lý các vấn đề liên quan đến ô tô tự hành.

NẮM BẮT SỰ KHÔNG CHẮC CHẮN

Suy nghĩ khi đặt cược có nghĩa là chấp nhận sự không chắc chắn. Chúng tôi kiểm tra khả năng cơ bản rằng một cái gì đó sẽ xảy ra. Nếu nó đủ khả

năng, hãy rẽ trái; nếu không, đi bên phải. Nói cách khác, chúng tôi biểu thị quy tắc quyết định của mình phụ thuộc vào dự đoán như một điểm cắt. Điều này hoạt động tốt nếu dự đoán là rất chính xác. Hãy nhớ lại quyết định của Michael Jordan về việc có nên thi đấu và có nguy cơ bị chấn thương thêm hay không. Nếu các bác sĩ nói rằng không có khả năng chấn thương kết thúc sự nghiệp, thì cả Jordan và Reinsdorf sẽ không do dự. Quyết định là khó khăn vì dự đoán liên quan đến sự không chắc chắn. Jordan đánh giá rằng tự tin 90% là đủ. Reinsdorf không đồng ý.

Ý tưởng đưa ra các quyết định dựa trên các ngưỡng liệt kê độ tin cậy là một ý tưởng hấp dẫn. Ví dụ, hãy xem xét quá trình đánh giá người tị nạn để được nhập cảnh vào các quốc gia, một quyết định đầy rẫy sự không chắc chắn.

Dựa trên lời khai của người yêu cầu tị nạn, các thẩm phán xét xử người tị nạn cần quyết định xem họ có thấy yêu cầu của người đó đáng tin cậy hay không và liệu người đó có bị tổn hại hay không, như được quy định trong Công ước Liên Hợp Quốc về Người tị nạn, nếu yêu cầu của họ bị từ chối. Hơn nữa, các tài liệu hỗ trợ nhất thiết phải thừa thót và các thẩm phán nhận được rất ít phản hồi về việc liệu các quyết định trước đây của họ có đúng hay không.



Hiện tại, các thẩm phán làm hết sức mình để cân nhắc các bằng chứng và quyết định. Người phân xử có xu hướng khá tự tin vào quyết định của họ. Như một học giả đã nói, “Một số người dường như nghĩ rằng trực giác của họ là người phân xử sự thật có giá trị duy nhất, và nếu trực giác mách bảo họ rằng ai đó đang nói dối, thì chắc chắn họ đang nói dối”.

Niềm tin đặt nhầm chỗ. Để đưa ra quyết định được cân nhắc kỹ lưỡng hơn, sẽ rất hữu ích nếu có sẵn một dự đoán đã gán xác suất cho khả năng một người yêu cầu nhất định có nói dối hay không. Mục tiêu sẽ là cải thiện kết quả quyết định hơn là tăng tỷ lệ yêu cầu được chấp nhận.

Hiện tại, không có dữ liệu về kết quả để xem liệu các quyết định chấp nhận hoặc từ chối người tị nạn có dẫn đến kết quả mà các thẩm phán cho rằng họ sẽ làm hay không. Nếu dữ liệu đó được thu thập, có thể xây dựng một cỗ máy dự đoán mà sau đó có thể đánh giá những người yêu cầu bồi thường trong tương lai. Với chiếc máy đó, chúng tôi sẽ có đánh giá tự tin hơn dựa trên bằng chứng. Trong một trường hợp ở Canada, một người xin tị nạn từ Đức khẳng định rằng cô ấy đang bị ban giám hiệu tại trường học của con trai mình ngược đãi và cảnh sát Đức không thể hỗ trợ cô ấy. Có rất nhiều dữ liệu

về phản ứng của cảnh sát Đức đối với các báo cáo về tội phạm, vì vậy có thể đưa ra dự đoán chắc chắn rằng cảnh sát sẽ bảo vệ cô ấy, ít nhất là theo yêu cầu đối với yêu cầu tị nạn. Người xét xử cũng có thể tin tưởng rằng bằng chứng sẽ hỗ trợ cho tuyên bố về một nhà hoạt động LGBTQ đến từ Yemen hoặc một thành viên của nhóm thiểu số bị ngược đãi ở Sudan.

Nhiều trường hợp khác để lại chỗ cho sự nghi ngờ. Thường có thông tin không đầy đủ về việc liệu lực lượng cảnh sát có đáp ứng các yêu cầu bảo vệ khỏi bạo lực gia đình hay không, hoặc liệu hồ sơ của người yêu cầu có đủ để thu hút sự chú ý của chính phủ hay không. Trong những trường hợp này, thiếu dữ liệu có nghĩa là không chắc chắn. Làm cho sự không chắc chắn đó có thể nhìn thấy được sẽ làm giảm sự tự tin thái quá.

Người xét xử cần áp dụng phán quyết bằng cách so sánh các dự đoán không chắc chắn với đánh giá về “sai lầm nào tồi tệ hơn: từ chối yêu cầu một yêu cầu tị nạn đáng lẽ phải được chấp nhận hay chấp nhận một yêu cầu đáng lẽ phải bị từ chối?” Câu hỏi có vẻ đơn giản, nhưng cũng rất rủi ro. Từ chối yêu cầu tị nạn hợp pháp có thể dẫn đến tra tấn hoặc tử vong. Chấp nhận một yêu cầu sai có nghĩa là mọi người lợi dụng sự hào phóng của một quốc gia. Theo Công ước Liên Hợp Quốc, việc chấp nhận



một yêu cầu sai thậm chí còn tồi tệ hơn. Rõ ràng có một “sai sót sai lầm” trong việc từ chối một yêu cầu chính đáng.

Bằng cách tách biệt dự đoán và phán đoán, đồng thời làm rõ ràng ngay cả với những AI tốt nhất, những dự đoán về yêu cầu tị nạn là không chắc chắn, AI có thể dẫn đến một quy trình tốt hơn. Cuối cùng, nếu cỗ máy dự đoán chấp nhận sự không chắc chắn vốn có trong dự đoán, thì sẽ có nhiều tuyên bố hơn được chấp nhận. Chi phí nói không sẽ là quá cao.

Hệ thống hiện tại không hoạt động theo cách này. Những người xét xử ít sử dụng một cỗ máy dự đoán truyền đạt sự không chắc chắn. Họ không được đào tạo để giải thích sự không chắc chắn của thống kê và luật không rõ ràng về mức độ sai lầm thực sự tồi tệ đến mức nào. Ngay cả khi những người xét xử đã được đào tạo, nó cũng không thể bị bỏ rơi vào quy trình hiện có. Chấp nhận tất cả các yêu cầu tị nạn không chắc chắn sẽ tạo ra áp lực chính trị khiến người tị nạn khó đến hơn. Nó cũng có thể tạo ra động cơ để che giấu thông tin. AI cho các yêu cầu tị nạn, mặc dù có tiềm năng cho một quy trình công bằng hơn, nhưng sẽ không khả thi nếu không thay đổi hệ thống. Một phần của hệ thống mới này sẽ là sự hiểu biết rõ ràng về phán đoán, về cách đo lường rủi ro tương đối của một sai lầm.

SỰ THIẾU PHÁN ĐOÁN HẠN CHẾ AI

Sự phán đoán là một biểu hiện của những gì bạn muốn. Nhưng nếu có một bối cảnh mới hoặc thông tin không phải là thứ bạn đã xử lý trước đây, bạn có nhất thiết phải biết mình muốn gì không? Làm thế nào một người xét xử cho các yêu cầu tị nạn giải thích tuyên bố rằng có 40% khả năng rằng một yêu cầu nhất định là hợp pháp? Trong quá khứ, các thẩm phán kết hợp dự đoán và phán đoán trong các quyết định của họ. Đối với nhiều ứng dụng mới của AI, phán đoán – tách rời khỏi dự đoán – có thể chưa tồn tại. Bởi vì không có khả năng dự đoán điều gì đang xảy ra, nên không có hành động nào có thể được thực hiện tùy thuộc vào dự đoán đó và do đó, không có lý do gì để tìm ra kết quả từ hành động đó.

Điều này có nghĩa là dự đoán và phán đoán có khả năng là vấn đề con gà và quả trứng, do đó tạo ra rào cản đối với việc áp dụng máy dự đoán và xây dựng hệ thống AI mới. Đầu tư và áp dụng các dự đoán tốt hơn chỉ có giá trị nếu bạn sử dụng được thông tin đó. Tìm ra những gì bạn có thể làm nếu theo giả thuyết bạn có dự đoán tốt hơn. Vì lý do này, bản thân việc không có khả năng phán đoán là một hạn chế đối với việc bạn sẵn sàng đầu tư vào dự đoán tốt hơn và ngược lại.



TÌM KIẾM PHÁN ĐOÁN

Phán quyết có thể được xây dựng từ việc làm việc thông qua các kết quả dự kiến của bạn trước thực tế. Bằng cách nghiên cứu, đánh giá và học hỏi từ những người khác, bạn có thể thiết lập những kết quả có thể xảy ra trong các tình huống khác nhau. Đây là cách hầu hết chúng ta học cách không chạm vào bếp nóng. Ai đó nói với bạn rằng bạn sẽ bị bỏng nếu chạm vào bếp. Bạn đã học được sự phán xét mà không tự mình trải qua vết bỏng. Người khác đã chuyển bản án. Lợi ích của việc đó là tránh mắc phải những sai lầm tốn kém trong quá trình thực hiện.

Một số bạn có thể đã hoài nghi. Trẻ em được khuyên không nên làm tất cả mọi thứ, và nhiều thứ gây ra rất ít hậu quả. Những độc giả nổi loạn hơn có thể đã chạm vào bếp nóng. Sau đó, quá trình tiến hóa bắt đầu, và bạn cảm thấy đau khi chạm vào bếp nóng. Bạn đã học được cách khác, từ kinh nghiệm.

Bạn đưa ra lựa chọn và sau đó nhận phản hồi. Kết quả cho bạn biết chi phí và lợi ích của các con đường khác nhau. Bạn càng có nhiều kinh nghiệm bằng cách đưa ra các lựa chọn khác nhau trong các bối cảnh khác nhau, kết quả của bạn càng được vạch ra cho bạn. Bức tranh mà bạn xây dựng từ trải

nghiệm đó là sự phán đoán cho phép bạn biết phải làm gì trong tương lai.

Đây là hai cách phán đoán được xây dựng. Nó liên quan đến việc học có kế hoạch từ người khác bằng cách đọc, hướng dẫn hoặc văn hóa, hoặc nó được học bằng kinh nghiệm. Bây giờ chúng ta lần lượt xem xét từng cách này.

LÊN KẾ HOẠCH

Việc có sẵn các dự đoán chất lượng hoặc chi phí thấp giúp cho nhiệm vụ thu thập phán đoán thông qua kinh nghiệm trở nên dễ dàng hơn. Nhưng nếu những dự đoán đó thay vào đó đòi hỏi một số đầu tư và phát triển thì sao? Dự đoán có thể rẻ sau sự phát triển đó, nhưng chi phí liên quan đến việc thu thập dữ liệu và đào tạo và sau đó thử nghiệm các thuật toán kết quả, có thể yêu cầu giải thích rõ ràng về việc sử dụng chúng. Tiến hành phân tích cẩn thận những lựa chọn nào có thể được đưa ra nếu có sẵn các dự đoán – tức là suy nghĩ trước về các kết quả để có được phán quyết cần thiết – có thể được đảm bảo. Ví dụ, nhiều nhà đầu tư mạo hiểm tài trợ cho các công ty khởi nghiệp với dự báo thành công không quá chắc chắn, thế nhưng trước khi thực hiện các khoản đầu tư đó, hãy cân nhắc xem nếu thành công, việc thoái vốn sẽ được thực hiện thông qua IPO hay hình thức mua (bán) lại.



Khi xem xét vấn đề này, một giả định tự nhiên là các kết quả sẽ được phân tích trong những kịch bản được coi là thường xuyên hơn so với các kịch bản ít xảy ra hơn. Tuy nhiên, khi lập kế hoạch lựa chọn cho các kịch bản khác nhau, vấn đề không hẳn là liệu dự báo có thể phân biệt giữa các kịch bản xảy ra thường xuyên và ít thường xuyên hơn hay không, mà liệu nó có thể phân biệt giữa các kịch bản khi các hành động khác nhau đáng kể được đảm bảo hay không.

Chúng ta hãy xem xét ứng dụng của AI trong gian lận thẻ tín dụng. Khi bạn quẹt thẻ tín dụng, thao tác này sẽ kích hoạt một bộ thuật toán xác định xem nên xử lý hay từ chối giao dịch. Nó có thể bị từ chối vì bạn đã hết tín dụng hoặc vì nghi ngờ gian lận. Mạng thẻ tín dụng không muốn cho phép các giao dịch mà họ nghi ngờ là gian lận vì họ sẽ phải chịu trách nhiệm về các chi phí liên quan đến hành vi trộm cắp đó. Mặt khác, toàn bộ hoạt động kinh doanh của thẻ tín dụng dựa trên việc nó trở nên dễ dàng và không gây khó khăn cho khách hàng cũng như người bán. Vì vậy, từ chối một giao dịch hợp pháp là có hại. Người tiêu dùng có thể bức bối hoặc tệ hơn là chuyển sang dùng thẻ khác.

Khi thuật toán của mạng lưới thẻ tín dụng nghi ngờ rằng một giao dịch có thể là gian lận, nó sẽ cho

điểm khả năng đó. Trên thực tế, số điểm đó thể hiện khả năng giao dịch là gian lận. Nhưng cần phải có sự phân xét để tìm ra cách phản ứng với thông tin đó. Phán quyết đó không nằm ở thương gia hay bất kỳ ai trên mặt đất. Thay vào đó, cách sử dụng thông tin được lập trình vào hệ thống và quyết định chấp nhận hay từ chối thẻ được tự động hóa. Làm sao có thể khác được?

Điều này có nghĩa là phán đoán hướng dẫn cách điểm số chuyển thành hành động chấp nhận hoặc từ chối phải được suy nghĩ trước, điều này rất có thể đến từ một ủy ban đã đánh giá các lựa chọn. Nếu điểm dự đoán luôn là 100% hoặc 0%, thì bạn không cần phải phán đoán nhiều để quyết định hướng hành động đúng đắn. Tuy nhiên, những gì bạn đang làm là thiết lập một ngưỡng cho điểm, trên đó giao dịch bị từ chối và dưới đó giao dịch được chấp nhận. Và đối với phần lớn các giao dịch, nó sẽ được chấp nhận, cho thấy rằng điểm số gian lận tương đối cao không thường xuyên xảy ra. Điều này có lẽ giải thích tại sao, trước khi tính điểm theo thuật toán, công ty thẻ tín dụng để người bán quyết định có chấp nhận thẻ hay không.

Ngưỡng đó phải được chọn để cân bằng hai lỗi. Đầu tiên là các giao dịch gian lận có thể được cho phép. Chi phí đó chỉ đơn giản là chi phí liên



quan đến công ty thẻ, chứ không phải là người bán hoặc chủ thẻ chịu chi phí giao dịch, có thể được tính toán từ dữ liệu lịch sử. Chi phí thứ hai là các giao dịch hợp pháp có thể bị từ chối. Chi phí ở đây khó tính toán hơn, và do đó, khó đưa ra phán đoán hơn. Loại chủ thẻ có thể đóng một vai trò quan trọng ở đây. Một công ty thẻ tín dụng có thể lo lắng về việc mắc phải loại lỗi này nếu chủ thẻ là khách hàng cao cấp, người có thể chuyển tất cả các giao dịch sang một thẻ khác nếu không hài lòng. Như vậy, phán quyết được áp dụng liên quan đến các đặc điểm khác của khách hàng. Điều này có thể tương tác với chính điểm gian lận. Xét cho cùng, việc dự đoán một giao dịch gian lận phụ thuộc vào việc ngoại suy điều gì đó bất thường từ giao dịch đang được điều tra. Đối với chủ thẻ cao cấp thường xuyên đi du lịch, điều này có thể khó dự đoán hơn so với chủ thẻ thông thường có mô hình tiêu dùng ổn định hơn.

Dễ dàng nhận thấy việc phán đoán bằng cách lập kế hoạch trước có thể là một bài tập phức tạp với nhiều khía cạnh khác nhau để xem xét. Những kích thích đó cần được dịch thành thứ gì đó có thể mô tả được để ít nhất là trong trường hợp thẻ tín dụng, nó có thể được mã hóa trong quy trình tự động. Với tự động hóa, phán đoán liên quan đến

việc các cá nhân đánh giá những gì quan trọng trước khi đưa ra quyết định cụ thể. Sự phức tạp này tạo ra một rào cản đối với việc áp dụng các hệ thống AI. Những người thực thi phán xét thay đổi. Thay vì các thương nhân áp dụng dự đoán kết hợp và phán đoán xem liệu ai đó có tôn trọng khoản tín dụng của họ hay không, công ty thẻ tín dụng kết hợp các dự đoán từng thời điểm với phán đoán cẩn thận và có kế hoạch trên quy mô lớn.

HÀNH TRÌNH TRẢI NGHIỆM

Kinh nghiệm, bằng cách đưa ra phán đoán – biết phải làm gì trong những trường hợp cụ thể – có thể dẫn đến những quyết định tốt hơn. Tuy nhiên, hành trình mà điều đó xảy ra có thể không đơn giản. Rốt cuộc, liệu chúng ta có trải nghiệm một cái gì đó phụ thuộc vào (a) những gì xảy ra và (b) biết rằng những gì đã xảy ra đã thực sự xảy ra. Nếu bạn vô tình chạm vào lò nướng nóng, cả hai bạn đã có một trải nghiệm mới (hy vọng) và có thể tìm ra hậu quả của việc đó. Nhưng điều đó đòi hỏi một tai nạn. Nếu bạn biết rằng lò nướng có thể nóng và không bao giờ chạm vào nó, bạn sẽ không thực sự biết hậu quả sẽ ra sao. Chúng tôi không nói rằng đó là một chiến lược tồi. Chúng tôi chỉ đưa ra quan điểm rằng các lựa chọn của bạn có thể dẫn dắt trải nghiệm của bạn.



Để làm cho điều này chính xác hơn, hãy xem xét một tình huống mà bạn có hai hành động để lựa chọn. Một hành động, mà chúng tôi sẽ gọi là hành động nguyên trạng, là điều bạn luôn làm. Bạn biết chính xác những gì bạn nhận được từ hành động đó, và một bản chất khác của hành động này là những gì bạn nhận được luôn giống nhau. Đó là một hành động với kết quả nhất định. Một hành động còn lại, mà chúng tôi sẽ gọi là hành động mạo hiểm, là điều bạn chưa bao giờ thực hiện. Đơn giản là bạn sẽ không biết điều gì sẽ xảy ra khi thực hiện hành động mạo hiểm. Chẳng hạn, đó có thể là tuyển dụng một người không đáp ứng các tiêu chí thông thường. Hoặc đó có thể là cấp vốn cho một công ty khởi nghiệp không hoàn toàn phù hợp với quan điểm đầu tư thông thường của bạn. Ở đây, ngay cả khi nhận được một số tín hiệu liên quan đến lựa chọn rủi ro giúp bạn hiểu rõ hơn bối cảnh của quyết định, bạn vẫn có thể không biết liệu nó có đáng để theo đuổi hay không.

Trong tình huống này, bạn có thể bị mắc kẹt. Dự đoán cung cấp cho bạn thông tin có thể có sẵn, nhưng nếu không biết mình sẽ làm gì với thông tin đó, bạn có thể không chọn “trả giá” cho dự đoán đó. Nhưng nếu không có dự đoán đó, bạn sẽ chỉ dính vào hiện trạng và không bao giờ tìm ra những hành

động rủi ro có sẵn cho bạn. Một lần nữa, những thách thức trong việc xây dựng khả năng phán đoán là rào cản đối với việc xây dựng hệ thống AI.

Nếu lợi ích của một giải pháp hệ thống AI đủ cao, thì sẽ rất đáng để đầu tư vào việc xây dựng khả năng phán đoán, vì vậy tình huống con gà và quả trứng này có thể không phải là vô vọng. Những người có vị trí tốt nhất để xây dựng phán đoán đó, thông qua kinh nghiệm hoặc lập kế hoạch, có thể khác với những người hiện đang đưa ra quyết định dựa trên dự đoán và phán đoán tổng hợp.

FDA CHO MỌI THỨ

Trong nhiều trường hợp, chúng ta không biết chính xác một người sẽ phản ứng như thế nào khi dùng thuốc để điều trị bệnh. Chúng ta thường biết rằng một số người sẽ phải chịu những tác dụng phụ khủng khiếp. Những người khác nhau là khác nhau, tác dụng dược phẩm là xác suất. Bởi vì ngay cả những loại thuốc tốt cũng không có tác dụng với tất cả mọi người, đôi khi rất khó để phân biệt giữa thuốc tốt và thuốc xấu. Vấn đề này có thể đã ngăn cản thị trường thuốc phát triển vượt qua những người bán dầu rắn vào đầu thế kỷ XX. Tuy nhiên, chúng tôi đã phát triển một quy trình quản lý (do Cục Quản lý Thực phẩm và Dược phẩm



đứng đầu trong trường hợp được phẩm ở Hoa Kỳ) để cân nhắc lợi ích tổng thể của từng loại thuốc đối với từng chỉ định so với chi phí. Quy trình quản lý công nhận bản chất xác suất của hiệu quả của thuốc và sử dụng các thử nghiệm ngẫu nhiên (được mô tả trong Chương 3) để xác định xem một loại thuốc có tác dụng hay không. Hơn nữa, chúng tôi xem xét tỷ lệ chi phí lợi ích trong tổng thể cũng như đối với các nhóm hẹp cụ thể trong toàn bộ phân phối tuổi và các phân phối khác. Một ví dụ gần đây là sự chấp thuận theo giai đoạn của FDA đối với vắc-xin Covid-19, đầu tiên là cho người lớn, sau đó là cho trẻ em.

Khi chúng tôi xem xét các thiết kế hệ thống mới thay thế các phương pháp xác suất bằng các phương pháp xác suất, chúng ta có thể cần xem xét các phương pháp tiếp cận quy định tương tự cho các lĩnh vực trước đây không được quy định tương tự. Ví dụ, mặc dù chúng tôi yêu cầu những người mới lái xe thực hiện một bài kiểm tra lái xe đơn giản nhưng không bao giờ kiểm tra đánh giá của họ về chi phí có thể gây hại cho người khác, chúng tôi có thể cần một cái gì đó giống như FDA cho các AI lái xe tự động để kiểm tra xem hành động của phương tiện có an toàn hay không, đến các hướng dẫn đã được thiết lập.

Tương tự như vậy, chúng ta có thể cần một cái gì đó giống như FDA cho AI cấp các khoản vay ngân hàng. Cơ quan quản lý này sẽ kiểm tra các hành động cấp khoản vay của AI để xác định xem những hành động đó có đáp ứng các yêu cầu pháp lý hay không. Ngoài ra, chúng ta có thể cần một cái gì đó giống như FDA cho AI của hệ thống điều khiển rô-bốt kho hàng, nơi rô-bốt làm việc gần con người, để kiểm tra xem hành động của rô-bốt có an toàn hay không so với một số điểm chuẩn.

Giống như ngành công nghiệp dược phẩm mang tính xác suất, khó kiểm chứng được hưởng lợi từ quy trình quản lý được thiết lập để trấn an người dân rằng, mặc dù có một số rủi ro, nhưng lợi ích nhìn chung là tích cực, chúng ta có thể cần suy nghĩ tương tự khi thiết kế các giải pháp cấp hệ thống để gặt hái toàn bộ lợi ích của AI. Khi chúng ta chuyển qua The Between Times sang một kỷ nguyên mới của AI phổ biến, nơi hầu hết các hệ thống chuyển đổi từ xác định sang xác suất, chúng ta có thể cần FDA cho mọi thứ. Các cơ quan quản lý này sẽ trở thành một phần của hệ thống mới.



AI LÀ PHÁN ĐOÁN ĐÚNG?

AI có quyền phán đoán, cách nó được tiếp thu và cách nó thực sự được triển khai trong các quyết định – cho dù là ngưỡng hay thứ gì đó phức tạp hơn – đều là những lựa chọn chính trong thiết kế hệ thống được xây dựng dựa trên khả năng dự đoán của AI. Hãy nhớ lại rằng dự đoán AI thường có nghĩa là giờ đây bạn có thể chọn ai là người đánh giá đúng thay vì kết hợp đúng giữa người dự đoán và người đánh giá, vì các chức năng này trở nên tách rời. Tuy nhiên, tách rời có nghĩa là bạn cần chọn người nhận được những dự đoán đó và cũng hiểu cách các dự đoán sẽ được sử dụng. Các dự đoán có thể đến một nơi – ví dụ: một thuật toán kết hợp phán đoán được chỉ định trước để tạo ngưỡng cho các dự đoán kích hoạt hành động – hoặc đến nhiều nơi, chẳng hạn như dự đoán cho tuyến đường điều hướng tốt nhất có sẵn cho nhiều người lái xe.

Những thay đổi này có thể gây rối, tạo ra sự bất hòa khi chúng được thực hiện. Tuy nhiên, cơ hội để suy nghĩ lại về hệ thống được tạo ra, bắt đầu bằng việc tách rời và sau đó dựa trên việc tìm kiếm những người đánh giá phù hợp. Những thẩm phán này có thể khác với những người ra quyết định ngày nay. Họ cần hiểu cách suy nghĩ khi đặt cược và có kỹ năng lập kế hoạch phù hợp, kinh nghiệm

hoặc cơ hội trải nghiệm để xây dựng phán đoán của mình một cách thích hợp.

NHỮNG ĐIỂM CHÍNH

- AI đưa tư duy xác suất vào một hệ thống. Khi chúng tôi điều tra một vụ tai nạn ô tô, chúng tôi hỏi liệu người lái xe có nhìn thấy người đi bộ trước khi họ va chạm hay không. Chúng tôi mong đợi một câu trả lời có hoặc không. Chúng ta ít quen với việc đối phó với “đại loại là” hoặc “một chút”. Tuy nhiên, đó chính xác là phản ứng mà AI cung cấp. Nó nhìn thấy thứ gì đó mà nó nghĩ là một con người đang tiến đến con đường với khả năng xảy ra là 0,01%. Khi chúng tôi đưa AI vào một hệ thống, chúng tôi thường chuyển đổi hệ thống đó từ xác định sang xác suất. Đôi khi hệ thống hiện tại được thiết kế tốt để chứa đầu vào xác suất. Trong những trường hợp khác, điều này tạo ra cơ hội nâng cao năng suất đáng kể thông qua thiết kế lại hệ thống.
- Để chuyển một dự đoán thành một quyết định, chúng ta phải áp dụng phán đoán. Nếu mọi người theo truyền thống đưa ra quyết định, thì bản án có thể không được mã hóa là khác biệt với dự đoán. Vì vậy, chúng ta cần phải tạo ra nó. Nó đến từ đâu? Nó có thể đến thông qua chuyển giao (học hỏi từ những người khác) hoặc thông qua kinh nghiệm. Nếu không có phán đoán hiện



tại, chúng ta có thể có ít động lực hơn để đầu tư vào việc xây dựng AI để dự đoán. Tương tự như vậy, chúng ta có thể ngần ngại đầu tư vào việc phát triển khả năng phán đoán liên quan đến một loạt các quyết định nếu chúng ta không có AI có thể đưa ra những dự đoán cần thiết. Chúng ta đang phải đối mặt với vấn đề con gà và quả trứng. Điều này có thể đưa ra một thách thức bổ sung cho việc thiết kế lại hệ thống.

- Để khai thác triệt để sức mạnh của AI, nhiều ứng dụng sẽ yêu cầu các giải pháp cấp hệ thống được thiết kế mới không chỉ bao gồm dự đoán và phán đoán mà còn có chức năng điều tiết được thiết kế để đảm bảo xã hội khi các hệ thống chuyển đổi từ tất định sang xác suất. Chúng tôi không biết trước hệ thống sẽ hoạt động như thế nào trong tất cả các tình huống vì nó không được mã hóa cứng. Tương tự như ngành công nghiệp dược phẩm xác suất, khó xác minh được hưởng lợi rất nhiều từ quy trình quản lý để đảm bảo với người dân rằng bất chấp rủi ro do tác dụng phụ, lợi ích tổng thể là tích cực, chúng tôi có thể cần chức năng quản lý kiểu FDA để kiểm tra máy móc quyết định chống lại một khuôn khổ thử nghiệm được thiết lập. Trong nhiều trường hợp, điều này có thể rất quan trọng đối với sự thành công của việc thiết kế lại hệ thống dựa trên thông tin xác suất.

Chương 15

NHỮNG “THẨM PHÁN” MỚI



Chì là một chất độc thần kinh chết người Ảnh hưởng đến sự phát triển não bộ của trẻ em. Theo quy định, chì phải bị loại bỏ khỏi sơn từ những năm 1960 và khỏi xăng từ những năm 1970. Hầu hết các loại sơn chứa chì đã được thay thế. Những chiếc xe sử dụng xăng pha chì cũng bị loại bỏ từ lâu. Những thay đổi này đã cải thiện sức khỏe của hàng triệu người trên thế giới.

Hoa Kỳ đã cấm sử dụng các đường ống dẫn nước uống bằng chì vào năm 1986, nhưng không áp dụng cho những đường ống đã được lắp đặt. Vì các đường ống có thể tồn tại hàng trăm năm mà không cần thay thế nên chúng sẽ tiếp tục gây hại. Một sự thay đổi trong cách xử lý nước ở Flint, Michigan đã dẫn đến lượng chì tăng đột biến trong nước uống và khiến vấn đề thay thế trở nên nghiêm trọng. Vấn đề là các quan chức thành phố không biết đường ống nào chứa chì và đường ống nào vô hại. Kiểm tra từng đường ống riêng lẻ rất tốn kém.



Sự không chắc chắn tạo cơ hội để triển khai một cỗ máy dự đoán. Các giáo sư Eric Schwartz và Jacob Abernethy của Đại học Michigan đã chấp nhận thử thách này. Với một nhóm các nhà nghiên cứu, họ bắt đầu dự đoán những đường ống nào có khả năng chứa chì. Họ đã xây dựng thành công một AI được triển khai ở Flint. Thành phố chỉ kiểm tra các đường ống trong những ngôi nhà dự đoán có thể có chì. Khi máy dự đoán xác định được một đường ống có thể khả năng nhiễm chì, tỷ lệ chính xác của dự báo này lên tới 80%. Nhờ vậy, đường ống dẫn nước của hàng nghìn cư dân Flint đã được thay thế vào năm 2016 và 2017.

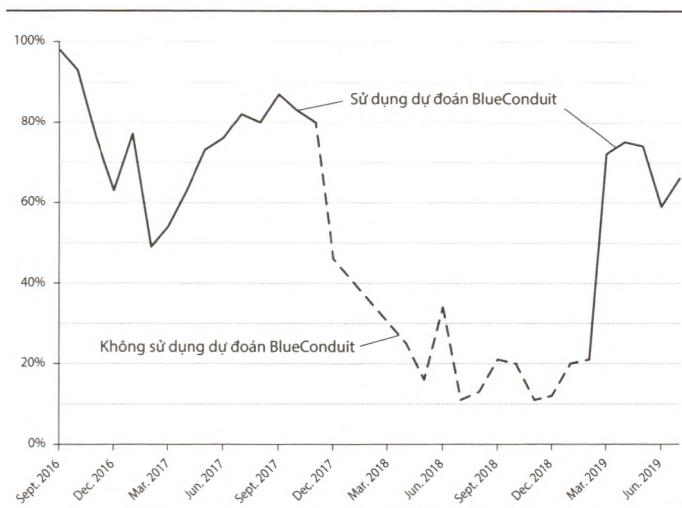
Tuy nhiên, một bộ phận cư dân tỏ ra không hài lòng. Dự đoán có thể chỉ ra rằng, chẳng hạn, chỉ một dãy nhà trong khu phố có khả năng nhiễm chì (có lẽ vì những ngôi nhà cũ hơn), khiến những cư dân khác trong khu phố lo lắng về đường ống của chính họ. Một số khu phố, những khu vực đặc biệt nghèo, có nhiều khả năng bị kiểm tra chì hơn các khu vực giàu có hơn. Một số cư dân giàu có tức giận vì đường ống của họ không được kiểm tra sớm hơn. Thị trưởng của Flint đã thuê một nhà thầu mới để quản lý việc thay thế đường ống chì, yêu cầu công ty này phải đào khắp các phường khác nhau của thành phố và trong mọi ngôi nhà mà không quan

tâm đến khả năng đường ống nước uống có thực sự bị nhiễm chì hay không.

Tỷ lệ thành công giảm mạnh xuống còn 15% (xem Hình 15-1). Trong khi nhiều cư dân hiện đã yên tâm rằng nhà của họ không có đường ống chì, quá trình xác định và thay thế đường ống bị đình trệ. AI đã đưa ra một dự đoán xuất sắc, nhưng phán quyết vẫn mang tính chính trị. Như người quản lý dự án mới đã nói, chính quyền thành phố “không muốn phải giải thích với một ủy viên hội đồng tại sao không có việc làm trong quận của họ” và “Thành phố không muốn bỏ lại bất kỳ ai”. Trong quá trình này, trong khi cư dân của một số nơi được đảm bảo rằng đường ống của họ không có chì, thì cư dân của Khu vực Năm của thành phố lại có ít cuộc khai quật nhất, mặc dù AI dự đoán rằng 80% ngôi nhà trong khu vực sẽ có đường ống dẫn bằng chì. Theo quyết định của các chính trị gia địa phương, Flint đã quyết định không sử dụng máy dự đoán.



Hình 15-1: Độ chính xác của các dự đoán cho tìm kiếm đường ống chứa chì tại Flint



Nguồn: Phòng theo dữ liệu của Jared Webb, Jacob Abernethy và Eric Schwartz, “Getting the Lead Out: Data Science and Water Service Lines in Flint” (bài nghiên cứu, Đại học Michigan, 2020), hình 3, truy cập trực tuyến tại [https://storage.googleapis.com/flint-storage-bucket/d4gx_2019%20\(2\).pdf](https://storage.googleapis.com/flint-storage-bucket/d4gx_2019%20(2).pdf) vào ngày 10 tháng 5 năm 2022.

Nhưng đó không phải là kết thúc của câu chuyện. Vào ngày 26 tháng 3 năm 2019, một thỏa thuận dàn xếp được tòa án Hoa Kỳ phê chuẩn yêu cầu thành phố sử dụng các dự đoán của Schwartz và Abernethy. Tòa án bỏ quyền tự quyết về chính trị. Thay vào đó, nó chỉ định phán xét. Về cơ bản, nó đưa ra phán quyết yêu cầu loại bỏ chì có giá trị như nhau ở mọi phường và mọi khu phố của thành phố. Điều quan trọng là khả năng có chì. Tỷ lệ

thành công nhanh chóng tăng lên 70% và hàng nghìn cư dân Flint khác đã được xác định và thay thế các ống chì của họ. Time đã gọi AI của Schwartz và Abernethy là một trong những phát minh tốt nhất của năm 2021. Hiện được tiếp thị bởi một công ty vì lợi nhuận có tên là BlueConduit, khoảng 50 thành phố sử dụng những dự đoán này để giúp họ tiết kiệm tiền trong khi xác định và loại bỏ chì khỏi hàng triệu ngôi nhà.

Điều thú vị ở đây là *sự sẵn có của một cỗ máy dự đoán mới đã gây ra cuộc tranh giành quyền quyết định như thế nào*. Các dự đoán gây bất tiện cho các chính trị gia của Flint, vì vậy họ đã từ bỏ các dự đoán. Những người khác nhận ra rằng những dự đoán có thể cải thiện cuộc sống, miễn là mỗi ngôi nhà được đánh giá là có giá trị như nhau. Sau một phiên tòa, quyền quyết định đã thay đổi. Việc giải quyết tòa án định trước bản án. Các chính trị gia địa phương mất quyền quyết định, và một hệ thống tập trung chiếm ưu thế.

Khi một cỗ máy dự đoán khiến phán đoán bị tách rời khỏi dự đoán, sẽ có cơ hội chuyển vị trí phán đoán sang người khác. Và như chúng tôi đã lưu ý, những người nắm giữ phán quyết cuối cùng sẽ quyết định. Ai nên và ai có quyền đưa ra quyết



định có thể thay đổi. Chương này xem xét thời điểm các thẩm phán mới có khả năng xuất hiện và chịu trách nhiệm về các quyết định.

Quan trọng hơn, khi làm như vậy, chúng tôi nhấn mạnh một nguồn quan trọng chống lại sự thay đổi trong việc áp dụng các hệ thống mới. Khi giới thiệu sự gián đoạn, chúng tôi lưu ý rằng thường có người thắng và người thua. Những người thua cuộc có thể là toàn bộ các bộ phận của tổ chức – chẳng hạn như nhượng quyền thương mại của Blockbuster chống lại việc phát trực tuyến video. Tổ chức đã bị cản trở trong việc áp dụng thay đổi trong nội bộ, do thay đổi về sức mạnh kinh tế. Nhưng sức mạnh kinh tế cũng phù hợp với việc ai sẽ đưa ra quyết định. Trong trường hợp của Flint, Michigan, dự đoán AI có khả năng lấy đi các quyết định từ các chính trị gia và dẫn đến mất quyền lực gây ra xích mích trong việc áp dụng chúng. Điều này đòi hỏi cơ quan ra quyết định phải thay đổi một lần nữa để loại bỏ xích mích đó. Do đó, khi chúng ta thay đổi người đưa ra quyết định, chúng ta tác động đến sự phân bổ quyền lực và chính điều đó có thể tạo ra sự phản kháng đối với việc áp dụng các hệ thống mới.

AI ĐƯỢC ĐƯA RA QUYẾT ĐỊNH?

Việc chuyển quyền phán quyết từ các chính trị gia địa phương sang danh sách ưu tiên do tòa án xác định có khả năng cải thiện cuộc sống. Những người khác nhau có những ưu đãi khác nhau và khi máy dự đoán có thể thay đổi thời gian và địa điểm của các quyết định, các cơ hội mới sẽ xuất hiện để đưa ra các quyết định tốt hơn.

Trong kinh doanh, khi chúng ta xem xét liệu một nhà quản lý cấp trên hay một trong những cấp dưới của họ có nên đưa ra một quyết định cụ thể hay không, tiêu chí chính là “ai sẽ đưa ra quyết định vì lợi ích tốt nhất của tổ chức với chi phí thấp nhất?”; đó là hiệu quả. Có rất nhiều lý do tại sao một người được phân bổ quyền quyết định vì lợi ích của hiệu quả. Một là người đó có quyền truy cập vào thông tin quan trọng sẽ hướng dẫn quyết định. Ví dụ, đây có thể là người quản lý hiện trường đưa ra quyết định về nơi triển khai tài nguyên cục bộ. Mặc dù họ có thể lấy thông tin họ thu thập và truyền đạt thông tin đó trong chuỗi, nhưng điều này có thể mất thời gian và tốn kém (đối với tất cả các bên); do đó, đôi khi chúng ta giữ quyết định gần với những người có kiến thức trực tiếp về một tình huống.

Một lý do khác là trình độ kỹ năng của người liên quan. Ra quyết định có thể khó khăn. Đặc biệt,



bạn phải giải thích thông tin và sau đó tham gia vào cả dự đoán và phán đoán để đưa ra lựa chọn. Tất cả những hoạt động đó đều cần có kỹ năng và những kỹ năng đó có thể không phổ biến. Do đó, bạn phân bổ quyền quyết định trên cơ sở kỹ năng.

Liên quan đến điều này là các ưu đãi. Bạn muốn một quyết định được đưa ra vì lợi ích của tổ chức. Tuy nhiên, khi đưa ra quyết định, mọi người có sở thích riêng ảnh hưởng đến phán đoán của họ. Điều này có thể gây ra sự lệch lạc với lợi ích của tổ chức. Mặc dù có thể sắp xếp các lợi ích bằng cách sử dụng các biện pháp khuyến khích (miễn là các yếu tố dẫn đến sự không phù hợp có thể được đo lường theo một cách nào đó), trong các tình huống khác, một số người sẽ có nhiều lợi ích phù hợp với tổ chức hơn những người khác và điều này sẽ thúc đẩy sự phân bổ của cơ quan có thẩm quyền quyết định.

Cuối cùng, các quyết định không có tác động cụ thể và đôi khi có tác động vượt ra ngoài tầm nhìn của người ra quyết định. Ví dụ, bán hàng và tiếp thị phải phù hợp với sản xuất và vận hành. Trong trường hợp này, một người hiểu được mối quan hệ giữa các quyết định khác nhau có thể là người đưa ra tất cả. Do đó, các quyết định có thể được nhóm lại với nhau để chúng có thể được thực hiện theo

cách phối hợp – ngay cả khi điều này tạo ra một số nhược điểm về sự sẵn có của thông tin, giao tiếp và thậm chí cả kỹ năng.

Toàn bộ vấn đề trao quyền quyết định cho ai đó là trao cho họ quyền đi kèm với quyết định đó. Họ cần có khả năng quyết định nơi triển khai tài nguyên, thông tin nào cần xem xét và vì lợi ích của ai để đưa ra quyết định cuối cùng. Với sức mạnh này, tự nhiên sẽ có khả năng nắm bắt nhiều giá trị hơn. Khi bạn được trao quyền quyết định, đó là vì bạn có thể tạo ra giá trị cho tổ chức bằng cách đưa ra các quyết định tốt hơn. Hơn nữa, lý do bạn, chứ không phải ai khác, có thẩm quyền đó là thông tin, kỹ năng và sự liên kết sở thích không nhất thiết phải được phổ biến rộng rãi. Nó có thể là bất cứ ai có thể đưa ra quyết định. Tuy nhiên, việc chúng ta dành thời gian xem xét ai là người đưa ra quyết định cho thấy rằng họ mang theo quyền lực.

Điều xảy ra với dự đoán AI là dự đoán và phán đoán trở nên tách rời. Quan trọng hơn, khả năng dự đoán của AI có nghĩa là khả năng phán đoán sẽ trở thành cơ sở lý giải tại sao một người nào đó được trao quyền đưa ra quyết định một cách hiệu quả hơn. Rốt cuộc, dự đoán AI phải giống nhau bất kể ai đang sử dụng nó, thứ gì đó sẽ tự động loại trừ nó như một yếu tố trong việc phân bổ quyết định.



TÀI NĂNG RA QUYẾT ĐỊNH

Con người có kỹ năng là những gì chúng ta thường gọi là *tài năng*. Vấn đề đối với nhân tài là những kỹ năng mà chúng ta đánh giá cao ở họ lại là những kỹ năng tạo ra những quyết định tốt hơn, những kỹ năng mà cho đến khi máy dự đoán ra đời, có nghĩa là khả năng phán đoán và dự đoán vượt trội. Sự sẵn có của một dự đoán máy móc đặt ra câu hỏi về những kỹ năng cụ thể mà con người đóng góp. Trong khi trước đây, thật khó để tách biệt khả năng ra quyết định tốt của con người với khả năng dự đoán và phán đoán tốt, thì cỗ máy dự đoán tập trung vào khả năng phán đoán khéo léo của con người.

Tin tốt, như đã được đề cập trong Chương 11, là sự phán đoán phải đến từ một người. Tin xấu là người đó không nhất thiết phải là người đang đưa ra phán đoán đó trước khi AI dự đoán.

Khi nào dự đoán AI sẽ thực sự thay đổi ai là người phù hợp để đưa ra phán đoán? Trong nhiều trường hợp, dự đoán của AI có thể củng cố tài năng hiện đang đưa ra quyết định, trong trường hợp đó, họ sẽ không bị gián đoạn. Tuy nhiên, đối với những người khác, phán đoán có thể nằm ở nơi khác một cách hiệu quả. Những yếu tố nào bạn nên tìm kiếm để xác định hướng thay đổi?

Việc tách rời dự đoán và phán đoán có thể thay đổi quyền lực nếu kiểu người cung cấp phán đoán hiệu quả nhất thay đổi. Sự phán xét của ai đó trở nên có giá trị hơn và của người khác trở nên kém giá trị hơn. Điều này sẽ không xảy ra trong mọi ngành và mọi tình huống. Trong nhiều trường hợp, những tiến bộ gần đây trong AI sẽ không ảnh hưởng đến việc ra quyết định. Trong các trường hợp khác, khả năng dự đoán tốt hơn sẽ cho phép các công ty cải thiện các phân tích dự đoán hiện có hoặc từng bước cải thiện các quy trình hiện có.

Tuy nhiên, đôi khi, dự đoán tốt hơn có nghĩa là vị trí phán đoán sẽ thay đổi. Khi điều này xảy ra, người ra quyết định sẽ thay đổi và quyền lực sẽ được phân bổ lại. AI sẽ có những tác động khác nhau đến giá trị của phán đoán tùy thuộc vào việc nó có được triển khai hay không trong tình huống mà trước khi có AI, dự đoán được dựa vào để đưa ra quyết định.

Tất cả các tài xế taxi phải dự đoán tuyến đường nhanh nhất giữa hai điểm tại bất kỳ thời điểm cụ thể nào. Ví dụ, tại thành phố London ở Vương quốc Anh, các tài xế taxi cần 3 năm học để học “Kiến thức” và cuối cùng được kiểm tra trí nhớ của họ về tên đường, địa điểm cơ sở và khả năng xác định tuyến đường nhanh nhất giữa hai điểm bất kỳ, bất



kỳ lúc nào, bất kỳ ngày nào trong tuần. Do đó, dự đoán là một phần cốt lõi của công việc. Vì vậy, khi một cỗ máy có thể đưa ra những dự đoán đó thông qua một ứng dụng, không có gì thay đổi đối với các tài xế taxi hiện tại, nhưng tất cả các tài xế khác giờ đây có cơ hội áp dụng phán đoán của riêng họ, dựa vào dự đoán của AI thay vì kỹ năng dự đoán của riêng họ. AI đã phá vỡ ngành công nghiệp taxi không phải vì nó thay đổi giá trị phán đoán của các tài xế taxi, mà vì nó nâng cao giá trị phán đoán của những người khác, những người hiện có thể lái xe cho Uber và Lyft. Dự đoán đã thay đổi người có thể đưa ra quyết định bằng cách mở rộng nhóm công nhân có thể thúc đẩy những người khác.

Trong trường hợp ống chì của Flint, dự đoán đã thay đổi người ra quyết định từ các chính trị gia địa phương thành thẩm phán và các bên trong vụ kiện pháp lý đã đồng ý với quyết định sửa đổi.

Thay đổi người ra quyết định có thể tạo ra sự phản đối và hoài nghi. Các nhà khí tượng học tác nghiệp dự báo thời tiết. Họ dự báo nhiệt độ hàng ngày, lượng mưa và thời tiết nguy hiểm, như bão, lốc xoáy và bão tuyết. Họ giao tiếp thời tiết, điều này đặc biệt quan trọng đối với các sự kiện thời tiết khắc nghiệt. Todd Lericos, cựu chủ tịch Hiệp hội Thời tiết Quốc gia Hoa Kỳ, lưu ý: “Những gì chúng

tôi đang làm là đánh giá rủi ro. Công chúng có nguy cơ như thế nào và chúng ta cần thông báo gì với họ để họ thực hiện các hành động giảm thiểu?”

Dự báo là bước đầu tiên, nhưng những dự báo tuyệt vời sẽ không phù hợp nếu thông tin liên lạc không thay đổi được hành vi. Hãy xem xét điều gì đã xảy ra khi một cơn lốc xoáy tấn công Joplin, Missouri, vào năm 2011. Cảnh báo lốc xoáy đã được đưa ra 4 giờ trước đó và cảnh báo, kèm theo còi báo động, xuất hiện 17 phút trước khi cơn lốc xoáy ập đến. Nhưng phần lớn cư dân Joplin được khảo sát đã không tìm nơi trú ẩn. Bi kịch thay, 158 người đã chết, và nhiều người nữa bị thương. Dự báo là một phần của quy trình để đảm bảo mọi người đưa ra quyết định đúng đắn khi thời tiết trở nên nguy hiểm.

Khi dự báo đã được cải thiện, giao tiếp đã trở nên khó khăn hơn. Giả sử nguy cơ lốc xoáy là 5%. Các nhà dự báo phải cân bằng lợi ích ngắn hạn của việc cảnh báo mọi người về nguy cơ thực sự đối với tính mạng và tài sản của họ với chi phí dài hạn của sự mệt mỏi trong cảnh báo nếu không có gì xấu xảy ra 19/20 lần. Điều này đã thay đổi vai trò hằng ngày của nhà khí tượng học hoạt động. Như Lericos đã nói,

Theo truyền thống, khách hàng chính của dịch vụ thời tiết là công chúng. Bây giờ, công chúng đang trở thành một khách hàng gián tiếp. Phần lớn



công việc nặng nhọc trong các dịch vụ của chúng tôi là làm việc với chính quyền địa phương để đưa ra các quyết định quan trọng... Họ là những người có ảnh hưởng. Bạn có thể đã nhìn thấy những trường hợp có bão tuyết mùa đông lớn và thị trường ở đó trước khi bão tuyết đi, “chúng tôi cần mọi người tránh ra đường”.

Nói cách khác, Dịch vụ thời tiết quốc gia cung cấp hỗ trợ quyết định. Như tác giả Andrew Blum lưu ý, giờ đây nó dành “nhiều thời gian hơn để giải thích cho các nhà quản lý tình trạng khẩn cấp và các quan chức công trình công cộng về khả năng xảy ra một sự kiện thời tiết và mức độ nghiêm trọng của các tác động của nó. Tạm thời, điều đó đã tạo ra nhiều công việc hơn. Trong tương lai, nó có thể là công việc duy nhất”.

Nguyên nhân của điều này là dự báo chính xác hơn. Khi dự báo sai, trường học chỉ có thể đóng cửa khi tuyết bắt đầu rơi. Dự báo tốt hơn có nghĩa là các hành động có thể diễn ra trước nhiều ngày:

Điều này đặt ra một thách thức mới: Nếu dự báo thời tiết gần như hoàn hảo, bạn có thể làm gì với nó? Làm thế nào để bạn học cách đưa ra quyết định bằng cách sử dụng nó? Trước đây, khí tượng học đã chậm giải quyết thực tế này. “Đó luôn là

vấn đề của khoa học của chúng tôi, đó là vấn đề của người khác”, [nhà khí tượng học và phó chủ tịch cấp cao của Công ty thời tiết Peter] Neilley giải thích. “Khoa học của chúng ta từ lâu đã nói, ‘Chúng ta chỉ tập trung vào độ chính xác và sau đó khi chúng ta đạt đến độ chính xác không tương, xã hội sẽ ở trong tay’. Nhưng chúng tôi đã nhận ra rằng điều đó không hoàn toàn đúng”. Công việc của họ đã được mở rộng. Nó hiện bao gồm “toàn bộ chuỗi giá trị, từ sản xuất trong dự báo bắt đầu từ các mô hình cho đến quyết định hiệu quả của một cá nhân”, Neilley nói.

Cho đến nay, các nhà khí tượng học tại các văn phòng địa phương đã làm điều này tốt nhất. Nhận cảnh báo tuyết rơi. Nơi chúng tôi sống, ở Toronto, dự báo tuyết rơi dày 1 inch qua đêm sẽ cho bạn biết nên lên kế hoạch dành thêm vài phút để phủi bụi xe trước khi ra ngoài. Dự báo tương tự 1 inch đó có thể đóng cửa Atlanta. Ở Las Vegas, nó có thể còn phức tạp hơn. Dự báo có thể chỉ phù hợp ở phía tây của thành phố, nơi có độ cao cao hơn sẽ có nhiều khả năng có tuyết hơn. Đưa ra phán quyết đúng đắn rất phức tạp và đòi hỏi sự hiểu biết về cách mọi người sống.



Lericos đã mô tả cách một giải pháp hệ thống AI cho thời tiết có thể hoạt động. Anh ấy bắt đầu bằng cách mô tả giải pháp điểm: “Dự đoán tốt hơn sẽ khiến việc truyền đạt các tác động và rủi ro trở thành một phần quan trọng hơn trong công việc của nhà khí tượng học”. Điều đó sau đó dẫn đến một sự thay đổi ở cấp độ hệ thống:

Có một biên giới khác cho AI trong khí tượng học. Lấy dự đoán thời tiết chính xác hơn làm đầu vào và sau đó kết hợp thông tin đó với dữ liệu cá nhân và xã hội khác để dự đoán tốt hơn hồ sơ rủi ro của một người (hoặc một công ty) và các hành động cần thiết để giảm thiểu rủi ro đó. Hãy tưởng tượng một thế giới nơi Cảnh báo thời tiết chung chung không còn được ban hành. Thay vào đó, các cảnh báo thời tiết được cá nhân hóa sẽ tự động được gửi đến các cá nhân hoặc công ty. Tạm biệt thời tiết con người hoạt động các nhà khí tượng học? Có vẻ như khả năng phán đoán của con người chúng ta sẽ phù hợp hơn khi làm việc để tạo ra AI có thể đưa ra các loại quyết định phù hợp để cung cấp cho khách hàng.

Làm điều này sẽ đòi hỏi sự hiểu biết về hành vi cá nhân để cho họ biết dự báo có ý nghĩa gì đối với họ. Lericos cung cấp một ví dụ. Nếu bạn sống ở phía đông của Las Vegas, bạn không cần phải thay

đổi hành vi của mình. Trừ khi trường học của con bạn ở phía tây. Trong trường hợp đó, bạn có thể dự đoán rằng trường học sẽ đóng cửa. Nếu không, bạn cần tìm ra cách để đưa chúng về nhà. Điều đó cũng phụ thuộc vào chiếc xe của bạn. Ở Nevada, rất nhiều người có xe thể thao dẫn động cầu sau. Khi các dự đoán trở nên tốt hơn, phán đoán sẽ trở nên sắc thái hơn. Lericos nhấn mạnh nhu cầu về những người có chuyên môn phù hợp, “các nhà xã hội học, chuyên gia giao thông vận tải và (có) nhà khí tượng học”. Ông thừa nhận rằng “có thể có những kịch bản mà các nhà khí tượng học không có chỗ ngồi trong bàn.... [Tôi] không phụ thuộc vào vấn đề chính xác mà bạn đang cố gắng giải quyết”. Nếu AI tốt hơn có nghĩa là nhà khí tượng học vận hành không còn là trung tâm trong các quyết định về thời tiết, thì chúng ta có thể mong đợi các nhà khí tượng đó cảnh giác với việc áp dụng AI trên quy mô lớn.

Các ví dụ về ống chì, ứng dụng điều hướng và dự báo thời tiết của Flint đều cho thấy rằng nếu giá trị phán đoán của người ra quyết định được AI nâng cao, thì họ sẽ giữ được quyền lực, trong khi nếu nó bị giảm giá trị thì họ sẽ không. Cuối cùng, nó phụ thuộc vào hiệu quả quyết định. Nếu có một cỗ máy dự đoán có nghĩa là những người có thông tin, kỹ năng, động lực và khả năng phối hợp sẽ thay đổi, thì khả năng phán đoán cũng thay đổi theo.



(PHI) TẬP TRUNG HÓA

AI cũng ảnh hưởng đến sự tập trung quyền lực. Nó có thể cho phép mở rộng khả năng phán đoán, do hệ quả trực tiếp của việc dự đoán AI là phần mềm có thể được giao tiếp và xác định qua một loạt các quyết định. Quy mô này thúc đẩy hiệu quả tiềm năng từ việc xây dựng phán đoán và mã hóa nó thành các quy trình tự động.

Để xem điều này có thể phát sinh như thế nào, hãy xem xét một số trường hợp mà dự đoán AI đã có tác động – X-quang và gian lận thẻ tín dụng. Chúng tôi đã giới thiệu phần sau trong Chương 14. Trước đây, dự đoán AI đã đe dọa công việc của các bác sĩ X-quang. Mặc dù khả năng dự đoán của AI có thể vượt trội hơn so với dự đoán của con người trong bối cảnh này, nhưng để đánh giá tác động của nó, chúng ta phải xác định xem khả năng dự đoán của AI cũng sẽ thay đổi người đưa ra phán đoán hay không.

Hãy nhớ lại phán đoán đó là kiến thức về sự đánh đổi về giá trị được tạo ra bởi các lựa chọn khác nhau. Khi bạn đưa ra quyết định dựa trên các dự đoán, đó là trọng lượng mà bạn đặt lên hậu quả của việc đưa ra quyết định sai lầm, bởi vì các dự đoán thường không hoàn hảo. Khi dự đoán AI xuất

hiện, ai có kiến thức về việc phải làm gì khi dự đoán không hoàn hảo?

Xem xét quyết định chẩn đoán đối với bệnh nhân nghi ngờ có khối u ác tính. Để chẩn đoán, bác sĩ X-quang kiểm tra hình ảnh chụp bệnh nhân. Tại Hoa Kỳ, bác sĩ X-quang thường không tiếp xúc với bệnh nhân, vì vậy dữ liệu duy nhất là hình ảnh.

Trong khi đó, trong một số trường hợp, tiên lượng sẽ rõ ràng, thì ở những trường hợp khác, bác sĩ X-quang đưa ra xác suất có khối u ác tính. Nếu họ chọn chẩn đoán rằng không có khối u ác tính khi hóa ra nó là (tức là âm tính giả), bệnh nhân sẽ không được điều trị và có thể tử vong. Điều đó dường như sẽ thúc đẩy mọi người hướng tới chẩn đoán tình trạng chỉ trong trường hợp. Nhưng tùy chọn “trong trường hợp” có chi phí riêng của nó. Chẩn đoán một bệnh nhân mắc bệnh ung thư ác tính không mắc bệnh này có nghĩa là một chế độ xét nghiệm và điều trị tiếp theo không thoải mái. Chẩn đoán cuối cùng có thể là một cuộc gọi phán xét. Hơn nữa, một phần của phán đoán đó có thể liên quan đến sự tự tin của chính bác sĩ X-quang vào khả năng của họ. Nếu bạn lo lắng rằng mình có thể bỏ lỡ điều gì đó, bạn có thể có xu hướng hướng tới kết quả dương tính giả hơn là âm tính giả. Phán đoán của bác sĩ X-quang đến từ đào tạo và kinh nghiệm. Mỗi bác



sĩ X-quang cá nhân áp dụng phán đoán của họ cho từng quyết định theo cách của họ.

Trong các tình huống khác, phán đoán không thuộc về nhiều cá nhân mà thay vào đó có thể là kết quả của một quy trình tập trung hơn, đó là điều xảy ra với mạng thẻ tín dụng. Trong trường hợp đó, một thuật toán có thể dự đoán liệu một giao dịch có gian lận hay không, nhưng phán đoán hướng dẫn cách xử lý dự đoán đó cần phải được suy nghĩ trước, mã hóa và áp dụng trên quy mô lớn. Phán quyết được cung cấp và nằm ở trung tâm.

Các ví dụ về bác sĩ chụp X-quang và mạng lưới thẻ tín dụng minh họa hai nguồn đánh giá rộng rãi. Các cá nhân có thể đưa ra đánh giá cục bộ cho từng quyết định hoặc trên toàn cầu để mở rộng quy mô cho một số lượng lớn các quyết định. Phán đoán địa phương có giá trị nhất khi bối cảnh của quyết định đối với các yếu tố địa phương là quan trọng hoặc khó hệ thống hóa phán đoán để áp dụng cho các quyết định. Đánh giá toàn cầu có giá trị nhất khi có lợi ích khi có các quy trình quyết định nhất quán trong toàn tổ chức và bối cảnh địa phương ít quan trọng hơn. Sự khác biệt giữa phán đoán có thể được điều chỉnh cho phù hợp với bối cảnh địa phương và phán đoán được áp dụng trên quy mô toàn cầu là rất quan trọng vì khi AI cải

thiện các dự đoán, nó có thể thay đổi nguồn phán đoán tốt nhất.

Điều này đã rõ ràng trong ví dụ về thẻ tín dụng. Trước thẻ tín dụng, các công cụ thanh toán chủ yếu là tiền mặt và séc. Tất nhiên, tiền mặt đáng tin cậy và khó có thể thực hiện được một cách gian lận. Séc là một vấn đề khác và các thương nhân quyết định xem họ có chấp nhận séc hay không từ một khách hàng cụ thể. Họ sẽ đưa ra dự đoán của riêng mình về khả năng tấm séc sẽ được thanh toán, nhưng cũng đối chiếu điều đó với phán đoán của riêng họ về hậu quả của việc chấp nhận séc hoặc đòi tiền mặt.

Khi các mạng thẻ tín dụng mở rộng quy mô, chúng đi kèm với một loạt dữ liệu cho phép chúng dự đoán, thậm chí trước cả AI, có nên tôn trọng thẻ hay không giao dịch. Ban đầu, việc quản lý gian lận tập trung vào việc phục hồi sau khi thực tế xảy ra. Nhưng với khả năng dự đoán tốt hơn, họ có thể mở rộng quy mô quyết định chấp nhận hoặc từ chối một cách đáng tin cậy hơn mà không cần yêu cầu người bán gọi cho các tác nhân đáng ngờ. Do đó, dự đoán tốt hơn đã dẫn đến sự thay đổi về nguồn gốc của phán đoán có thể đến từ: bối cảnh địa phương sang quy mô toàn cầu.

Các bác sĩ X-quang vẫn dự đoán và do đó áp dụng phán đoán trong bối cảnh đó. Tuy nhiên, khi



AI bắt đầu cải thiện, câu hỏi đặt ra là liệu các bác sĩ X-quang có còn là nguồn tốt nhất để phán đoán dẫn đến chẩn đoán hay không.

Bác sĩ X-quang làm nhiều hơn dự đoán. Như đã lưu ý trong Chương 8, có 30 nhiệm vụ khác nhau tạo nên quy trình làm việc của bác sĩ X-quang và chỉ một nhiệm vụ cơ bản là về các dự đoán bị ảnh hưởng trực tiếp bởi AI nhận dạng hình ảnh. Những người khác liên quan đến các hành động như tiến hành kiểm tra thể chất và đánh giá như lập kế hoạch điều trị. Một bác sĩ quang tuyến dành nhiều năm để được đào tạo sau khi tốt nghiệp trường y. Nhiều người dành những năm đó để học cách diễn giải hình ảnh. Khi AI thực hiện dự đoán, câu hỏi đặt ra là ai là người có vị trí tốt nhất để đảm nhận các nhiệm vụ khác mà các bác sĩ X-quang hiện đang làm. Nó có thể cần phải là một chuyên gia y tế hiểu được bối cảnh địa phương. Dự đoán tốt hơn trong hình ảnh y tế có thể có nghĩa là nhiều chuyên gia y tế hơn có thể sử dụng khả năng phán đoán, thay đổi quyền ra quyết định từ bác sĩ X-quang sang nhiều chuyên gia y tế hơn.

Dự đoán có thể tăng hoặc giảm mức độ tập trung của người đánh giá.

PHÁN ĐOÁN VÀ KIỂM SOÁT

Chúng ta đã thảo luận về hai khía cạnh phán đoán và kiểm soát. Đôi khi, việc tách biệt giữa dự đoán và phán đoán có nghĩa là những người khác nhau thực hiện phán đoán và đôi khi đó là cùng một người. Đôi khi, việc kiểm soát việc ra quyết định trở nên tập trung hơn và đôi khi ít hơn (Hình 15-2).

Hình 15-2: Đánh giá và kiểm soát

		Kiểm soát việc ra quyết định	
		Tập trung hơn	Ít tập trung hơn
Mọi người thực thi sự đánh giá	Tương đương	Hỗ trợ khách hàng, tuyển dụng	Hình ảnh y tế
	Khác biệt	Thẻ tín dụng, ống dẫn nước ở Flint	Tài xế Uber, khí tượng học

Khi một vài người đưa ra phán quyết hiệu quả nhất và AI có nghĩa là những người đó nên khác với những người hiện đang đưa ra quyết định, thì sẽ có khả năng xảy ra gián đoạn. Chúng tôi đã chứng kiến sự gián đoạn này trong ngành thanh toán, nơi phán quyết của các nhà quản lý tập trung mạng thay thế hàng triệu thương nhân. Nó xảy ra ở Flint, khi những cách làm việc cũ không bền vững khi đối mặt với dữ liệu. Khi những người khác nhau đánh giá, ngay cả các giải pháp về điểm AI cũng sẽ gặp phải sự phản kháng, điều này làm chậm sự lan tỏa của AI, kéo dài thời gian dành cho The Between



Times. Nó cũng củng cố khả năng gây gián đoạn và nhu cầu về các giải pháp hệ thống AI. Nếu những người có quyền lực trong hệ thống hiện tại không từ bỏ vị trí của họ, thì một hệ thống mới – có lẽ được phát triển bởi một nhà kinh doanh hệ thống AI dám nghĩ dám làm – sẽ phân bổ phán quyết cho những người có vị trí tốt nhất để cung cấp nó.

Tuy nhiên, ngay cả khi có những thẩm phán phù hợp, như chúng ta sẽ thấy trong Phần 6, điều quan trọng là phải nhận ra rằng các hệ thống bao gồm các quyết định liên quan đến nhau. Do đó, việc thay đổi cách thực hiện một quyết định có thể gây ra hậu quả cho nhiều quyết định khác. Điều này có hậu quả đối với thiết kế hệ thống.

NHỮNG ĐIỂM CHÍNH

- Khi việc triển khai AI dẫn đến việc tách rời dự đoán và phán đoán, có thể có cơ hội tăng cường tạo ra giá trị, nhưng điều đó có thể yêu cầu thiết kế lại hệ thống theo cách chuyển vị trí phán đoán từ những người ra quyết định hiện tại sang những người khác. Khi điều này xảy ra, quyền lực được phân bổ lại. Những người đưa ra phán quyết cuối cùng sẽ quyết định và do đó có quyền lực. Thiết kế hệ thống mới tận

dụng AI có thể làm giảm sức mạnh của một số cá nhân, những người do đó có thể chống lại sự thay đổi.

- Khi thiết kế một hệ thống mới, chúng ta phân bổ quyền quyết định như thế nào? Chúng ta chọn người hoặc nhóm có nhiều khả năng đưa ra quyết định vì lợi ích tốt nhất của tổ chức với chi phí thấp nhất. Đó là hiệu quả quyết định. Có 4 yếu tố chính cần xem xét: (1) thông tin: Ai có quyền truy cập hoặc ai nên được cấp quyền truy cập thông tin cần thiết để đưa ra quyết định?; (2) kỹ năng: Ai có kỹ năng và chuyên môn cần thiết để đưa ra quyết định?; (3) ưu đãi: Ai có ưu đãi phù hợp nhất với lợi ích của tổ chức đối với quyết định cụ thể này?; và (4) phối hợp: Nếu quyết định ảnh hưởng đến nhiều bộ phận của tổ chức, thì ai có thẩm quyền, thông tin và động lực cần thiết trong toàn tổ chức để đưa ra quyết định phù hợp nhất với lợi ích chung của tổ chức? Câu trả lời cho những câu hỏi này có thể rất khác khi yêu cầu là dự đoán cộng với phán đoán so với chỉ phán đoán, bởi vì AI đang đưa ra dự đoán.
- Thiết kế hệ thống mới có thể tập trung quyền lực nếu phán đoán có thể mã hóa được và do đó có thể mở rộng. Mạng lưới thẻ tín dụng và



khoa X-quang là hai ví dụ. Trong trường hợp mạng lưới thẻ tín dụng, quyền lực tập trung vào một vài công ty thẻ tín dụng thay vì tập trung vào nhiều thương nhân như trường hợp trước đây. Trong trường hợp chụp X-quang, một số người suy đoán rằng kỹ năng chính của nhận dạng mẫu và phát hiện điểm bất thường trong hình ảnh y tế phụ thuộc vào mức độ tập trung của dự đoán trong giải pháp AI. Trong trường hợp đó, nếu các bác sĩ X-quang không còn cần đến kỹ năng dự đoán của họ, liệu họ có phải là người thích hợp nhất để đưa ra phán đoán? Nếu không, y tá, nhân viên xã hội hoặc các chuyên gia chăm sóc sức khỏe được đào tạo khác có thể đưa ra phán quyết.

PHẦN 6

HÌNH DUNG CÁC HỆ THỐNG MỚI



Chương 16

THIẾT KẾ CÁC HỆ THỐNG

ĐÁNG TIN CẬY



Thomas Schelling, nhiều thập kỷ trước khi đoạt giải Nobel kinh tế, đã đưa ra thí nghiệm tưởng tượng sau:

Bạn sắp gặp ai đó ở thành phố New York. Bạn chưa được hướng dẫn địa điểm gặp mặt; bạn không biết trước với người đó về địa điểm gặp mặt; và bạn không thể giao tiếp với nhau. Bạn chỉ được thông báo rằng bạn sẽ phải đoán nơi gặp nhau, anh ấy cũng được cho biết điều tương tự và bạn sẽ phải cố gắng làm cho những dự đoán của mình trùng khớp.

Ngày nay, khi chúng ta đưa điều này cho học sinh, họ thậm chí không thể hiểu vấn đề này sẽ phát sinh như thế nào. Chắc chắn, bạn chỉ cần nhấn tin cho họ? Nhưng trước đây, đó là một câu hỏi học búa quen thuộc hơn. Và toàn bộ vấn đề là

để xem bạn có thể làm gì nếu không có khả năng giao tiếp.

Trong lớp, khoảng trước năm 2000, nếu hỏi sinh viên ở New York, họ sẽ nhanh chóng đến một địa điểm: Nhà ga Grand Central dưới chiếc đồng hồ lớn. Ở Melbourne, đó là những bậc thang trước ga Flinder Street. Ở Toronto, nó có thể ở bảng hiệu Toronto lớn ở Quảng trường Nathan Phillips. Đây là những tiêu điểm. Hầu hết các thành phố và thị trấn đều có một. Cả hai bạn đều biết rằng bạn đang cố gắng đạt được cùng một điều, và cả hai bạn đều biết rằng người kia cũng biết điều đó,...

Theo kinh nghiệm của chúng tôi, những người gặp khó khăn là sinh viên ngoại thành hoặc sinh viên nước ngoài. Họ biết những gì họ cần biết, nhưng họ không có kiến thức chung cần thiết. Họ có thể thực sự thấy dễ dàng hơn khi trả lời câu hỏi tiếp theo của Schelling: “Bạn đã được thông báo về ngày tháng nhưng không phải giờ họp... hai bạn phải đoán chính xác phút gặp nhau trong ngày. Bạn sẽ xuất hiện tại điểm hẹn lúc mấy giờ? Bất cứ ai cũng có thể trả lời điều này: buổi trưa. Lần này thậm chí còn là một tiêu điểm mạnh hơn vị trí.

Sự nghiệp của Thomas Schelling được định hình bởi chiến tranh, cụ thể là Chiến tranh Lạnh. Nghiên cứu của ông liên quan đến cách tránh nó và



cho phép mọi người phối hợp làm một việc gì đó vì lợi ích chung của họ. Các công cụ ông sử dụng là lý thuyết trò chơi. Khi suy nghĩ về cách phối hợp các quyết định của nhiều người khác nhau, ngay cả khi tất cả họ đều hướng tới một nguyên nhân chung, những công cụ này có thể cho biết khi nào thì sự phối hợp đó sẽ khó khăn và khi nào thì dễ dàng.

Điều chúng tôi quan tâm ở đây là cách đưa AI vào để trợ giúp cho một quyết định trong hệ thống sẽ thay đổi bản chất của cách phối hợp tất cả các quyết định (dù được thực hiện với AI hay không). Và câu trả lời xoay quanh độ tin cậy.

Bài tập về tiêu điểm của Schelling cho thấy việc dựa vào những người khác có cùng nền tảng kiến thức sẽ hữu ích như thế nào. Trong các tổ chức, điều này có thể quan trọng, nhưng thường thì sự tin tưởng của chúng ta xuất phát từ việc mong đợi những việc được thực hiện ở nơi khác sẽ rõ ràng phần nào. Khi những người khác tuân theo các quy tắc, nhiệm vụ này thường dễ dàng và chúng ta có thể gắn kết các bộ phận riêng biệt của tổ chức lại với nhau. Nhưng khi các quy tắc thay đổi thành các quyết định, những thách thức trong việc xây dựng một hệ thống trơn tru trở nên rõ ràng. Chúng ta thường không truyền đạt những gì chúng ta đang làm cho người khác mọi lúc bởi vì bản thân điều

đó có chi phí. Thay vào đó, chúng tôi tạo ra những kỳ vọng về những gì họ có thể đang làm và sau đó, khi đưa ra lựa chọn trong lĩnh vực của mình, chúng tôi điều chỉnh hành động của mình phù hợp với những kỳ vọng đó. Nếu những kỳ vọng đó không đáng tin cậy, thì chúng ta sẽ gặp khó khăn trong việc sắp xếp tất cả các quyết định.

Nó không phải là cách đó. Nếu bạn có thể thiết kế hệ thống sao cho các kỳ vọng vẫn đáng tin cậy, thì bạn có thể thu được lợi ích từ việc dự đoán tốt hơn để kích hoạt các cách phục vụ khách hàng mới.

AI BULLWHIP

Hãy tưởng tượng bạn đang điều hành một nhà hàng. Thực khách vào gọi món. Các đầu bếp sau đó làm cho các bữa ăn. Điều này có vẻ khá đơn giản, bởi vì các kỳ vọng được căn chỉnh. Tại bất kỳ thời điểm nào, các đầu bếp bị hạn chế về những món ăn mà họ có thể làm. Những hạn chế này được thúc đẩy bởi kỹ năng của họ, số lượng đơn đặt hàng và sự sẵn có của nguyên liệu và thiết bị. Nếu bạn cho phép khách hàng của mình gọi bất kỳ món ăn nào mà họ có thể thích, thì sẽ có vấn đề. Cái gì bạn làm, do đó, được thiết lập một menu. Bạn giới hạn các lựa chọn của khách hàng để bạn thực sự có thể làm những gì họ yêu cầu. Từ quan điểm của nhà



bếp, thực đơn tự nó tạo ra độ tin cậy và ngăn chặn những bất ngờ không mong muốn.

Mỗi tuần bạn cần đặt hàng nguyên liệu. Các thành phần được dựa trên menu. Nếu trong thực đơn có guacamole, bạn cần bơ. Bạn đặt hàng 100 bảng mỗi tuần. Đôi khi 100 bảng là quá nhiều, và bạn phải vứt bỏ phần dư thừa. Những lần khác, 100 bảng lại không đủ và bạn bỏ lỡ cơ hội bán hàng.

Sau đó, bạn áp dụng AI để dự báo nhu cầu và cách này hiệu quả. Bây giờ một số tuần thì bạn đặt hàng ít nhất là 30 pound. Những tuần khác, bạn đặt 300 bảng. Bạn lãng phí ít hơn và bán được nhiều hơn. Khả năng sinh lời tăng lên.

Nhà cung cấp địa phương của bạn trước đó nhập hàng trăm bảng Anh mỗi tuần theo đơn đặt hàng bạn. Bây giờ họ phải đối mặt với sự khó lường hơn từ bạn. Các khách hàng khác cũng đang sử dụng AI để dự báo nhu cầu. Nhu cầu bắt đầu dao động mạnh.

Vì vậy, nhà cung cấp quyết định áp dụng AI để dự báo nhu cầu của chính họ. Công ty này trước đó thường đặt hàng 25.000 bảng mỗi tuần. Bây giờ đơn đặt hàng của nó thay đổi từ 5.000 bảng Anh đến 50.000 bảng Anh. Ngược lại, nhà cung cấp của nó cần áp dụng AI và các đơn đặt hàng của nó bắt đầu dao động. Và cứ thế, cho đến tận những người

trồng trọt cần đưa ra quyết định về quy mô gieo trồng trước một năm hoặc hơn.

Chúng tôi gọi hiệu ứng này – khi triển khai AI giúp cải thiện chất lượng của một quyết định nhưng lại gây hại cho những quyết định khác trong hệ thống bằng cách hạ thấp độ tin cậy của các quyết định khác – là AI bullwhip. Giống như với chiếc roi, vết nứt nhỏ ở một góc có thể dẫn đến một vết nứt lớn ở nơi khác.

AI có thể được sử dụng để giải quyết sự không chắc chắn, nhưng trừ khi điều đó có thể chuyển thành các quyết định phù hợp hoàn toàn, nếu không thì vấn đề cơ bản – nhu cầu cần phải phù hợp với nguồn cung – vẫn chưa thực sự được giải quyết. Giống như cú vung roi, giải pháp của riêng bạn sẽ có tiếng vang trong tương lai.

Chúng tôi còn lại với một số nghịch lý. Giá trị của AI đến từ việc có thể đưa ra quyết định tốt hơn bằng cách kết hợp những gì bạn làm với các yếu tố có thể không chắc chắn. Tuy nhiên, kết quả là, quyết định của chính bạn trở nên kém tin cậy hơn đối với người khác. Bạn có khả năng vượt qua giới hạn không chắc chắn, điều đó có nghĩa là có thể tốt hơn nếu không điều chỉnh sự không chắc chắn bằng cách sử dụng AI mà gắn bó với một hệ thống đáng tin cậy hơn.



Có hai cách để xây dựng một giải pháp hệ thống AI để xử lý việc này: tăng khả năng phối hợp hoặc tăng tính mô đun.

GIÁ TRỊ CỦA SỰ PHỐI HỢP

Nhà hàng AI dự đoán nhu cầu. Người quản lý nhà hàng đưa ra một số quyết định khác; ví dụ, những gì để cung cấp trên menu. Nếu AI bullwhip có nghĩa là người trông không thể cung cấp đủ bơ, thì nhà hàng cần thay đổi thực đơn. Nó chỉ có thể làm điều này nếu nó biết rằng bơ sẽ không đến. Điều đó đòi hỏi sự phối hợp.

Những hiệp lực như vậy có nghĩa là điều quan trọng là phải xem xét làm thế nào để đạt được sự phối hợp giữa nhiều người ra quyết định, quản lý sự thay đổi và điều chỉnh hơn là một quá trình chuyển đổi và thay đổi.

Để hiểu điều này, hãy xem xét hoạt động của một đội 8 tay chèo. Hai điều xác định cách một đội sẽ thực hiện trong một cuộc đua. Đầu tiên, họ chèo đồng thanh. Thứ hai, họ điều chỉnh tốc độ chèo khi cuộc đua diễn ra để đảm bảo một hoặc nhiều tay chèo không tiêu hao hết sức lực trước khi về đích. Coxswain, người ngồi ở cuối thuyền, cần thiết cho chức năng thứ hai nhưng không phải là chức năng đầu tiên.

Điều này có vẻ đáng ngạc nhiên, vì coxswain đang hô “gạt, chèo, chèo” và phối hợp tất cả các tay chèo để giữ cùng một lúc. Nhưng điều đó không đòi hỏi một người riêng biệt. Một trong những tay chèo có thể làm được điều này, điều này xảy ra trong các cuộc đua không có coxswain. Nhưng khi theo dõi chiến lược trong cuộc đua và thu thập tín hiệu về trạng thái của từng tay chèo – nghĩa là thu thập và tổng hợp thông tin – đồng đội đóng vai trò quan trọng. Coxswain có thể đánh giá sự cần thiết phải thay đổi nhịp điệu và điều chỉnh thông điệp gửi đến người chèo cho phù hợp. Một lần nữa, nếu một tỷ lệ gây duy nhất được sử dụng cho toàn bộ cuộc đua, thì chức năng đó sẽ không cần thiết. Coxswain ở đó chính xác là vì nhóm muốn phản hồi thông tin nhưng cần đảm bảo rằng họ thích ứng theo cách phối hợp.

Đối với các vấn đề đồng bộ hóa như vậy, thiết kế tổ chức được xác định dựa trên nhu cầu về sự phối hợp, vì vậy mô-đun hóa không thể thực hiện được công việc. Kiểu phản ứng phối hợp tương tự đối với thông tin cũng cần thiết cho các vấn đề mà các nhà kinh tế học Paul Milgrom và John Roberts gọi là các bài toán phân công. Đây là những vấn đề mà bạn cần chỉ định tài nguyên cho một hoạt động, nhưng bạn biết rằng chỉ cần sử dụng một số lượng tài nguyên nhất định; nhiều hơn sẽ bị lãng



phí, trong khi ít hơn sẽ không đủ. Hãy xem xét vấn đề gửi xe cứu thương. Nếu có trường hợp khẩn cấp về y tế, một xe cứu thương là rất quan trọng, nhưng hai xe cứu thương thì lãng phí. Để đảm bảo rằng chỉ một phản hồi, bạn cần một người điều phối trung tâm, dù là con người hay phần mềm, người này nhận cuộc gọi (nghĩa là thông tin) liên quan đến trường hợp khẩn cấp và sau đó chỉ định một xe cứu thương để ứng phó. Giả sử, nếu tất cả các xe cứu thương đều nhận được một thông báo khẩn cấp và sau đó chọn riêng lẻ xem có phản hồi hay không, bạn có thể sẽ không có người phản hồi hoặc có quá nhiều người phản hồi. Trong trường hợp này, có một cách phối hợp là tốt nhất, chính xác là vì rủi ro rất cao và việc gửi một xe cứu thương thay vì một xe cứu thương khác đến trường hợp khẩn cấp sẽ ít gặp vấn đề hơn nhiều so với việc không gửi một hoặc gửi quá nhiều xe.

Thay vì chia nhỏ các quyết định và bảo vệ các bộ phận khác của tổ chức khỏi các quyết định do một bộ phận đưa ra để đáp ứng với dự đoán của AI, bạn đang cung cấp tài nguyên và nỗ lực cho một hệ thống liên lạc để đảm bảo rằng các kết quả xấu phát sinh từ một giải pháp điểm – thiếu đồng bộ hóa hoặc phân công tài nguyên kém – không xảy ra. Giờ đây, bạn có nhiều khả năng đưa ra các quyết định quan trọng phản hồi chính xác dự đoán

của AI vì chi phí giảm giá tiềm ẩn được giữ ở mức tối thiểu thông qua thiết kế hệ thống và giao tiếp hiệu quả. Sự kết hợp giữa dự đoán và phối hợp là giải pháp hệ thống. Mỗi quyết định được cải thiện vì nó đáp ứng các dự đoán mà không ảnh hưởng đến độ tin cậy.

GIÁ TRỊ CỦA TÍNH MÔ ĐUN

Tính mô-đun là một cách xây dựng bức tường xung quanh một quyết định do dự đoán của AI đưa ra để tránh các chi phí liên quan đến việc thiếu sự liên kết giữa quyết định đó và những quyết định khác trong tổ chức. Tính mô-đun làm giảm chi phí phối hợp, nhưng nó có thể phải trả giá bằng sự phối hợp có thể đạt được nếu dự đoán AI dẫn đến một quyết định cũng cho phép các quyết định khác di chuyển theo hướng tương tự. Khi không thể phối hợp, tính mô-đun cho phép một số quyết định được hưởng lợi từ AI, đồng thời bảo vệ những quyết định khác khỏi độ tin cậy bị giám sát.

Herbert Simon, người duy nhất đoạt giải Nobel về kinh tế và giải thưởng Turing về khoa học máy tính, từng đưa ra một câu chuyện ngụ ngôn về các tổ chức giải quyết các tình huống phức tạp hơn. Trong truyện ngụ ngôn, hai người thợ đồng hồ sản xuất những sản phẩm chất lượng cao. Cả hai đều có



nhu cầu và liên tục nhận được các yêu cầu từ khách hàng mới. Một bên thịnh vượng trong khi bên kia thất bại. Tại sao?

Đồng hồ bao gồm hàng nghìn bộ phận. Một cách tiếp cận là lắp ráp từng chiếc đồng hồ trong một lần, điều này dẫn đến chiếc đồng hồ có chất lượng cao hơn. Nhưng nếu người thợ đồng hồ bị gián đoạn trong quá trình (ví dụ, bởi một khách hàng khác), họ phải bắt đầu lắp ráp lại từ đầu. Cách tiếp cận khác là lắp ráp đồng hồ thành các bộ phận nhỏ hơn, mỗi bộ phận có khoảng 10 bộ phận nhỏ hơn. Sau đó, những thứ này phải khớp với nhau, điều này tốn nhiều thời gian hơn một chút và có lẽ kết quả cuối cùng không hoàn hảo bằng. Nhưng điều này có lợi thế là, nếu có sự gián đoạn trong quá trình, tất cả những gì bị mất chỉ là một yếu tố nhỏ. Cuối cùng, đó là một quá trình nhanh hơn nhiều và cho phép người thợ đồng hồ tạo ra nhiều đồng hồ hơn. Quy trình thứ hai, được gọi là mô-đun, linh hoạt hơn và có khả năng mở rộng hơn, đặc biệt đối với các sản phẩm phức tạp hơn.

Nếu bạn làm mọi thứ cùng một lúc, bạn phải phối hợp tất cả các quyết định. Một số đặc điểm kỹ thuật sai có thể dẫn đến các vấn đề. Ngược lại, nếu bạn có thể sắp xếp những gì bạn làm thành các

mô-đun, thì các bộ phận riêng lẻ có thể tự làm công việc của mình mà không cần quan tâm đến những gì đang xảy ra ở nơi khác. Điều này không có nghĩa là những gì chúng làm không quan trọng đối với kết quả cuối cùng – nếu một mô-đun không thực hiện công việc của nó, thì toàn bộ sản phẩm có thể bị lỗi. Nhưng nó có nghĩa là một vấn đề lớn hơn trở nên nhỏ hơn và dễ quản lý hơn.

Một ưu điểm khác đến từ khả năng phục hồi của hệ thống trước những thay đổi xảy ra ở một trong các mô-đun: mô-đun có thể cải thiện mà không làm gián đoạn phần còn lại của hệ thống. Đó là, các mô-đun có thể được đổi mới.

Lịch sử có rất nhiều ví dụ trong đó mô-đun hóa giúp đổi mới dễ dàng hơn; ví dụ: khi chúng tôi chuyển từ điện thoại tương tự sang điện thoại kỹ thuật số, điều đó liên quan đến việc thay đổi thiết bị quay số chứ không phải bản thân mạng. Nhưng vào những thời điểm khác, sự đổi mới đã bị hạn chế do thiếu tính mô-đun. Khi máy bay được nâng cấp từ cánh quạt lên động cơ phản lực, các kỹ sư nghĩ rằng khung máy bay có thể được sản xuất như trước đây. Tuy nhiên, các loại rung động mà động cơ mới gây ra quá khác biệt nên toàn bộ cấu trúc của máy bay phải được thiết kế lại. Điều này làm chậm quá trình chuyển đổi.



Tính mô-đun là cơ hội để một nhà hàng áp dụng AI mà không có hiệu ứng bullwhip AI, nhưng đó không phải là điều họ có thể tự quyết định. Nếu một nhà hàng muốn thay đổi thực đơn của mình, các nhà cung cấp của họ cần phải có hệ thống mô-đun riêng để xử lý sự thay đổi theo nhu cầu đó. Trong ví dụ của chúng tôi, việc cung cấp bị hạn chế bởi nhu cầu thay đổi trong toàn ngành. Với đủ nhà hàng trên nhiều khu vực được cung cấp, ngay cả khi nhu cầu nhà hàng cá nhân rất thay đổi, tổng thể có thể ổn định hơn. Quy mô có thể tạo cơ hội cho mô-đun hóa trong chuỗi cung ứng. Nói chung, đối với việc áp dụng AI, tính mô-đun có thể giúp giải quyết những khó khăn phát sinh do các quyết định có liên quan đến nhau.

GIÁ TRỊ CỦA THIẾT KẾ

Việc áp dụng AI cho một quyết định sẽ dễ dàng hơn nhiều nếu các quyết định được đưa ra không nhất thiết phải phù hợp với các quyết định ở những nơi khác trong hệ thống. Đây là một vấn đề của mức độ. Tất nhiên, về mặt khái niệm, sẽ tốt hơn nếu toàn bộ hệ thống có thể chuyển động như một. Câu hỏi đặt ra là, nếu điều đó không thể xảy ra, liệu có thể gặt hái được những lợi ích của AI hơn bất kỳ chi phí nào có thể được cảm nhận ở nơi khác hay không.

Để hiểu điều này, hãy xem xét hoạt động của Amazon, nơi cung cấp hàng triệu sản phẩm trên toàn thế giới. Amazon mua chúng, lưu trữ chúng trong kho, nắm bắt đơn đặt hàng của khách hàng và giao hàng cho những khách hàng đó. Nhưng nó cũng liên quan đến việc giúp khách hàng tìm ra thứ cần mua ngay từ đầu, nghĩa là cung cấp cho họ các đề xuất.

Về mặt khái niệm, Amazon phải đối mặt với vấn đề nhà hàng của chúng tôi. Nó muốn cung cấp cho khách hàng những gì họ muốn khi họ muốn, nhưng các sản phẩm không xuất hiện một cách kỳ diệu. Một chuỗi cung ứng kéo dài hàng nghìn kilomet và hàng tháng. Vì vậy, nếu nó giới thiệu một sản phẩm cho khách hàng không có sẵn, thì sao?

Thật hấp dẫn khi nghĩ rằng giải pháp là: đừng làm vậy. Nếu bạn không có sẵn sản phẩm, đừng giới thiệu nó cho khách hàng. Nhưng có một vấn đề với cách tiếp cận đó: Làm thế nào để bạn biết liệu có một sản phẩm không có sẵn mà khách hàng thực sự muốn hay không? Nếu chỉ giới thiệu những gì bạn có, bạn sẽ bỏ lỡ cơ hội phát triển và mở rộng.

Đây là lý do tại sao Amazon khuyến nghị các sản phẩm đã hết hàng hoặc có thể mất nhiều thời gian hơn để đến tay khách hàng. Các quyết định



được điều phối theo nghĩa là Amazon thông báo khả năng xảy ra chậm trễ cho khách hàng. Khách hàng có thể chọn sản phẩm có sẵn, nhưng đôi khi họ sẽ không chọn. Sau đó, Amazon biết được họ cần nỗ lực bao nhiêu trong việc kiểm kê hàng tồn kho cho những mặt hàng đó.

Đạt được sự cân bằng này đòi hỏi thiết kế cẩn thận. Amazon có một tổ chức mô-đun cho phép nó đưa dự đoán AI tốt hơn vào các khuyến nghị giảm thiểu tác động đến phần còn lại của tổ chức. Nhưng đưa đến một thái cực, đó sẽ là một bước quá xa. Do đó, các lựa chọn về hàng tồn kho và đặt hàng mà nó đưa ra không thể hoàn toàn độc lập với cách hoạt động của đề xuất AI, chính xác là do các lựa chọn và phản ứng của khách hàng làm phát sinh thông tin mà bộ phận hậu cần cần giao tiếp và hành động.

Việc áp dụng AI thường sẽ liên quan đến một giải pháp hệ thống tìm thấy sự cân bằng tối ưu giữa tính mô-đun và sự phối hợp. Tính mô-đun bảo vệ các quyết định khỏi sự thay đổi đến từ AI. Nó làm giảm tầm quan trọng của độ tin cậy. Sự phối hợp, ngược lại, trực tiếp tạo ra độ tin cậy. Các hệ thống AI thành công cho phép phối hợp khi có thể và mô-đun hóa khi cần thiết.

HỆ THỐNG THUYỀN BUỒM

Các nhà sản xuất thuyền buồm và thủy thủ đã dày công xây dựng và phát triển ngành nghề này trong 5.000 năm. Những đổi mới vẫn chưa dừng lại, ngay cả khi vận chuyển thương mại đã không còn dựa vào gió để vận hành. Người chiến thắng trong cuộc đua America's Cup khẳng định chèo thuyền là giải thưởng danh giá và chiếc cúp lâu đời nhất trong thể thao quốc tế. Tiêu chí của cuộc thi là về cả công nghệ được phát triển cho thuyền và kỹ năng của các thủy thủ.

Hàng triệu đô la được đầu tư vào thiết kế thuyền. Am hiểu về tính chất vật lý của gió, nước và tàu, các đối thủ cạnh tranh sử dụng trình mô phỏng để xác định các thiết kế hiệu quả nhất từ lâu trước khi bắt đầu xây dựng. Trình mô phỏng cho phép các thủy thủ lái thử một chiếc thuyền mà không cần đóng nó. Đội có trình mô phỏng tốt nhất sẽ giành được lợi thế. Đội New Zealand đã sử dụng trình giả lập của mình để giành chiến thắng vào năm 2017.

Khi nhóm lên kế hoạch cho cuộc đua năm 2021, họ tự hỏi liệu họ có thể đẩy nhanh quá trình thiết kế hay không. Sau khi hợp tác với công ty tư vấn toàn cầu McKinsey, họ đã xác định được nút thắt chính để tạo nên sự đột phá đổi mới: bản thân các thủy thủ. Khi con người chèo thuyền trong trình giả lập,



sẽ mất thời gian. Không có cách nào để tăng tốc độ phản ứng của con người với các điều kiện và phản ứng của con thuyền.

Sử dụng công nghệ tương tự như AI đã đánh bại những kỳ thủ cờ vây hàng đầu thế giới, nhóm đã dạy một cỗ máy chèo thuyền. Họ không phải quan tâm liệu các thủy thủ có thời gian. Bot không cần ngủ hay ăn và nó có thể chạy hàng trăm mô phỏng trong khoảng thời gian mà các thủy thủ con người chỉ chạy được một lượng nhất định. Sau 8 tuần, AI bắt đầu đánh bại các thủy thủ trong trình giả lập.

Đây là khi mọi thứ bắt đầu trở nên thú vị. Các thủy thủ AI bắt đầu dạy các thủy thủ con người những mảnh khöhe mới. Trước đây, điểm đổi mới trong thiết kế thuyền nằm ở tốc độ của con người. Nếu có một quá trình học tập để tìm ra cách tốt nhất để sử dụng một chiếc thuyền được thiết kế mới, thì quá trình đó diễn ra trong nhiều giờ, nhiều ngày hoặc nhiều tuần khi các thủy thủ thử nghiệm những điều khác nhau và học hỏi.

Ngược lại, AI có thể thử nghiệm các biến thể khác nhau của con thuyền. Nó có thể thử các chiến thuật đua xe khác nhau. Nó đẩy nhanh chu kỳ lặp lại thiết kế và cho phép phát triển các thao tác mới dành riêng cho thiết kế mới. Sau đó, khi AI đã tìm

ra giải pháp ưu việt, các thủy thủ con người có thể sao chép AI và học các thủ thuật mới từ các thủy thủ AI để chèo thuyền mô phỏng. Như một thành viên trong nhóm đã nói, “Việc đẩy nhanh quá trình học hỏi là vô cùng quý giá, cả về mặt cho phép nhóm thiết kế khám phá càng nhiều không gian thiết kế càng tốt và các thủy thủ tối đa hóa hiệu suất cho một thiết kế nhất định”. Cuối cùng, Đội New Zealand giành được cúp.

Trong ví dụ về giải pháp hệ thống AI này, AI đã dẫn đến những thay đổi trong nhiều quyết định. Cụ thể, việc chuẩn bị cho cuộc đua bao gồm hai loại quyết định. Có quyết định về thiết kế thuyền và quyết định về điều động thuyền. Mô phỏng từ lâu đã được sử dụng để thiết kế thuyền. Con người đã luôn thực hiện các thao tác chèo thuyền. Thủy thủ AI không thực sự lái thuyền trong cuộc đua – người thật vẫn lái thuyền thật. Thay vào đó, AI đã đẩy nhanh quá trình đổi mới, cho phép phối hợp tốt hơn giữa thiết kế thuyền và thao tác chèo thuyền. Hệ thống hoàn chỉnh của thuyền mô phỏng và thủy thủ AI đã kích hoạt các cải tiến cho cả hai.

•

HỆ THỐNG SONG SINH

Trình mô phỏng chèo thuyền là một ví dụ về “song sinh số” hoặc biểu diễn ảo của một đối tượng



hoặc hệ thống vật lý, cung cấp thông tin thay thế cho tài nguyên vật lý. Với các cảm biến phù hợp tại chỗ, chúng cho phép theo dõi thời gian thực và bảo trì dự đoán. Những đại diện ảo này có thể làm được nhiều hơn thế. Chúng cung cấp một khuôn khổ cho mô phỏng cấp hệ thống. Accenture gọi đây là “sân chơi đổi mới không rủi ro”. Như Michael Grieves, giám đốc điều hành của Digital Twin Institute, nói rằng, “Các hệ thống không bùng nổ khi được hình thành đầy đủ. Chúng tiến triển thông qua một vòng đời sáng tạo, sản xuất, vận hành và đào thải. Với các hệ thống ‘đơn thuần vật lý’, đây là một tiến trình tuyến tính. Digital Twin cho phép phát triển đồng thời, lặp đi lặp lại nhiều hơn”.

Khi kết hợp với AI, điều này tạo ra cơ hội để thiết kế một cách thức làm việc mới. Như Đội New Zealand đã phát hiện ra trong quá trình chuẩn bị cho America's Cup 2021, các hệ thống mô phỏng cho phép các đội tìm ra cách tốt nhất để phối hợp đưa ra các quyết định, chẳng hạn như thiết kế thuyền và điều khiển chèo thuyền. Bằng cách này, họ có thể giảm thử và sai. Khi người quản lý phát triển ý tưởng về cách thay đổi hệ thống, có thể mô phỏng tác động của ý tưởng đó mà không tốn chi phí chế tạo máy móc hoặc rủi ro ngừng hoạt động.

Các mô phỏng cũng có thể tập trung vào việc triển khai AI. Nếu một máy dự đoán được thêm

vào một phần của hệ thống, mô phỏng có thể giúp xác định những quyết định nào khác cần được phối hợp hoặc cách hệ thống có thể được tạo thành mô-đun hơn.

Các hệ thống rất phức tạp vì chúng là sự kết hợp của các quyết định tương tác với nhau. Hãy tưởng tượng một hệ thống chỉ có một quyết định nhị phân: nói lỏng cánh buồm (L) hoặc thắt chặt cánh buồm (T). Chỉ có hai lựa chọn [L, T]. Bây giờ, hãy tưởng tượng rằng quyết định thứ hai ảnh hưởng đến quyết định đầu tiên: đứng thẳng (S) hoặc nghiêng sang phải (R). Hiện có 4 tùy chọn (LS, LR, TS, TR). Bây giờ hãy tưởng tượng rằng có một quyết định thứ ba về việc thêm một cánh buồm khác (A) hay không (N) phụ thuộc vào hai quyết định đầu tiên. Bây giờ có 8 lựa chọn (LSA, LSN, LRA, LRN, TSA, TSN, TRA, TRN).

Trong trường hợp đầu tiên, chúng ta có $2^1 = 2$ lựa chọn; trong trường hợp thứ hai, ta có $2^2 = 4$ lựa chọn; và trong trường hợp thứ ba, chúng ta có $2^3 = 8$ lựa chọn. Vào thời điểm chúng tôi có 10 quyết định tương tác, chúng tôi có 1.024 lựa chọn và 20 quyết định tương tác tạo ra 1.048.576 lựa chọn. Đua thuyền buồm có thể liên quan đến hàng trăm quyết định cần phối hợp và sẽ không mất nhiều thời gian để có nhiều lựa chọn hơn số nguyên tử trong vũ trụ quan sát được.



Vấn đề ở đây là các hệ thống có các quyết định phụ thuộc lẫn nhau có thể trở nên rất phức tạp và rất nhanh chóng. Cái nhìn sâu sắc quan trọng đó đặt nền tảng cho lý do tại sao mô phỏng lại có tác dụng mạnh mẽ như vậy đối với thiết kế hệ thống. Nó tận dụng cùng một hiểu biết sâu sắc mà chúng ta đã thảo luận trong Chương 3 về lý do tại sao AI lại thành công như vậy khi chơi trò chơi: việc mô phỏng dữ liệu mới tương đối dễ dàng. Mặc dù rất tốn kém – trong một số trường hợp, về mặt vật lý là không thể – để thử nghiệm tất cả các tùy chọn để tìm ra cái tốt nhất, chúng ta có thể sử dụng các tài sản kỹ thuật số, bao gồm cả cặp song sinh kỹ thuật số của môi trường vật lý, để mô phỏng các tùy chọn khác nhau và sử dụng AI để dự đoán kết quả từ mỗi tùy chọn.

Vì vậy, chúng ta có thể khám phá nhiều lựa chọn hơn so với những gì chúng ta có thể làm trong thế giới vật chất và tăng cơ hội tìm thấy sự kết hợp tốt hơn so với sự kết hợp mà chúng ta có khả năng giải quyết khi không có mô phỏng.

Virtual Singapore là mô phỏng địa hình, vùng nước, thảm thực vật, cơ sở hạ tầng giao thông và các tòa nhà cho đến vật liệu xây dựng trong thành phố. Bản sao kỹ thuật số này là một công cụ cho

phép các nhà quản lý mô phỏng giải pháp hệ thống AI và tránh được một số lỗi tốn kém. Nó tốn hàng chục triệu đô la để phát triển. Mô hình này cho phép các nhà quy hoạch đánh giá tác động của các công viên hoặc tòa nhà mới đối với giao thông và đám đông, đồng thời khám phá phạm vi phủ sóng của mạng di động. Nó cũng có thể được sử dụng để đánh giá cách thêm dự đoán máy sẽ ảnh hưởng đến cuộc sống của công dân và cư dân Singapore. Ví dụ, AI có thể cho phép tối ưu hóa giao thông công cộng tốt hơn. Mô hình có thể đánh giá xem việc tối ưu hóa đó có yêu cầu thay đổi bổ sung đối với quản lý lưu lượng hay không. Nói cách khác, nó có thể đánh giá liệu các hệ thống giao thông công cộng và tư nhân có thể được coi là mô-đun hay cần được phối hợp với nhau. Sau đó, nó có thể phát triển một hệ thống giao thông hỗ trợ AI tốt hơn, bằng cách sử dụng tính mô-đun hoặc bằng cách đảm bảo sự phối hợp khi cần.

Công ty Xây dựng & Công nghiệp nặng Doosan của Hàn Quốc và Microsoft đã phát triển một trang trại gió song sinh kỹ thuật số để minh họa cho những lợi ích khác nhau của các hệ thống mô phỏng. Mô phỏng kết hợp các mô hình dựa trên vật lý và học máy để dự đoán sản lượng cho từng tuabin trong trang trại.



So sánh sản lượng dự đoán và sản lượng thực tế của tuabin cho phép người vận hành tinh chỉnh các điều khiển và tối ưu hóa sản xuất. Hơn nữa, bộ đôi kích hoạt sự đổi mới trong thiết kế và phát triển trang trại gió tổng thể, cải thiện độ tin cậy. Cuối cùng, nó cho phép phối hợp các quyết định. Các dự đoán chính xác về sản lượng năng lượng cho phép Doosan gia tăng các cam kết về sản lượng đối với các nhà vận hành lưới điện, đồng thời tránh bị phạt do không đáp ứng các cam kết. Nó làm giảm rủi ro khi xây dựng giải pháp hệ thống và gắn bó với giải pháp điểm ít giá trị hơn nhưng đơn giản hơn. Ví dụ, khả năng dự đoán tốt hơn giúp đưa ra quyết định tua-bin nào sẽ hoạt động và tua-bin nào cần bảo trì. Điều này lần lượt kích hoạt các quyết định về cam kết với lưới năng lượng.

Các hệ thống cho AI

Mô phỏng không phải là cách duy nhất để xây dựng một hệ thống do AI điều khiển, nhưng nó cho thấy cơ hội. Bằng cách tìm ra cách phối hợp phù hợp với những người ra quyết định, Đội New Zealand đã tìm ra con đường dẫn đến chiến thắng. Trí tuệ nhân tạo có thể được áp dụng khi một hệ thống ở dạng mô-đun, nhưng tác động của nó có thể lớn hơn nhiều nếu có thể phối hợp. Thách thức là tìm ra loại phối hợp nào là cần thiết.

NHỮNG ĐIỂM CHÍNH

- Các quyết định không hoạt động trong môi trường chân không. Thông thường, nhiều quyết định hoặc hành động khác bị ảnh hưởng bởi kết quả của một quyết định. Đó là lý do tại sao đôi khi chúng tôi sử dụng các quyết định (quy tắc) được xác định trước thay vì các quyết định theo thời gian thực vì các quy tắc nâng cao độ tin cậy, vì vậy chúng tôi chấp nhận các quyết định cục bộ tồi tệ hơn để đổi lấy độ tin cậy cao hơn vì lợi ích của toàn bộ hệ thống. Độ tin cậy là một tính năng chính của các hệ thống với các quyết định phụ thuộc lẫn nhau.
- Có hai phương pháp thiết kế hệ thống chính để giải quyết vấn đề giảm độ tin cậy do đưa ra quyết định dựa trên AI: (1) phối hợp và (2) mô-đun hóa. Phối hợp liên quan đến việc xác định mục tiêu tổng thể và sau đó thiết kế các luồng thông tin, khuyến khích và quyền quyết định để mỗi người ra quyết định trong hệ thống có thông tin và khuyến khích để tối ưu hóa cho mục tiêu tổng thể. Tính mô-đun liên quan đến việc xây dựng một bức tường xung quanh một quyết định AI được nâng cao để tránh các chi phí liên quan đến việc thiếu sự liên kết giữa quyết



định đó và những quyết định khác trong tổ chức. Tính mô-đun làm giảm chi phí phối hợp, nhưng phải trả giá bằng sự phối hợp.

- Hệ thống là sự kết hợp của các quyết định tương tác với nhau. Xem xét một tập hợp các quyết định nhị phân liên quan. 3 quyết định dẫn đến 8 kết hợp khác nhau. 10 quyết định tương tác tạo ra 1.024 kết hợp và 20 quyết định tương tác tạo ra 1.048.576 kết hợp. Các hệ thống có các quyết định phụ thuộc lẫn nhau có thể trở nên rất phức tạp rất nhanh. Đó là lý do tại sao mô phỏng rất hữu ích cho thiết kế hệ thống. Chúng ta có thể sử dụng cặp song sinh kỹ thuật số để mô phỏng các kết hợp khác nhau và sử dụng AI để dự đoán kết quả từ mỗi kết hợp.

Chương 17

TỜ GIẤY TRẮNG



Hãy tưởng tượng đến thăm bác sĩ của bạn để kiểm tra y tế. Cuối cùng, bác sĩ của bạn nói: “Có khả năng cao là bạn sẽ bị bệnh nặng trong vòng 3 năm. Cảm ơn vì đã đến”. Sau đó, bác sĩ của bạn bước ra ngoài và bắt đầu buổi khám với bệnh nhân tiếp theo của cô ấy. Bạn sẽ sững sốt. Tại sao cô ấy không nói với bạn những gì sẽ làm cho bạn bị bệnh? Tại sao cô ấy không giải thích bạn có thể làm gì để giảm nguy cơ mắc bệnh?

Mặc dù câu chuyện này có vẻ khó tin nhưng nó lại xảy ra hằng ngày trong ngành bảo hiểm. Các công ty bảo hiểm tính phí bảo hiểm cho một số người cao hơn những người khác. Tại sao? Bởi vì họ dự đoán rằng một số khách hàng có nhiều khả năng bị thua lỗ hơn những khách hàng khác. Làm thế nào để các công ty bảo hiểm biết ai có nguy cơ cao hơn? Bởi vì họ đầu tư rất nhiều vào việc thu thập và phân tích dữ liệu để dự đoán khả năng khách hàng bị thua lỗ và nộp đơn yêu cầu bồi thường.



Không có gì đáng ngạc nhiên khi các công ty bảo hiểm đang dẫn đầu về khoa học dữ liệu. Đó là công việc của họ để đưa ra dự đoán. Tuy nhiên, điều đáng ngạc nhiên là họ không chia sẻ những hiểu biết về rủi ro với khách hàng của mình. Thông tin có giá trị này có thể giúp khách hàng giảm thiểu rủi ro, thay vì chỉ bảo đảm cho họ trước rủi ro đó.

Ví dụ, các công ty bảo hiểm nhà ở đang áp dụng AI để tạo ra các dự đoán có độ chính xác cao hơn. Nhiều người hiện có thể dự đoán rủi ro ở mức nguy hiểm hoặc dưới mức nguy hiểm (ví dụ: nguy cơ chập điện do hệ thống dây điện kém, nguy cơ lũ lụt do đường ống bị rò rỉ). Vì vậy, nếu một công ty bảo hiểm nhà dự đoán rằng một chủ nhà cụ thể có nguy cơ đặc biệt cao về hỏa hoạn hoặc lũ lụt, thay vì chỉ tính phí bảo hiểm cao cho họ do khả năng họ nộp đơn yêu cầu bồi thường cao, thì công ty có thể chia sẻ thông tin này để khách hàng có thể hành động nếu họ muốn giảm rủi ro. Ví dụ, những khách hàng như vậy có thể đầu tư vào các thiết bị giá rẻ để phát hiện sớm nguy cơ hỏa hoạn hoặc lũ lụt tăng cao. Công ty bảo hiểm thậm chí có thể quyết định trợ cấp cho các công cụ giảm thiểu rủi ro này, bởi vì việc giảm tổn thất dự kiến có thể lớn hơn chi phí của các thiết bị.

Đáng ngạc nhiên là rất ít công ty bảo hiểm đã đi theo con đường này với bất kỳ quy mô đáng kể nào. Thay vào đó, hầu hết đang tập trung nỗ lực vào việc xây dựng và triển khai AI nhằm nâng cao các dự đoán bảo lãnh phát hành truyền thống của họ. Họ đang xây dựng các giải pháp điểm. Tại sao hầu hết các công ty bảo hiểm không theo đuổi các cơ hội để phục vụ khách hàng của họ tốt hơn bằng cách chuyển từ mô hình kinh doanh dựa trên việc tập hợp rủi ro giữa các khách hàng và chuyển rủi ro từ khách hàng sang hãng vận chuyển sang giảm thiểu rủi ro cho khách hàng của họ? Các đại lý có thể không thích điều này vì rủi ro thấp hơn có nghĩa là phí bảo hiểm thấp hơn, điều này có thể dẫn đến thu nhập thấp hơn. Tuy nhiên, nhìn chung, điều này có vẻ như sẽ tạo ra giá trị đáng kể cho khách hàng.

Trong nhiều trường hợp, có vẻ như các công ty bảo hiểm đã không đánh giá hết cơ hội này vì nó nằm ngoài mô hình kinh doanh thông thường của họ. Ngoài các khuyến khích của đại lý, có rất nhiều quy tắc kinh doanh, quy định của chính phủ và phương pháp thực hiện trong ngành này mà những gì có vẻ hiển nhiên từ quan điểm thuận lợi của người ngoài cuộc khó có thể nhìn thấy từ quan điểm của người trong cuộc. Đó là lý do tại sao chúng tôi đề xuất phương pháp phương tiện



chặn trống. Để tạo điều kiện thuận lợi cho việc này, chúng tôi giới thiệu Canvas khám phá hệ thống AI.

SUY NGHĨ NHƯ MỘT NHÀ KINH TẾ

Một trong những kỹ năng của chúng ta với tư cách là các nhà kinh tế học là lấy một thứ gì đó thú vị và không thể hiểu được rồi phân tích nó thành một thứ gì đó nhằm chán và dễ hiểu. Mặc dù điều đó không khiến chúng ta trở thành những vị khách tuyệt vời của bữa tiệc, nhưng đôi khi nó cho phép chúng ta nhìn thấy những điều mà người khác bỏ lỡ. Chúng tôi cũng đã thiết kế một khuôn khổ để giúp bạn làm điều này. Khung khám phá hệ thống AI rất hữu ích nếu bạn muốn phát triển tư duy hệ thống để đánh giá giá trị của AI.

Trong chương này, chúng tôi cung cấp cho bạn một công cụ mà bạn có thể sử dụng để xây dựng trên một phương tiện trống. Ý tưởng là xác định các quyết định quan trọng tối thiểu mà bạn cần đưa ra trong ngành của mình, nếu bạn có các máy dự đoán có độ chính xác cao, để đạt được sứ mệnh của mình. Dự đoán AI, về bản chất, sẽ xâm nhập vào bất kỳ tổ chức nào ở cấp độ quyết định. Nhưng biết làm thế nào một quyết định hoặc loại quyết định tác động đến người khác là bước quan trọng đầu tiên trong việc phát triển quan điểm về cách AI có thể có tác

động trên toàn hệ thống. Có hai lý do tại sao bài tập này có giá trị khi bạn đang cố gắng đánh giá việc áp dụng khả năng dự đoán của AI có thể dẫn đến sự gián đoạn như thế nào và cách xem xét liệu bạn có cần đổi mới ở cấp độ hệ thống hay không. Đầu tiên là có thể có nhiều quy tắc trong một tổ chức và khả năng một số chức năng đã được xây dựng để che giấu sự không chắc chắn liên quan đến các quy tắc đó. Phương tiện trống yêu cầu bạn quay lại các nguyên tắc đầu tiên và xem xét các quyết định liên quan đến việc hoàn thành sứ mệnh của tổ chức bạn. Trong quá trình này, một số quyết định đó có thể đã tồn tại dưới dạng quy tắc và một số có thể mang đến cơ hội áp dụng các dự đoán sẽ biến các quy tắc đó thành quyết định. (Tuy nhiên, chúng tôi sẽ trì hoãn việc sử dụng phương tiện chặn trống cho loại mục đích này cho đến chương tiếp theo.)

Một lý do khác khiến bài tập này có giá trị là bạn có thể sử dụng nó để đánh giá tác động mà các giải pháp AI cụ thể có thể có đối với các hệ thống. Bằng cách sử dụng một phương tiện chặn trống, bạn có thể có góc nhìn rộng hơn về cách quyết định được thông báo bởi các dự đoán AI có thể tương tác với các quyết định hoặc quy tắc khác trong tổ chức của bạn. Trong chương này, chúng tôi sẽ chứng minh tính hữu ích của việc bắt đầu với một phương



tiện trống để đánh giá ý nghĩa hệ thống của các giải pháp AI cụ thể.

Ví dụ, trong lĩnh vực bảo hiểm, một số doanh nhân đã phát triển các ứng dụng theo đó người dùng chụp ảnh một chiếc ô tô hoặc ngôi nhà sau khi bị hư hỏng và công ty sẽ tự động sử dụng những bức ảnh đó để tính toán tài khoản yêu cầu bồi thường và thanh toán ngay cho việc sửa chữa. Người tiêu dùng không phải đợi chuyên gia đánh giá hoặc lái xe vòng quanh để tìm báo giá. Chỉ cần mở ứng dụng, chụp một vài bức ảnh, và bạn đã hoàn tất. Một bộ ứng dụng khác có trong các thiết bị giám sát việc lái xe hoặc ngôi nhà của bạn. Các ứng dụng đó có thể nhanh chóng xác định xem bạn có đang làm điều gì đó rủi ro hay không – không chỉ đơn giản là yêu cầu bạn dừng lại mà còn cảnh báo bạn rằng nếu hành vi đó vẫn tiếp diễn, phí bảo hiểm của bạn có thể khác vào tháng tới hoặc năm tới.

Thật dễ hiểu tại sao các doanh nhân có thể nhắm mục tiêu đến các công ty bảo hiểm với các giải pháp ứng dụng này. Nhưng liệu các giải pháp đó có hữu ích không? Để giải quyết vấn đề đó, bạn cần hiểu rõ bản chất của một ngành cụ thể, chẳng hạn như ngành bảo hiểm ô tô hoặc nhà ở. Điều đó có nghĩa là tìm ra những quyết định mà ngành cần đưa ra và liệu giải pháp AI cụ thể này có cung cấp thông tin cho các quyết định được đưa ra hay không. Sau đó,

bạn có thể vạch ra một kế hoạch hành động cụ thể. Liệu có ai đã đang chịu trách nhiệm về quyết định đó chưa? Hay liệu đã có người ra quyết định rõ ràng chưa? Hay đang phụ thuộc vào các quy tắc? Nếu bạn muốn thay đổi quy tắc đó thành quyết định, điều gì khác có khả năng tác động? Để đặt ra những câu hỏi quan trọng đó, bạn cần một điểm khởi đầu. Tiếp theo chúng ta sẽ thảo luận về vấn đề này.

BIỂU ĐỒ KHÁM PHÁ HỆ THỐNG AI

Theo thời gian, chúng tôi nhận ra rằng, những người thực hiện những thay đổi thường thích trình bày các phương án của họ trên một khung vẽ, được gọi là canvas. Canvas là một biểu đồ trống, trên đó bạn có thể bắt đầu ở bất kỳ đâu, miễn là khi kết thúc quy trình, bạn đã suy nghĩ thấu đáo về toàn bộ biểu đồ. Nó không phải là sổ tay hướng dẫn từng bước mà là cách sắp xếp suy nghĩ của bạn.

Hình 17-1 là hình một biểu đồ cho phép bạn vạch ra các quyết định quan trọng trong một ngành. Nhiệm vụ quan trọng là xác định sứ mệnh của doanh nghiệp bạn. Đây không cần phải là một tuyên bố chính xác mà thay vào đó là một lời nhắc nhở rộng rãi về các mục tiêu của bạn.

Trọng tâm là xác định các quyết định cần thiết để đạt được sứ mệnh đó. Rõ ràng, có thể có nhiều quyết định (về mặt lý thuyết có thể lên tới hàng triệu).



Mục đích ở đây không phải là xác định được tất cả các quyết định. Thay vào đó, mục đích là nêu ra các loại quyết định cần thiết. Nếu bạn có những cỗ máy dự đoán rất mạnh mẽ để nâng cao khả năng ra quyết định của mình, thì số lượng quyết định nhỏ nhất cần thiết để đạt được sứ mệnh của bạn là bao nhiêu? Chỉ xác định những quyết định quan trọng nhất hoặc cốt lõi.

Sau khi xác định các quyết định, đã đến lúc đi sâu vào. Bạn cần thu thập thông tin gì để đưa ra quyết định? Đây không chỉ đơn giản là thông tin bạn có hoặc có thể dễ dàng có được mà thay vào đó, thông tin bạn có thể tưởng tượng là quan trọng. Hầu hết các quyết định được đưa ra trong điều kiện không chắc chắn. Tuy nhiên, với một dự đoán, bạn có thể có thông tin cần thiết để đưa ra quyết định tốt hơn. Các dự đoán là những gì AI có thể cung cấp, vì vậy bài tập này liên kết các dự đoán với các quyết định cốt lõi trong tổ chức của bạn.

Hình 17-1: Khám phá hệ thống AI

Bạn có thể thu hẹp doanh nghiệp của mình về những quyết định?

1. Nhiệm vụ			
2. Quyết định			
3. Dự đoán			
4. Đánh giá			

Cuối cùng, không có dự đoán nào là hoàn hảo. Nếu bạn có một dự đoán hoàn hảo, việc ra quyết định sẽ dễ dàng và có khả năng tự động. Canvas không hẳn là phi thực tế. Do đó, đối với mỗi quyết định, bạn cần nói rõ những đánh đổi chính có liên quan. Thật vậy, chúng tôi ủng hộ việc sử dụng “khung lỗi”. Nếu dự đoán của tôi sai hoặc không tồn tại, chúng ta có thể mắc phải những loại sai lầm nào? Điều này cho bạn cảm giác về rủi ro của quyết định. Trong phần lựa chọn ở Chương 3, nếu dự báo thời tiết của bạn sai, bạn có thể mang theo ô một cách không cần thiết hoặc bạn bị ướt. Đánh giá của bạn là cách bạn xếp hạng những lỗi đó. Đối với mỗi quyết định trên khung vẽ của mình, bạn muốn xác định hậu quả của các sai sót và khả năng bạn sẽ xếp hạng chúng như thế nào – chẳng hạn như tính toán chi phí rõ ràng hoặc điều gì đó chủ quan hơn.

Tiếp theo, bạn có thể thực hiện bất kỳ dự đoán tiềm năng nào do AI điều khiển và liên kết nó với một quyết định để đánh giá (a) liệu tổ chức của bạn có thực hiện các quyết định đó một cách rõ ràng hay không; (b) người hiện đang sở hữu quyết định; và (c) nếu bạn sử dụng AI để đưa ra quyết định đó, điều đó có thể gây ra sự gián đoạn gì cho phần còn lại của tổ chức hiện tại của bạn? (Chúng ta xem xét những bước cuối cùng này trong Chương 18 và phần kết.)



Hiện tại, bạn muốn có điểm bắt đầu – bảng trống – và hệ thống làm nền tảng cho ngành của bạn.

NGÀNH BẢO HIỂM

Theo nhiều cách, không có ngành nào ổn định hơn bảo hiểm. Bắt đầu từ nhiều thế kỷ trước, nó đã phát triển thành một yếu tố chính của cuộc sống hiện đại. Sản phẩm bảo hiểm tiêu dùng rất đơn giản. Mọi người trả phí bảo hiểm hằng năm và đổi lại, nhận được khoản thanh toán nếu họ bị tai nạn xe hơi, hư hỏng nhà cửa hoặc trộm cắp hoặc họ chết. Cuộc cách mạng công nghệ thông tin mang lại một số tiến bộ. Việc tính toán các bảng tính toán bảo hiểm trở nên dễ dàng hơn, sau đó có thể sửa đổi các bảng này để cung cấp nhiều loại sản phẩm bảo hiểm hơn. Nhưng cuối cùng, sự khác biệt chính trong các sản phẩm đó đến từ những thứ dường như nằm ngoài tầm kiểm soát của khách hàng, chẳng hạn như tuổi tác hoặc nơi họ sinh sống.

Cần những gì để cung cấp các sản phẩm bảo hiểm? Hãy xem xét bảo hiểm của chủ nhà. Sự mệnh của một doanh nghiệp trong lĩnh vực đó có thể được phát biểu như sau: “Mang đến cho chủ nhà sự an tâm chống lại sự mất mát thảm khốc đối với nhiều tài sản quý giá nhất mà họ sở hữu”. Bạn sẽ

đặt cái này vào ô trên cùng của canvas (xem Hình 17-2). Trong hình, chúng ta đã xác định được 3 loại quyết định – tiếp thị, bảo lãnh phát hành và yêu cầu bồi thường. Đây thường là những bộ phận trong các công ty bảo hiểm, khiến ngành này trở thành một ngành khá đơn giản để phân tích.

Tiếp thị chịu trách nhiệm thu hút khách hàng: tìm kiếm những người có thể sử dụng bảo hiểm và bán sản phẩm cho họ. Quyết định của các nhà tiếp thị xoay quanh việc đặt các nguồn lực ở đâu để nhắm đến khách hàng. Bảo lãnh phát hành tạo ra các sản phẩm bảo hiểm và đánh giá hồ sơ rủi ro của khách hàng để xác định phí bảo hiểm cũng như liệu một khách hàng có trở thành khách hàng được bảo hiểm hay không. Nói cách khác, các nhà bảo hiểm chịu trách nhiệm định giá các sản phẩm bảo hiểm, có nghĩa là họ hiểu chi phí bảo hiểm cho các khách hàng hoặc nhóm khách hàng cụ thể với các đặc điểm nhất định. Cuối cùng, bộ phận yêu cầu bồi thường xác định có nên thanh toán yêu cầu bồi thường hay không. Trên thực tế, nó xử lý những gì trong các doanh nghiệp khác sẽ được gọi là trải nghiệm của khách hàng: thật thú vị biết bao khi trao lợi ích cho khách hàng, nhưng có lẽ điều này cũng được thực hiện với mục đích không cung cấp những lợi ích đó nếu có thể.



Hình 17-2 Khám phá hệ thống AI: Bảo hiểm nhà ở

Bạn có thể thu hẹp doanh nghiệp của mình về những quyết định?

1. Nhiệm vụ	Để giúp chủ nhà yên tâm trước sự mất mát to lớn đối với tài sản quý giá nhất mà họ sở hữu		
2. Quyết định	Marketing: Quyết định mục tiêu tiếp thị của ai	Bảo lãnh phát hành: Quyết định giá (phí bảo hiểm)	Khiếu nại: Quyết định xem có nên trả tiền yêu cầu bồi thường hay không
3. Dự đoán	Dự đoán mức sẵn lòng chi trả của từng khách hàng tiềm năng	Dự đoán khả năng chủ nhà sẽ nộp đơn yêu cầu bồi thường trong một phạm vi giá trị	Dự đoán khả năng yêu cầu được gửi là hợp lệ và phải được thanh toán
4. Đánh giá	Xác định chi phí nhằm mục tiêu đến người không mua so với không nhằm mục tiêu đến người sẽ mua	Đặt chiến lược (tăng trưởng so với lợi nhuận): xác định chi phí định giá quá thấp (thua lỗ) so với quá cao (mất khách hàng)	Xác định chi phí của việc không thanh toán yêu cầu bồi thường chính đáng (khách hàng thất vọng, danh tiếng) so với việc thanh toán yêu cầu bồi thường bất hợp pháp (chi phí)

Kinh doanh bảo hiểm cho một tài sản nào đó, chẳng hạn như một ngôi nhà có phương thức sinh lời rất đơn giản. Bạn muốn bán các hợp đồng bảo hiểm mà khoản lỗ dự kiến từ các khoản thanh toán thấp hơn phí bảo hiểm mà bạn có. Khách hàng không chỉ quan tâm đến phí bảo hiểm mà còn cả các dịch vụ liên quan, như có dễ dàng đăng ký một hợp đồng bảo hiểm hay các thủ tục yêu cầu bồi thường có phức tạp hay không. Trong ngành kinh

doanh cạnh tranh này, một công ty bảo hiểm lâu đời không thể làm gì nhiều để cải thiện mức phí bảo hiểm, nhưng nếu có thể giảm tổn thất dự kiến, nó sẽ thu được nhiều lợi nhuận hơn.

Vậy làm thế nào để một công ty bảo hiểm làm điều đó? Công ty muốn tìm những khách hàng có tổn thất dự kiến thấp hơn và bán hợp đồng cho họ, đồng thời đảm bảo rằng họ chỉ bán hợp đồng cho những khách hàng có tổn thất dự kiến cao với mức phí bảo hiểm cao hơn. Nhưng nếu không có thông tin tốt về người có mức lỗ dự kiến thấp hơn hoặc cao hơn, nhiều khách hàng sẽ nhận được mức phí bảo hiểm tương tự. Những khách hàng có tổn thất dự kiến thấp hơn phải trả quá nhiều, còn những khách hàng có tổn thất dự kiến cao hơn thì quá ít. Hơn nữa, nếu không có thông tin phù hợp, cạnh tranh không thể giải quyết vấn đề đó. Do đó, công ty muốn dự đoán những khách hàng nào có khả năng nộp đơn khiếu nại thấp hơn và nhắm mục tiêu tiếp thị cho họ. Những dự đoán này được liên kết với các quyết định tiếp thị và bảo lãnh phát hành. Công ty cũng muốn đảm bảo thanh toán các yêu cầu bồi thường khi thích hợp chứ không phải ngược lại – tức là, công ty muốn tránh gian lận. Việc đưa ra những sai lầm trong tuyên bố cuối cùng sẽ ảnh hưởng đến khả năng cạnh tranh của công ty vì điều đó làm tăng chi phí của công ty. Có như vậy



mới đánh giá được sai phạm tác động như thế nào đến hoạt động kinh doanh bảo hiểm.

Bằng cách đưa ra những quyết định này, chúng ta cũng có thể thấy chúng được liên kết với nhau như thế nào. Như chúng tôi đã lưu ý, dự đoán AI mang đến cơ hội cho ngành bảo hiểm, đặc biệt là đối với các nhà bảo hiểm, những người có nhiệm vụ dự đoán hồ sơ rủi ro của khách hàng, một ứng dụng gần như hoàn hảo cho AI. Đồng thời, bằng cách đẩy nhanh tốc độ này, công việc của các nhà tiếp thị trở nên dễ dàng hơn, vì nhân viên bán hàng có thể phản hồi khách hàng tiềm năng một cách nhanh chóng. Các vấn đề pháp lý ở các khu vực pháp lý khác nhau hạn chế cách các công ty sử dụng AI để tạo hồ sơ rủi ro, nhưng hoạt động bảo lãnh phát hành và tiếp thị lại phù hợp với giá trị của chúng. AI cũng có thể dễ dàng đánh giá tính hợp lệ của các khiếu nại, điều này sẽ quay trở lại hoạt động tiếp thị và bảo lãnh phát hành. Nhưng trên thực tế, bộ phận yêu cầu bồi thường đang làm tốt hơn công việc của mình.

Chúng ta có thể thấy các ứng dụng AI được đề cập trước đó có thể phù hợp với hệ thống này như thế nào. Ứng dụng cho phép đánh giá khiếu nại bằng cách nhấp vào máy ảnh được thiết kế để tự động hóa các quyết định khiếu nại. Nó chỉ là một

cách khác để đưa ra quyết định đó, sắp xếp gọn gàng vào bộ phận đó. Khi đó, việc có trải nghiệm khách hàng tốt hơn với các tuyên bố sẽ giúp công việc tiếp thị trở nên dễ dàng hơn và các nhà tiếp thị có thể chọn phân bổ nguồn lực của họ cho những khách hàng coi trọng trải nghiệm tốt hơn đó. Tất nhiên, điều này có thể gây ra nhiều vấn đề phức tạp hơn cho việc bảo lãnh phát hành. Liệu những khách hàng cảm thấy việc yêu cầu bồi thường dễ dàng hơn sẽ gửi nhiều yêu cầu hơn không? Tiếp thị có nhắm mục tiêu đến những khách hàng có nhiều khả năng đưa ra yêu cầu không? Chi phí của bộ phận yêu cầu bồi thường sẽ tăng lên (với nhiều yêu cầu hơn) hay giảm (với chi phí đánh giá thấp hơn)? Do đó, trong khi ứng dụng nằm gọn trong các quyết định khiếu nại, việc áp dụng nó có thể ảnh hưởng đến các quyết định khác. Điều thú vị là, về cơ bản, nó không thay đổi những gì mà các quyết định đó đang thực hiện, phán đoán liên quan hoặc các dự đoán cần thiết. Có khả năng là các ứng dụng như vậy có thể được thông qua (hoặc không) mà không có sự thay đổi trên toàn hệ thống.

Còn ứng dụng cho phép giám sát và phản hồi cho khách hàng về hồ sơ rủi ro và hành vi của họ thì sao? Trọng tâm của hệ thống hiện tại của ngành bảo hiểm là đánh giá rủi ro đó, thường là ở giai đoạn mua lại khách hàng, mặc dù trong một số trường



hợp, nó được đánh giá lại khi tái tục hợp đồng bảo hiểm. Nếu khách hàng đang mua bảo hiểm cho một ngôi nhà, họ có thể giảm phí bảo hiểm bằng cách xác minh rằng họ có hệ thống báo động có thể ngay lập tức gọi trợ giúp từ bên ngoài hoặc hệ thống giám sát nước tự động ngắt nước nếu phát hiện đường ống bị vỡ. Nhưng rủi ro của các sự kiện bất lợi dường như không được xác định bởi đặc điểm của ngôi nhà mà bởi hành vi. Ví dụ, Hiệp hội Phòng cháy chữa cháy Quốc gia báo cáo rằng nấu ăn gây ra 49% vụ cháy nhà ở Hoa Kỳ. Đi sâu vào các báo cáo của nó, và bạn thấy rằng nó không chỉ là nấu ăn mà còn là nấu ăn với dầu và đặc biệt là chiên rán. Điều này có ý nghĩa và hầu như không phải là tin tức. Sau đó, câu hỏi đặt ra là tại sao một người hiểm khi nấu ăn ở nhà lại trả mức phí bảo hiểm tương đương với một gia đình đông người chiên thức ăn của họ mỗi ngày.

Câu trả lời khá đơn giản; ngoài việc xé toạc hoàn toàn căn bếp, một công ty bảo hiểm không thể theo dõi xem ai đó có đang nấu ăn hay không, chứ chưa nói đến việc họ có đang nấu ăn bằng dầu hay không. Điều tốt nhất mà công ty có thể làm là thay đổi khoản thanh toán bảo hiểm dựa trên những sự thật này, nhưng điều đó có nghĩa là nó sẽ không bảo hiểm cho mọi người khỏi phần

xui xẻo khi nấu ăn trái ngược với hồ sơ rủi ro tổng thể của họ.

Nhưng các công nghệ AI có thể bù đắp những lỗ hổng này và theo dõi rủi ro đang diễn ra theo cách tiết kiệm chi phí. Một số là cơ chế can thiệp tự động như AI giám sát nước (ví dụ: Phyn) hoặc nguy cơ điện (ví dụ: Ting) hoạt động giống như máy dò khói. Các công ty bảo hiểm đã khuyến khích sử dụng chúng. Đối với bảo hiểm xe hơi, có các thiết bị giám sát người lái xe không chỉ xem xét mức độ lái xe của một người mà còn cả chất lượng lái xe của họ. Cài đặt những thứ này và phí bảo hiểm của bạn có thể được giảm tương ứng.

Nhưng nấu ăn, sưởi ấm, hút thuốc hoặc sử dụng nệm đều có những phẩm chất hành vi khó khăn hơn. Tuy nhiên, tất cả các hành vi này có thể được theo dõi và các số liệu liên quan để đánh giá rủi ro được gửi đến bảo hiểm các công ty trên cơ sở liên tục, với phí bảo hiểm được điều chỉnh theo thời gian thực. Tất nhiên, loại giám sát này giới thiệu quyền riêng tư và các mối quan tâm liên quan. Nhưng giống như các công ty bảo hiểm xe hơi đã có thể khiến khách hàng tự nguyện đồng ý giám sát việc lái xe của họ, họ có thể làm điều tương tự đối với bảo hiểm nhà ở. Nếu việc giám sát như vậy có nghĩa là những người nhận thấy việc giảm 25%



phí bảo hiểm khiến việc điều chỉnh hành vi của họ trở nên đáng giá và giảm 25% rủi ro hỏa hoạn, thì đó là một thỏa thuận tốt cho tất cả những người có liên quan.

Tuy nhiên, tất cả các sản phẩm bảo hiểm đáp ứng hành vi này chưa được cung cấp trước đây vì không có sự giám sát như vậy. Nhiều người không nhất thiết phải liên quan đến người tiêu dùng, nhưng bảo hiểm cho các doanh nghiệp có rủi ro rất khó đánh giá.

Việc tạo ra những sản phẩm mới này đòi hỏi sự phối hợp giữa các bộ phận hiện tại. Đặc biệt, ranh giới giữa tiếp thị và bảo lãnh phát hành trở nên mờ nhạt. Nếu bộ phận tiếp thị hình dung ra một sản phẩm mới để tận dụng lợi thế của một số công nghệ dự đoán AI mới, thì nó sẽ yêu cầu bảo lãnh phát hành điều chỉnh các quy trình của chính mình để phù hợp với nó. Ngoài ra, bộ phận nào sẽ chịu trách nhiệm giám sát và điều chỉnh phí bảo hiểm? Nó sẽ là bảo lãnh phát hành, với chuyên môn trong việc thiết lập phí bảo hiểm, hoặc yêu cầu bồi thường, với kinh nghiệm xác minh? Khi các ranh giới phân chia trở nên mờ nhạt, sẽ có áp lực phải phân bổ lại thẩm quyền ra quyết định và thay đổi người chịu trách nhiệm xử lý thông tin trong các công ty bảo hiểm.

Có lẽ lý do khiến các công ty bảo hiểm không trực tiếp giải quyết vấn đề giảm thiểu rủi ro, sử dụng công nghệ để xây dựng sản phẩm mới, chính là vì điều này khó hoặc không thể thực hiện được với các hệ thống hiện tại coi mức độ rủi ro là một thực tế của cuộc sống. Ngoài ra, giảm rủi ro có nghĩa là giảm phí bảo hiểm, điều này có thể vấp phải sự phản đối từ các đại lý và mọi người khác trong hệ thống mà khoản bồi thường gắn liền với phí bảo hiểm. Nhưng nếu một công ty bảo hiểm đặt việc giảm thiểu rủi ro thay vì chuyển giao rủi ro làm cốt lõi, thì họ sẽ thiết kế một hệ thống khuyến khích phù hợp với lợi ích của mọi người để tập trung vào việc giảm thiểu rủi ro. Trong khi phí bảo hiểm trung bình sẽ thấp hơn do rủi ro thấp hơn, công ty cũng có thể thấy lợi nhuận cao hơn và nhiều chính sách hơn. Với khả năng dự đoán nâng cao, các công ty bảo hiểm sẽ biết nhiều hơn về các nguồn rủi ro cụ thể liên quan đến một ngôi nhà hơn là chính chủ nhà. Kết quả là, nó sẽ mang lại lợi ích to lớn cho xã hội (và là một công việc kinh doanh tốt!) để chuyển trọng tâm của đề xuất giá trị từ chuyển giao rủi ro sang quản lý rủi ro. Để làm được điều này, các công ty bảo hiểm cần một hệ thống mới không chỉ đòi hỏi công nghệ mới mà còn cả sự thay đổi về tổ chức.



TÁC ĐỘNG CỦA TÙY CHỈNH

Một trong những hứa hẹn của dự đoán AI là khả năng cung cấp các sản phẩm tùy biến cao hơn, phản ánh đúng hơn hoàn cảnh, nhu cầu và mong muốn của khách hàng. Chúng tôi đã thấy điều này với quảng cáo được cá nhân hóa và giáo dục kinh doanh. Bằng cách kết hợp thông tin về sản phẩm với dự đoán về nhu cầu và thị hiếu của khách hàng, một công ty có thể cung cấp hàng hóa hoặc dịch vụ được cá nhân hóa hơn. Kết quả là, nó tạo ra giá trị bởi vì khách hàng đang nhận được thứ gì đó phù hợp với sở thích của họ.

Việc tùy chỉnh thường đòi hỏi phải tăng cường tự động hóa các quy trình. Nếu bạn chuyển từ cung cấp vài trăm hoặc nghìn sản phẩm riêng biệt sang cung cấp hàng triệu sản phẩm và thiết kế để chúng phù hợp với số lượng người tiêu dùng ít hơn, con người sẽ gặp khó khăn trong việc quản lý quy trình. Do đó, bạn muốn có một hệ thống có thể tự động hóa cả dự đoán và cả việc phân phối sản phẩm cho khách hàng. Quá trình tự động hóa đó là một thách thức đối với thiết kế và cũng nhất thiết tác động đến những người đã làm việc trong các tổ chức lâu đời. Nó tạo ra xung đột về phân bổ quyền lực có thể cản trở việc thiết kế một hệ thống mới.

Sử dụng Canvas khám phá hệ thống AI, chúng tôi có thể phân tích tác động tiềm năng của việc tùy chỉnh đối với bảo hiểm. Các công ty bảo hiểm từ lâu đã cố gắng thu thập thông tin để giúp bảo lãnh các chính sách và ấn định phí bảo hiểm phù hợp. Ngôi nhà nằm ở đâu (nói lên nguy cơ lũ lụt và hỏa hoạn), ngôi nhà có hệ thống phát hiện khói hay không và vật liệu xây dựng ngôi nhà đó đều có thể đóng một vai trò trong việc bảo lãnh phát hành. Nhưng dự đoán AI mang lại tiềm năng nhiều hơn nữa. Bằng cách thu thập thêm dữ liệu về các khiếu nại, công ty có thể cải thiện đáng kể ước tính thiệt hại dự kiến của một ngôi nhà cụ thể. Đây chính là điều mà các công ty “công nghệ bảo hiểm” như Lemonade đang cố gắng thực hiện. Nhưng chúng tôi vẫn không biết liệu AI có thể tác động đến hoạt động bảo lãnh phát hành một cách có ý nghĩa hay không.

Nhưng giả sử rằng một công ty công nghệ bảo hiểm có thể kiểm tra một ngôi nhà và đưa ra mức tổn thất dự kiến chính xác hơn và điều chỉnh phí bảo hiểm một cách thích hợp. Cũng giả sử rằng điều này cho phép công ty bảo hiểm xác định chính sách giá rõ ràng hơn dựa trên đặc điểm của ngôi nhà, báo hiệu cho chủ sở hữu biết liệu có đáng để mua hay không để thực hiện các thay đổi để tối ưu hóa phí bảo hiểm của họ. Điều này có hai tác động lớn: một là cạnh tranh và hai là về mặt tổ chức.



Hiệu quả cạnh tranh là nếu công ty công nghệ bảo hiểm xác định được những khách hàng có rủi ro thấp hơn, thì họ có thể giảm phí bảo hiểm để thu hút họ, trong khi các công ty khác khi không thể xác định được những khách hàng ít rủi ro tương tự thì không thể. Điều này rất phức tạp bởi vì nếu công ty công nghệ bảo hiểm gây được tiếng vang đủ lớn, các công ty bảo hiểm lâu đời có thể bắt chước mức phí bảo hiểm do công ty công nghệ bảo hiểm đưa ra, tìm hiểu một cách an toàn khách hàng nào sẽ nhắm mục tiêu bằng cách quan sát công ty công nghệ bảo hiểm. Tuy nhiên, quá trình này có thể sẽ mang lại lợi thế cạnh tranh cho công ty công nghệ bảo hiểm.

Các công ty bảo hiểm hiện tại cảm thấy thôi thúc phải có biện pháp tương ứng. Không giống như công ty công nghệ bảo hiểm, họ không phải là công ty mới thành lập, vì vậy họ sẽ phải thay đổi để áp dụng chính sách bảo hiểm chính xác hơn. Quy trình bán bảo hiểm truyền thống liên quan đến việc thu thập một số thông tin cơ bản từ khách hàng, thông qua các chuyên gia đánh giá thông tin đó và sau đó trả lại phí bảo hiểm. Quá trình này là một phần của tiếp thị tới khách hàng. Công ty công nghệ bảo hiểm tự động hóa quy trình này mà không cần có sự đồng ý của con người trước khi cung cấp

phí bảo hiểm cho khách hàng. Điều này có lợi thế về tốc độ, nhưng nó không có yếu tố con người. Nhiều công ty công nghệ bảo hiểm coi đây là một lợi thế chính và quảng cáo khả năng cung cấp bảo hiểm của họ với ít người hơn. Ví dụ, vào năm 2018, Lemonade tuyên bố họ có thể viết 2.500 chính sách cho mỗi nhân viên, so với 1.200 của Allstate và 650 của GEICO. Hệ quả của việc áp dụng AI trong bảo hiểm sẽ là ít người bảo hiểm, nhân viên bán hàng và người giám sát trực tiếp của họ hơn.

Nhiều người sẽ chống lại sự thay đổi. Chúng ta có thể tưởng tượng những phản đối mà họ sẽ đưa ra nếu họ bị loại khỏi vòng lặp. Các công ty bảo hiểm hiện tại đã bày tỏ sự hoài nghi. Một ấn phẩm thương mại của ngành bảo hiểm gọi đợt IPO của Lemonade là “một con kỳ lân nôn ra cầu vồng”. Chúng ta có thể lập luận rằng việc ấn định phí bảo hiểm và chính sách bán hàng không hoàn toàn khách quan, và có những yếu tố chủ quan mà một nhà bảo hiểm lành nghề có thể xác định được. Các công ty bảo hiểm sẽ cho rằng Lemonade sẽ không thể giảm tổn thất dự kiến bằng cách bỏ qua điều này. Còn những khách hàng hiện tại có tổn thất dự kiến tăng lên, do đó phải chịu phí bảo hiểm cao hơn thì sao? Liệu một công ty bảo hiểm hiện tại có thể làm điều đó mà không làm tổn hại đến thương hiệu



của chính mình không? Tất cả những phản đối này đều có một phần sự thật chính xác là do không chắc chắn rằng một tổ chức mới do AI điều khiển sẽ hoạt động. Nghịch lý là nếu một công ty bảo hiểm lâu đời không muốn đặt cược tổ chức hiện tại của mình vào sự không chắc chắn nếu được giải quyết trong AI ủng hộ, thay đổi có thể đến quá muộn. Các tổ chức hiện tại phải đối mặt với tình thế tiến thoái lưỡng nan này khi áp dụng những đổi mới có thể yêu cầu các hệ thống mới. Trong phạm vi mà các hệ thống mới mang lại sức mạnh cho những người đổi mới của họ, thì họ lại lấy đi sức mạnh của những hệ thống cũ đang vận hành.

NHỮNG ĐIỂM CHÍNH

- Hầu hết các công ty đã tạo ra các hệ thống bao gồm rất nhiều quy tắc phụ thuộc lẫn nhau, cùng với rất nhiều giàn giáo liên quan để quản lý sự không chắc chắn, đến nỗi thật khó để nghĩ về cách hoàn tác các phần của nó và suy ngẫm về khả năng thiết kế hệ thống mới mà các dự đoán của AI có thể mang lại. Vì vậy, thay vì suy nghĩ về ý nghĩa của việc thay đổi một số quy tắc hoặc giàn giáo và cách những thay đổi đó sẽ tác động đến các phần khác của hệ thống, thay vào đó, chúng tôi khuyên bạn nên bắt đầu lại

từ đầu: phương tiện trống. Khung khám phá hệ thống AI bao gồm 3 bước: (1) nêu rõ nhiệm vụ; (2) giảm doanh nghiệp xuống mức ít quyết định nhất có thể cần thiết để đạt được sứ mệnh nếu bạn có AI siêu mạnh, độ trung thực cao; và (3) nêu rõ dự đoán và phán đoán liên quan đến từng quyết định chính.

- Với bảo hiểm nhà ở, doanh nghiệp có thể rút gọn thành 3 quyết định chính: (1) tiếp thị: quyết định cách phân bổ nguồn lực tiếp thị để thu hút khách hàng nhằm tối ưu hóa lợi nhuận hoặc tăng trưởng; (2) bảo lãnh phát hành: quyết định phí bảo hiểm cho bất kỳ chính sách sở hữu nhà nhất định nào để tối đa hóa khả năng sinh lời hoặc tăng trưởng (có thể hoàn toàn không được cung cấp nếu dự đoán là rủi ro quá cao để chính sách có thể sinh lãi, do các hạn chế về giá theo quy định); và (3) yêu cầu bồi thường: quyết định xem bất kỳ yêu cầu nào được đưa ra có hợp pháp hay không và nếu có, hãy thanh toán cho yêu cầu đó. Nếu 3 AI siêu mạnh, có độ chính xác cao có thể dự đoán: (1) giá trị lâu dài của một khách hàng tiềm năng nhân với xác suất chuyển đổi, (2) khả năng nộp đơn yêu cầu nhân với mức độ yêu cầu và (3) tính hợp pháp của tuyên bố, sau đó bạn có thể thiết kế lại một doanh nghiệp bảo hiểm nhà ở nhanh, hiệu quả, chi phí thấp



và có lợi nhuận cao được cải thiện nhiều, vượt trội so với các đối thủ cạnh tranh về cả giá cả và sự thuận tiện. Đó chính xác là mục tiêu của một số công ty công nghệ bảo hiểm mới.

- Canvas khám phá hệ thống AI cũng có thể cung cấp thông tin chi tiết về các cơ hội kinh doanh mới. Ví dụ: nếu một AI dự đoán khả năng gửi khiếu nại nhân với mức độ của khiếu nại trở nên đủ tốt để có thể tạo dự đoán ở mức độ nguy hiểm hoặc mức độ nguy hiểm phụ (ví dụ: cảm biến để cung cấp khả năng phát hiện sớm nguy cơ hỏa hoạn do điện tăng cao hoặc nguy cơ lũ lụt do đường ống bị rò rỉ tăng lên), thì công ty có thể dự đoán giải pháp giảm thiểu rủi ro nào sẽ có lợi tức đầu tư đủ cao để đảm bảo chi phí thực hiện. Sau đó, công ty bảo hiểm có thể trợ cấp cho thiết bị giảm thiểu rủi ro và giảm phí bảo hiểm, cung cấp một đề xuất giá trị hoàn toàn mới cho khách hàng: giảm thiểu rủi ro. Công ty bảo hiểm không chỉ chuyển rủi ro từ chủ nhà sang người vận chuyển, mà còn làm giảm rủi ro – một dịch vụ có giá trị mà ngành bảo hiểm trước đây không cung cấp trừ một số ít trường hợp. Để khai thác triệt để cơ hội này, cần phải thiết kế một hệ thống mới được tối ưu hóa để giảm thiểu rủi ro.

Chương 18

DỰ ĐOÁN THAY ĐỔI HỆ THỐNG



Một bệnh nhân đến khoa cấp cứu vì đau ngực. Liệu đó có phải là một cơn đau tim? Một bác sĩ có thể tìm ra bằng cách tiến hành một bài kiểm tra. Một xét nghiệm dương tính sẽ cho phép bác sĩ điều trị bệnh nhân nhanh chóng, với những lợi ích rõ ràng. Nhưng các xét nghiệm đắt tiền và cũng xâm lấn. Xét nghiệm hình ảnh sử dụng bức xạ, có thể làm tăng nguy cơ ung thư lâu dài. Một phiên máy chạy bộ có nguy cơ ngừng tim nhỏ nhưng được xác định. Thông tin liên quan đến bức xạ cũng như nguy cơ tổn thương động mạch. Đó không phải là một quyết định dễ dàng.

Bác sĩ cần cân nhắc những lợi ích và chi phí này. Khả năng bệnh nhân thực sự lên cơn đau tim là bao nhiêu? Đây là một dự đoán. Nếu dự đoán là khả năng xảy ra cao, điều đó sẽ thúc đẩy bác sĩ tiến tới xét nghiệm và điều trị. Nếu nó thấp, khả năng xét



nghiệm bị lãng phí càng lớn, gây rủi ro cho bệnh nhân vì những lý do nhỏ nhất.

Khi quyết định có tiến hành xét nghiệm hay không, lợi ích là tiết lộ thông tin về các biện pháp can thiệp tiếp theo (chẳng hạn như đặt stent). Nếu bệnh nhân bị đau tim, thì cô ấy sẽ được hưởng lợi từ các phương pháp điều trị đau tim; nếu không, thì cô ấy sẽ không có lợi ích gì. Vì vậy, giá trị của xét nghiệm hoàn toàn bắt nguồn từ giá trị quyết định mà nó tạo ra, trong việc nhắm mục tiêu can thiệp vào những bệnh nhân sẽ được hưởng lợi nhiều nhất.

Thử nghiệm không phải là một bữa trưa miễn phí. Các bài kiểm tra căng thẳng có thể tốn hàng nghìn đô la. Catheterization có thể tốn hàng chục nghìn đô la. Bạn vui vẻ trả tiền cho một bài kiểm tra căng thẳng để tránh chi phí cao gấp 10 lần.

Đây chỉ là chi phí tiền tệ. Một số xét nghiệm yêu cầu theo dõi và quan sát qua đêm. Sau đó, các bài kiểm tra tự tạo ra rủi ro của chính chúng: Trong tất cả các bài kiểm tra hình ảnh, các bài kiểm tra căng thẳng mang theo liều bức xạ ion hóa cao nhất, được cho là làm tăng đáng kể nguy cơ ung thư lâu dài. Tập thể dục trên máy chạy bộ trong bối cảnh đau tim có nguy cơ ngừng tim nhỏ nhưng chắc chắn. Đưa bệnh nhân trực tiếp đến đặt ống thông tiểu có lợi ích là việc điều trị (thường là đặt stent) có

thể được hoàn thành như một phần của quy trình giống như chẩn đoán.

Sau đó, một lần nữa, bạn không nhất thiết muốn tránh kiểm tra gắng sức và các rủi ro liên quan vì quy trình đặt ống thông tiểu cũng có rủi ro. Một thủ thuật xâm lấn bao gồm một lượng lớn bức xạ ion hóa cũng như tiêm chất cản quang vào tĩnh mạch có thể gây suy thận và có nguy cơ tổn thương động mạch và đột quỵ do suy nhược. Vì vậy, trước khi quyết định có nên điều trị cho một bệnh nhân bị đau tim hay không, các quyết định về việc có nên kiểm tra xem ai đó có bị đau tim hay không và liệu có nên tiến hành kiểm tra gắng sức trước hay đưa bệnh nhân trực tiếp đến thông tim.

Các quyết định xét nghiệm thuộc về bác sĩ. Các xét nghiệm có xu hướng là tiền thân của thủ tục. Nhưng các bác sĩ áp dụng rất nhiều phán đoán trong việc quyết định phải làm gì. Tuổi của bệnh nhân là bao nhiêu? Mọi người có cần được chăm sóc nhiều hơn không (chẳng hạn như trong viện dưỡng lão)? Bệnh nhân có những tình trạng nào khác (ví dụ ung thư)? Tất cả những cân nhắc vào những quyết định.

Bây giờ, giả sử bác sĩ có sự trợ giúp – sự trợ giúp siêu phàm – với phần dự đoán của phương trình: nghĩa là, giả sử một AI có thể đánh giá liệu một



bệnh nhân có cần xét nghiệm đó hay không và thực hiện nó một cách nhanh chóng. Không khó để thấy những lợi ích tiềm năng ở đây. Đây không phải là giả thuyết. Các nhà kinh tế Sendhil Mullainathan và Ziad Obermeyer đã xây dựng một AI dựa trên cùng một thông tin mà các bác sĩ có trong khoa cấp cứu khi chẩn đoán bệnh nhân. Họ đã chỉ ra rằng AI của họ chính xác hơn các bác sĩ trong việc dự đoán các cơn đau tim. Quyết định của các bác sĩ cho thấy nhiều thử nghiệm quá mức. Bệnh nhân được xét nghiệm mà họ không cần. Với một số ưu đãi trong hệ thống chăm sóc sức khỏe của Hoa Kỳ, có lẽ kết quả này rất dễ đoán trước. Không ai muốn đối mặt với trách nhiệm pháp lý liên quan đến việc không thử nghiệm, đặc biệt là khi bạn được trả nhiều tiền hơn để thử nghiệm nhiều hơn.

Đáng ngạc nhiên, họ cũng tìm thấy một số lượng đáng kể các hoạt động kiểm tra dưới mức. Hàng nghìn bệnh nhân mà AI dự đoán có nguy cơ cao không bao giờ được xét nghiệm. Những bệnh nhân mà thuật toán AI dự đoán là có rủi ro cao cuối cùng lại có kết quả tồi tệ hơn, đó là phải quay lại bệnh viện hoặc tử vong.

Thuật toán AI dường như tuyệt vời trên mọi mặt. Nó rẻ, nhanh và dường như ít bị lỗi hơn theo cả hai hướng. Bạn có thể thực hiện cùng một khối

lượng xét nghiệm và phân bổ lại chúng từ những bệnh nhân có nguy cơ thấp đến nguy cơ cao và đạt được kết quả tốt hơn xung quanh, giúp bệnh nhân và có khả năng giảm rủi ro trách nhiệm pháp lý. Ngoài ra, bạn có thể giảm xét nghiệm và giữ nguyên chất lượng chăm sóc sức khỏe.

GIẢI PHÁP ĐIỂM VÀ ỨNG DỤNG

Điều này có vẻ dễ hiểu đối với việc áp dụng AI. Chẩn đoán bằng AI trong các khoa cấp cứu cung cấp một giải pháp điểm tuyệt vời và hiệu quả. AI hỗ trợ bác sĩ, người sau đó có thể đưa ra quyết định tốt hơn và từng bước cải thiện sức khỏe của bệnh nhân với chi phí thấp hơn một chút. Quy trình làm việc không thay đổi và công việc của mọi người không bị đe dọa. Bác sĩ không dành nhiều thời gian cho bước chẩn đoán, vì tách rời khỏi phán đoán.

Câu hỏi đặt ra là liệu những lợi ích này có xứng đáng với chi phí triển khai một công cụ mới hay không. Các nhà quản lý chăm sóc sức khỏe đang phải đối mặt với nhiều công nghệ mới đầy hứa hẹn. Mỗi cái đều yêu cầu đào tạo và điều chỉnh các quy trình. Cái nào cũng có rủi ro. Trên thực tế, các công nghệ hiếm khi tốt như trong quá trình thử nghiệm. Các quản trị viên có thể quyết định rằng lợi ích gia tăng của giải pháp điểm AI cho chẩn đoán cơn đau tim là không xứng đáng với chi phí bỏ ra.



Tuy nhiên, các quản trị viên có thể thấy một giải pháp ứng dụng AI hấp dẫn. Thay vì một AI cung cấp dự đoán cho bác sĩ, họ triển khai một AI xác định xem có nên xét nghiệm hay không. Bác sĩ được rút quyết định cho bệnh nhân. Thay vào đó, khi bệnh nhân đến phòng cấp cứu, máy sẽ dự đoán khả năng bệnh nhân bị đau tim. Nếu dự đoán nằm dưới ngưỡng nào đó, bệnh nhân sẽ được đưa về nhà. Nếu dự đoán nằm trong một số phạm vi trung gian, bệnh nhân sẽ được gửi đến một bài kiểm tra căng thẳng. Nếu AI dự đoán rằng bệnh nhân rất có thể bị đau tim, thì bệnh nhân sẽ được chuyển trực tiếp đến thông tim. Ngưỡng giữa các quyết định gửi về nhà, kiểm tra căng thẳng và đặt ống thông tiểu là sự phán xét. Trong trường hợp này, phán quyết đó có thể đến từ lãnh đạo bệnh viện hoặc từ một ủy ban gồm các bác sĩ và chuyên gia y tế khác.

HỆ THỐNG CÓ THỂ XỬ LÝ KHÔNG?

Các bệnh viện thường được tổ chức thành hai bộ phận bao trùm, mỗi bộ phận có trách nhiệm riêng. Cơ quan quản lý chịu trách nhiệm về mặt tiền, thu các khoản thanh toán (hoặc các khoản bồi hoàn từ chính phủ hoặc cơ quan khác bảo hiểm), tuyển dụng nhân viên và mua sắm các nguồn lực. Y học chịu trách nhiệm chẩn đoán và điều trị cho bệnh nhân. Trong các bệnh viện, có các bộ phận trong

mỗi bộ phận, nhưng trong việc phân bổ quyền quyết định, một bộ phận có tiền và các nguồn lực trong khi bộ phận kia có y tế. Xung đột giữa hai người vẫn đang tiếp diễn. Nhưng hầu hết các bệnh viện đang hoạt động đã giải quyết các điều khoản về sự tham gia giữa các bộ phận cho phép mỗi bộ phận đưa ra quyết định của riêng mình, tùy thuộc vào những ràng buộc do bộ phận khác áp đặt.

Đối với AI chẩn đoán trong khoa cấp cứu, người ta có thể hình dung ra các nguồn kháng cự tiềm tàng. Các bác sĩ nhận được một số lợi ích cá nhân từ việc thực hiện nhiều xét nghiệm hơn, dưới hình thức bảo vệ khỏi rủi ro sơ suất hoặc có thể là thêm thu nhập. Nếu AI đang đưa ra những dự đoán tốt hơn bác sĩ, thì việc đào tạo và kinh nghiệm của bác sĩ sẽ trở nên ít quan trọng hơn. Giá trị của chúng bị giảm đi so với những người xây dựng những AI đó. Quản trị viên có thể lo lắng về chi phí thực hiện. Trong giải pháp ứng dụng, các bác sĩ không được đưa ra quyết định và quá trình đào tạo cũng như kinh nghiệm của họ có thể không còn phù hợp. Các bác sĩ đã đưa ra nhiều quyết định trong quá khứ có thể nghi ngờ rằng AI có thể làm tốt hơn họ.

Tiềm năng của AI có thể còn lớn hơn, nhưng điều đó đòi hỏi một số thay đổi hệ thống. Ở hầu hết các khoa cấp cứu của bệnh viện, khi bệnh nhân



đến, các bác sĩ sẽ quyết định cho bệnh nhân về nhà, tiến hành kiểm tra gắng sức hay đặt ống thông tiểu. Rất lâu trước khi quyết định đó được đưa ra, chính quyền quyết định sẽ cung cấp các bài kiểm tra nào. Hiện tại, các ưu đãi dường như được liên kết. Mullainathan và Obermeyer ước tính rằng 15% bệnh nhân mà bác sĩ gửi đi xét nghiệm thực sự bị đau tim. Với độ chính xác đó, cả bác sĩ và quản lý dường như đồng ý rằng gửi bệnh nhân đi kiểm tra căng thẳng trước là tốt nhất.

Một AI từng bước cải thiện dự đoán này sẽ cải thiện kết quả của bệnh nhân với ít thay đổi về động cơ khuyến khích. Với độ chính xác 20%, cho dù AI được triển khai như một giải pháp điểm hay giải pháp ứng dụng, cả bác sĩ và ban quản lý đều có thể muốn gửi bệnh nhân đi kiểm tra căng thẳng trước. Những cải tiến như vậy có thể tăng dần đến mức cả bác sĩ và chính quyền đều không nghĩ rằng nó đáng để áp dụng.

Nếu có thể xây dựng một AI gần như hoàn hảo, thì sẽ có cơ hội cải thiện việc chăm sóc bệnh nhân và thiết kế lại hệ thống. Nếu AI dự đoán rằng một bệnh nhân đang lên cơn đau tim với độ chính xác 99%, thì cả bác sĩ và ban quản lý đều nhận ra rằng tốt nhất là tiến hành đặt ống thông tiểu và bỏ qua bài kiểm tra căng thẳng. Khi AI đủ tốt trên tất cả bệnh nhân, mọi người sẽ đồng ý rằng các bài kiểm

tra căng thẳng là không cần thiết. Chính quyền sẽ ngừng cung cấp bài kiểm tra căng thẳng như một tùy chọn và các bác sĩ sẽ không bao giờ muốn sử dụng bài kiểm tra này.

Giữa độ chính xác ngày nay và một AI gần như hoàn hảo, có một khoảng trung gian mà động cơ khuyến khích của bác sĩ và quản lý có thể thay đổi và xung đột. Quản trị viên có thể nhận thấy chi phí kiểm tra căng thẳng không cần thiết lớn hơn so với bác sĩ, vì quản trị viên không muốn sử dụng tài nguyên hoặc vì họ ít lo lắng hơn về rủi ro trách nhiệm pháp lý. Trong trường hợp đó, với độ chính xác 50%, các bác sĩ có thể vẫn muốn kiểm tra căng thẳng, nhưng các quản trị viên muốn đưa bệnh nhân đi đặt ống thông.

Giải pháp, dựa trên sự phân chia được mô tả, có vẻ dễ dàng. Nó cho thấy một sự thay đổi hệ thống tương đối đơn giản, trong đó chính quyền lấy quyết định từ các bác sĩ. Chính quyền sẽ từ chối cung cấp các bài kiểm tra căng thẳng. Sau đó, các bác sĩ chỉ có thể lựa chọn giữa việc đưa bệnh nhân trực tiếp đến đặt ống thông tiểu hoặc đưa bệnh nhân về nhà. Trong trường hợp đó, bệnh nhân tiến hành đặt ống thông tiểu ngay lập tức, các quản trị viên có được những gì họ muốn, và các bác sĩ cầu nhàu nhưng chấp nhận bàn tay của họ.



Đối với chúng tôi, điều đó dường như không thể. Các bác sĩ sẽ đẩy lùi. Một cơ quan quản lý có thể được gọi vào. Quyền của bệnh nhân sẽ được thảo luận. Khi những người ra quyết định không còn liên kết với nhau, cách thức phân bổ các quyết định trong hệ thống hiện tại có thể không còn được chấp nhận. Ngược lại, sự thay đổi trong sự liên kết quyết định này có thể có nghĩa là công cụ AI không bao giờ được áp dụng. Cho rằng AI cải thiện với dữ liệu phản hồi, sự sai lệch trong trung hạn có thể ngăn không cho AI có lợi lâu dài, có lợi cao.

Để khắc phục điều này, các bác sĩ và quản trị viên sẽ cần phải cùng nhau thiết lập một cấu trúc quyết định mới. Điều này có nghĩa là một sự thay đổi hệ thống lớn hơn là bỏ qua một quyết định thử nghiệm. Đối với các công cụ AI như công cụ mà chúng tôi đã mô tả cho các cơn đau tim, sự thay đổi hệ thống lớn hơn đó có thể không đáng.

Mặt khác, một khi bạn nhận ra khả năng thay đổi hệ thống, sẽ có cơ hội sử dụng Khung khám phá hệ thống AI để mô phỏng lại hình thức của thuốc cấp cứu.

Y HỌC CẤP CỨU

Xây dựng một canvas là một bài tập cần thiết trong dự báo. Cụ thể, bạn sẽ lựa chọn một ngành nghề phức tạp và mô tả nó ở dạng cơ bản nhất. Như

chúng ta đã vạch ra trong chương trước, công việc này bắt đầu với nhiệm vụ. Nhiệm vụ của khoa cấp cứu có thể là “cải thiện tình trạng của những bệnh nhân bị bệnh nặng và bị thương thông qua dịch vụ chăm sóc chất lượng cao, hiệu quả về chi phí”.

Để cung cấp dịch vụ chăm sóc đó, các quản trị viên và bác sĩ đưa ra hàng nghìn quyết định. Bài tập với canvas là tóm tắt các quyết định này thành các danh mục thiết yếu của chúng. Đối với thuốc cấp cứu, một cách để chia nó chỉ thành hai quyết định cốt lõi (xem Hình 18-1). Có quyết định điều trị khi một chuyên gia y tế quyết định cung cấp dịch vụ y tế nào cho bệnh nhân và có quyết định cung cấp nguồn lực khi chính quyền quyết định trang bị và nhân sự nào sẽ có sẵn trong khoa. Phương pháp điều trị phụ thuộc vào chẩn đoán và sự hiểu biết về bằng chứng y tế hỗ trợ các phương pháp điều trị khác nhau được đưa ra trong chẩn đoán. Chẩn đoán, như chúng ta đã thảo luận, là một vấn đề dự đoán. Nguồn cung cấp cũng phụ thuộc vào chẩn đoán, nhưng không phải bệnh nhân của bệnh nhân. Thay vào đó, việc cung cấp nguồn lực phụ thuộc vào dự đoán về sự phân bố các chẩn đoán theo thời gian.

Hãy đưa bài tập suy đoán này đến giới hạn, nơi AI để chẩn đoán trở nên hữu ích trong một loạt cài đặt. Eric Topol, một bác sĩ tim mạch mà chúng tôi



đã giới thiệu trong Chương 8, tin rằng những dự đoán tốt hơn của AI sẽ báo trước một thời kỳ hoàng kim của y học, nơi các bác sĩ có thể tập trung vào các khía cạnh chăm sóc sức khỏe của con người và để lại các quy trình cơ học cho máy móc. Dự đoán cơn đau tim chỉ là một ví dụ cho thấy AI đang bắt đầu vượt trội hơn các bác sĩ trong việc chẩn đoán.

Hình 18-1: Khám phá hệ thống AI: Cấp cứu

Bạn có thể thu hẹp doanh nghiệp của mình về những quyết định nào?

1. Nhiệm vụ	Cải thiện kết quả điều trị của bệnh nhân bị bệnh nặng và bị thương thông qua dịch vụ chăm sóc chất lượng cao, tiết kiệm chi phí	
2. Quyết định	Điều trị: Quyết định phương pháp điều trị nào sẽ được kê đơn	Nguồn lực: Quyết định số lượng từng loại thiết bị và nhân sự sẽ được triển khai
3. Dự đoán	Chẩn đoán: Dự đoán nguyên nhân gây ra các triệu chứng của bệnh nhân	Số lượng và loại bệnh nhân: Dự đoán số lượng bệnh nhân và phân bố chẩn đoán
4. Đánh giá	Hậu quả của việc điều trị quá mức, điều trị thiếu và điều trị không đúng cho bệnh nhân là gì?	Hậu quả của việc có quá ít thiết bị và nhân viên so với việc có quá nhiều là gì?

Khi các dự đoán cơ bản về quyết định số một trở nên tốt hơn và nhanh hơn, bệnh nhân ngày càng nhận được các phương pháp điều trị thích hợp. Với đủ dữ liệu, có thể chuyển hoàn toàn những dự đoán đó ra khỏi không gian phân loại trong khoa cấp cứu của bệnh viện và quay trở lại nhà của bệnh

nhân. Do đó, trước khi xe cứu thương được gọi, có thể hình dung các dự đoán chất lượng cao để đưa ra chẩn đoán .

Chẩn đoán như vậy có thể cho phép thực hiện nhiều thay đổi ở cấp độ hệ thống. Bệnh nhân có thể hoàn toàn bỏ qua khoa cấp cứu và được chuyển thẳng đến khoa y tế liên quan, có thể là khoa tim mạch hoặc khoa chỉnh hình. Nhiều bệnh nhân sẽ không bao giờ cần đến bệnh viện sau khi được chẩn đoán mắc bệnh mà được sĩ hoặc bác sĩ chăm sóc chính có thể giúp điều trị. Các nhân viên y tế có thể đưa bệnh nhân đến các bệnh viện khác nhau, dựa trên khả năng chuyên môn có liên quan – và sự chậm trễ trong hệ thống – có thể ở đâu.

Vai trò của nhân viên y tế cũng có thể thay đổi. Nhân viên y tế có thể đào tạo để chuyên về các điều kiện y tế cụ thể. Sau đó, khi bệnh nhân gặp trường hợp khẩn cấp cần có chuyên môn về tim mạch, nhân viên y tế được đào tạo phù hợp có thể được gửi đến, cùng với xe cứu thương chứa đầy thiết bị tập trung vào tim mạch. Bệnh nhân sẽ không phải chờ đợi đến bệnh viện để được điều trị.

Chúng tôi biết bạn đang nghĩ gì. Không thể nào. Ngay cả khi các dự đoán đủ tốt, các nhân viên y tế không thể chuyên biệt như vậy. Họ cần phải là những người nói chung vì họ cần xử lý bất cứ điều



gì có thể xảy ra theo cách của họ. Nếu chúng tôi cần các chuyên gia, mỗi trạm cứu thương sẽ cần nhiều nhân viên y tế hơn nhiều so với khả năng đào tạo và chi trả của chúng tôi.

Đó là lúc quyết định thứ hai xuất hiện: cung cấp nguồn lực. Nếu dự đoán về sự phân bố nhu cầu là đủ tốt (và mật độ dân số đủ cao), thì có thể đặt các thiết bị cần thiết và nhân viên được đào tạo tốt cần thiết vào đúng nơi, đúng thời điểm. Dự đoán AI để chẩn đoán bệnh nhân trở thành phần bổ sung cho dự đoán AI về phân phối nhu cầu theo thời gian.

Ở mức giới hạn, phiên bản thuốc cấp cứu này có hầu hết bệnh nhân được điều trị tại nhà bằng thiết bị chuyên dụng do nhân viên y tế được đào tạo về các kỹ năng liên quan cụ thể mang đến. Những bệnh nhân được gửi đến bệnh viện là những người cần phải ở lại lâu dài hoặc có nhiều nhân viên y tế để điều trị cho họ. Kết quả bệnh nhân tốt hơn. Các bệnh viện trở nên nhỏ hơn. Đào tạo y tế và thay đổi nhân sự. Sức mệnh cải thiện kết quả thông qua dịch vụ chăm sóc chất lượng cao với chi phí hiệu quả có thể đạt được một cách ngoạn mục.

Đây không phải là sắp xảy ra. Nó có lẽ thậm chí không thể. AI vẫn chưa đủ tốt và có thể không bao giờ. Chi phí để đại tu hệ thống sẽ là đáng kể, mặc dù khoản tiết kiệm dài hạn cũng có thể là như

vậy. Sức đề kháng của các chuyên gia y tế và các nhà quản lý trước sự thay đổi mạnh mẽ như vậy sẽ là điều phi thường. Tuy nhiên, nó có thể xảy ra ở quy mô nhỏ hơn. Nhiều khu vực pháp lý đã cử bác sĩ cùng với nhân viên y tế. Máy dự đoán có thể giúp xác định khi nào cần đến bác sĩ. Một tổ chức kinh doanh đang thực hiện bước đầu tiên cho một hệ thống mới, tích hợp AI vào công văn để xác định bác sĩ nào (và thiết bị nào) là cần thiết.

NHỮNG LỰA CHỌN HỆ THỐNG

Chúng tôi chắc chắn không thành thật khi tuyên bố rằng thay đổi hệ thống là phức tạp và sau đó đưa ra một quan điểm rất đơn giản về những lựa chọn nào xác định một hệ thống và sự thay đổi của chúng đi kèm với việc áp dụng công nghệ mới như thế nào. Tuy nhiên, đôi khi để minh họa cho sự phức tạp, đơn giản hóa quá mức cũng có giá trị. Ý tưởng là loại bỏ một mớ hỗn độn để tìm ra những phần quan trọng trong bản chất của nó. Đó là những gì mà Canvas khám phá hệ thống AI thực hiện và là những gì thử nghiệm tưởng tượng của chúng tôi về AI trong y học khẩn cấp đã đạt được.

Có hai loại lựa chọn rộng rãi xác định cái mà chúng ta gọi là “hệ thống”. Đầu tiên là ai nhìn thấy gì. Thứ hai là ai quyết định cái gì. Điều này cung cấp một khuôn khổ khác để hiểu những thay đổi hệ



thống mà chúng tôi đã mô tả. Đối với những người nhìn thấy những gì, công việc của một tổ chức là lọc thông tin. Một lượng lớn thông tin có thể được đưa vào tổ chức, nhưng nhiệm vụ chính là đưa vào thông tin có liên quan đến việc ra quyết định hiện tại. Do đó, một tổ chức bắt đầu phân bổ các vai trò trong việc quan sát và xử lý thông tin cho các khu vực cụ thể – các bộ phận như hành chính và y tế. Trong một số trường hợp, các bộ phận thu thập thông tin và giữ nó trong bộ phận. Ở những nơi khác, các bộ phận lọc và truyền thông tin đến các bộ phận khác. Y học nhìn thấy chẩn đoán dự đoán cho từng bệnh nhân khi họ đến khoa cấp cứu. Quản trị chỉ nhìn thấy phân phối sau khi thực tế. Vấn đề là sẽ có những thông tin hoàn toàn không được đưa vào tổ chức và thậm chí còn có ít thông tin được sử dụng phổ biến trong toàn bộ tổ chức.

Đối với ai quyết định cái gì, việc phân bổ thẩm quyền ra quyết định là sự kết hợp giữa ai có kỹ năng sử dụng thông tin tốt nhất, ai có thông tin và ai có động cơ phù hợp để đưa ra quyết định phù hợp với lợi ích của tổ chức. Về nguyên tắc, nếu có một siêu cá nhân, thì một người chỉ có thể đưa ra mọi quyết định. Nhưng không có cá nhân như vậy tồn tại. Như vậy, trong việc phân bổ thẩm quyền ra quyết định, tổ chức thực hiện sự đánh đổi ở mọi nơi.

Các bộ phận tận dụng lợi thế của chuyên môn hóa và giảm thiểu nhu cầu phối hợp giữa chúng, để mọi người có thể tiếp tục công việc của mình. Nhưng điều đó cũng có nghĩa là không có ai có thể nhìn thấy mọi thứ. Thiết kế tổ chức tốt đảm bảo rằng các bộ phận có thể nhận ra khi một tình huống nằm ngoài phạm vi của họ và chuyển nó cho người khác. Nhưng không có cách hoàn hảo để làm điều này. Do đó, các tổ chức phân bổ các quyết định và thông tin cho một cái gì đó hoạt động tốt. Tổ chức càng lớn thì việc dàn xếp này càng cần thiết hoặc càng có khả năng xảy ra. Các tổ chức nhỏ hơn có ít bộ phận hơn, nhưng điều đó hạn chế quy mô của chúng.

Các hệ thống hiện tại được xác định bởi trách nhiệm của các bộ phận sẽ rất tốt trong việc áp dụng các công nghệ mới mà lợi ích của chúng chỉ giới hạn trong một bộ phận. Trong những trường hợp này, có việc áp dụng các giải pháp điểm AI và giải pháp ứng dụng AI. Các cơ quan quản lý bệnh viện đã áp dụng AI để lên lịch cho nhân viên bảo trì và sàng lọc sơ yếu lý lịch, cả hai đều liên quan rõ ràng đến việc phân bổ nguồn lực. Các công nghệ phục vụ sứ mệnh của bộ phận và không yêu cầu thay đổi người quyết định điều gì, điều này có thể dẫn đến sự phản kháng trong nội bộ vì những thay đổi về trách nhiệm thường theo sau những thay đổi về quyền lực.



Ngược lại, đối với những công nghệ mới mà lợi ích của nó được phân bổ giữa các bộ phận hoặc tệ hơn là liên quan đến chi phí cho bộ phận này và lợi ích ở bộ phận khác, thì việc áp dụng sẽ khó khăn hơn nhiều. Ngay cả khi những cơ hội đó có thể được nhận ra, thì việc đạt được sự phối hợp đòi hỏi phải phân bổ lại ai là người quyết định điều gì. Đó là cơ sở cho những thách thức trong việc áp dụng AI để chẩn đoán tại các khoa cấp cứu. Do đó, cần phải đối mặt và hiệu chỉnh lại các thỏa thuận được đàm phán cẩn thận và mạnh mẽ. Thay đổi như vậy, nói một cách nhẹ nhàng, là đột phá và do đó có thể dễ dàng đạt được hơn khi bắt đầu lại từ đầu.

Khi nói đến dự đoán AI, những cải tiến mới có thể xuất hiện dưới nhiều hình thức. Nhiều giải pháp điểm hoặc giải pháp ứng dụng mà các bộ phận có thể áp dụng mà không gây ra xung đột hoặc khó khăn trong việc phối hợp. Đối với những người khác, điều đó là không thể. Việc áp dụng chúng liên quan đến sự gián đoạn và thay đổi. Câu hỏi ở đây là làm thế nào các nhà lãnh đạo doanh nghiệp có thể biết liệu họ có đang bỏ lỡ những cơ hội sau này hay không bởi vì tổ chức của họ đã được thiết kế để không nhìn thấy chúng, chứ đừng nói đến việc đánh giá tiềm năng của chúng một cách thuận lợi khi chúng phát sinh.

Do đó, để tiến hành các thử nghiệm suy nghĩ về việc liệu bạn có thể áp dụng đổi mới AI hiện tại hoặc tiềm năng trong tổ chức của mình hay không, chúng tôi đưa ra một quy trình hai bước như sau:

BUỐC 1: Xác định thông tin mà dự đoán cung cấp hoặc sự không chắc chắn mà nó giải quyết và quyết định hoặc các quyết định mà nó sẽ cải thiện.

BUỐC 2: Thông tin đó hoặc những quyết định đó có trong một bộ phận hay trên nhiều bộ phận không?

Bây giờ, hãy quay lại việc áp dụng AI trong các khoa cấp cứu, cụ thể là AI dự đoán liệu bệnh nhân có bị đau tim hay không. Ở bước 1, AI dự đoán chẩn đoán. AI được thiết lập để sử dụng thông tin chỉ có trong hệ thống bệnh viện cho các bác sĩ tại thời điểm bệnh nhân vào phòng cấp cứu. Quyết định chính mà chẩn đoán này đưa ra là nên cung cấp các xét nghiệm và phương pháp điều trị nào. Đối với bước 2, rõ ràng thông tin này và quyết định này đều nằm trong tầm nhìn của các bác sĩ trong bộ phận y tế. Việc lựa chọn có áp dụng các thuật toán như vậy để hỗ trợ loại phân loại này có vẻ là một lựa chọn dễ dàng, dường như khiến nó có nhiều khả năng được áp dụng như một giải pháp điểm hơn.



Sau đó, chúng tôi khám phá câu hỏi liệu khoa cấp cứu có nên đưa ra bài kiểm tra căng thẳng hay không. Điều này thêm một quyết định mới vào bước 1: các tài nguyên nên có sẵn. Nhìn vào bước 2, điều này không còn bị hạn chế đối với một bộ phận và như đã lưu ý trước đó, những thách thức ở cấp hệ thống có thể ngăn cản việc áp dụng. Cũng có thể có những cách để cải thiện AI yêu cầu điều phối thông tin giữa các bộ phận. Chẩn đoán AI trước khi bệnh nhân đến bệnh viện có thể yêu cầu quyền truy cập vào dữ liệu bệnh nhân trong nhiều tuần và có lẽ nhiều năm trước sự kiện. Điều đó yêu cầu chính quyền phê duyệt việc thu thập dữ liệu bên ngoài giới hạn của bệnh viện. Các quy định sẽ cần phải được điều chỉnh. Bệnh nhân sẽ cần phải được thuyết phục. Và trong đó có xung đột tiềm năng. Cơ cấu tổ chức hiện tại của các bệnh viện không thể dễ dàng áp dụng các công nghệ dự đoán AI và những thay đổi dẫn đến hoạt động của chúng. Nó đòi hỏi một hệ thống mới.

THẬT KHÓ ĐỂ ĐƯA RA DỰ ĐOÁN VỀ CÁC HỆ THỐNG

Đến năm 1880, rõ ràng là điện có tiềm năng to lớn để cải thiện cách thức hoạt động của các nhà máy. Nhưng phải mất 40 năm nữa để hiểu làm thế nào để thiết kế một hệ thống nhà máy tận dụng

năng lượng điện. Theo hiểu biết tốt nhất của chúng tôi, không ai hình dung ra các hệ thống cuối cùng được xây dựng bằng điện. Quá trình khám phá mất nhiều thời gian, khi con người tiến hóa trong sự hiểu biết của họ về những gì điện có thể làm.

Với AI, chúng ta đang tiến gần đến năm 1880 hơn là năm 1920. Có vẻ như AI sẽ dẫn đến các hệ thống hoàn toàn mới trong nhiều ngành công nghiệp. Chúng tôi tin rằng chính vì vai trò của dự đoán AI trong việc ra quyết định và thực tế là khi nó tạo ra các quyết định mới, các hành động được thực hiện và kết quả từ những quyết định đó phản ánh thay vì bảo vệ hoặc che giấu sự không chắc chắn tiềm ẩn. Bởi vì các quyết định thường tương tác với các quyết định khác, nên có khả năng ảnh hưởng đến toàn hệ thống. Thực tế là nếu không có dự đoán, những cách thức hoạt động hiện tại có thể đáng tin cậy – ngay cả khi chúng lãng phí về mặt kỹ thuật – cho chúng ta biết rằng nếu không có sự đổi mới hệ thống, việc áp dụng AI sẽ bị tắt tiếng.

Trong vài chương trước, chúng tôi đã cung cấp các hướng dẫn và phương pháp giúp giải nén các hệ thống mà AI có thể xâm nhập. Bảo hiểm và chăm sóc sức khỏe là những ngành có thể chín muồi để thay đổi. Nhưng chúng ta đang ở những ngày đầu của Giai đoạn Chuyên giao. Ngay cả khi bạn sử



dụng các phương pháp này để hiểu hệ thống có thể thay đổi như thế nào, vẫn còn nhiều bước nữa trước khi bạn có thể xác định liệu thay đổi đó có đáng giá hay không và những tiến bộ dự đoán AI nào là cần thiết để khiến việc tái cấu trúc trở nên đáng giá. Ngay cả khi đó, đổi mới hệ thống đòi hỏi phải có sự đột phá, kéo theo những thay đổi trong việc phân bổ quyền lực giữa người thắng và người thua. Sự cân bằng giữa những người muốn thay đổi và những người chống lại nó sẽ quyết định liệu nó có xảy ra hay không và nhanh như thế nào.

Do đó, chúng tôi nhấn mạnh lời cảnh báo rằng lịch sử công nghệ cho chúng ta biết nhiều về việc liệu một công nghệ có mục đích chung, như AI có vẻ như, sẽ gây rối và mất thời gian hay không. Một phân tích rõ ràng và khách quan về chính xác những gì AI đang làm, cụ thể là dự đoán và nơi AI xâm nhập vào các hệ thống – trong việc đưa ra các quyết định bổ sung – cung cấp một chiếc la bàn để định hướng cho chúng ta thay vì một bản đồ chỉ ra đích đến của chúng ta. Cuối cùng, ngay cả khi nghiên cứu về máy dự đoán, câu ngạn ngữ cũ vẫn được áp dụng: “Thật khó để đưa ra dự đoán, đặc biệt là về tương lai”.

NHỮNG ĐIỂM CHÍNH

- Hai nhà kinh tế học đã xây dựng một AI có khả năng siêu phàm trong việc dự đoán thời điểm ai đó lên cơn đau tim. Nó rẻ hơn, nhanh hơn, và dường như ít mắc lỗi hơn so với bác sĩ bình thường về cả dương tính giả và âm tính giả. Máy dự đoán này có thể được triển khai như một giải pháp điểm, chỉ tác động đến một quyết định duy nhất: có thực hiện xét nghiệm hay không. Ứng dụng giải pháp điểm của AI này sẽ nâng cao năng suất của bệnh viện bằng cách phân bổ tốt hơn các xét nghiệm phát hiện cơn đau tim.
- Mặc dù giải pháp điểm có thể có tác động có ý nghĩa trong việc cải thiện dịch vụ chăm sóc sức khỏe thông qua phân bổ xét nghiệm tốt hơn, AI dự đoán cơn đau tim có độ chính xác cao có khả năng làm nền tảng cho một giải pháp cấp hệ thống có tác động lớn hơn nhiều. Sử dụng Canvas khám phá hệ thống AI, chúng tôi thấy rằng một trong những quyết định quan trọng là có nên thử nghiệm hay không và nó dựa trên dự đoán về một người bị đau tim. Nếu dự đoán đó trở nên đủ tốt và có thể được tạo bằng dữ liệu được thu thập dễ dàng – chẳng hạn như



bằng đồng hồ thông minh – thì có thể chuyển những dự đoán đó ra khỏi không gian phân loại trong khoa cấp cứu của bệnh viện đến nhà bệnh nhân. Nhiều bệnh nhân sẽ không bao giờ cần đến bệnh viện, sau khi được chẩn đoán mắc bệnh gì đó mà được sĩ hoặc bác sĩ chăm sóc chính có thể giúp điều trị tại nhà.

- Một thuộc tính quan trọng của Khung khám phá hệ thống AI là nó trừu tượng hóa tổ chức đối với các quyết định cốt lõi của nó. Bằng cách đó, nó giải phóng sứ mệnh của tổ chức, vốn vẫn cố định, khỏi vô số quy tắc và quyết định liên quan đến hiện trạng, vốn không cần thiết. Sau đó, các nhà thiết kế có quyền tự do tưởng tượng vô số giải pháp cấp hệ thống khác nhau có thể được kích hoạt bởi các máy dự đoán mạnh mẽ làm nền tảng cho các quyết định chính. AI dự đoán cơn đau tim duy nhất có thể cho phép không chỉ một mà một số giải pháp cấp hệ thống thay thế. Quá trình suy nghĩ bắt đầu bằng việc xác định các quyết định quan trọng, suy đoán xem điều gì có thể xảy ra nếu các dự đoán trở nên chính xác cao, sau đó hình dung lại các loại hệ thống có thể khai thác các dự đoán đó theo cách được tối ưu hóa để đạt được thành công trong nhiệm vụ.

PHẦN KẾT



XU HƯỚNG VÀ HỆ THỐNG AI

Nhìn lại, nó không chắc sẽ diễn ra tốt đẹp. Nhưng vào năm 2016, các nhà nghiên cứu của Microsoft đã phát hành một thuật toán AI có tên là Tay để học cách tương tác trên Twitter. Trong vòng vài giờ, nó đã học được và bắt đầu đăng những dòng tweet xúc phạm. Tay không đơn độc khi trở thành kẻ tồi tệ nhất trong chúng tôi. Những câu chuyện như thế này rất nhiều và khiến nhiều người, bao gồm cả các doanh nghiệp, miễn cưỡng áp dụng AI, không phải vì khả năng dự đoán của AI kém hơn con người. Thay vào đó, AI có thể quá giỏi trong việc hành xử giống như chúng.

Đây không phải là một bất ngờ. Dự đoán AI yêu cầu dữ liệu và đặc biệt đối với dữ liệu liên quan đến dự đoán điều gì đó về con người, dữ liệu đào tạo đến từ con người. Điều này có thể có giá trị, chẳng hạn như khi luyện tập để chơi một trò chơi với con người, nhưng con người không hoàn hảo và AI thừa hưởng những điểm không hoàn hảo đó.



Điều mà nhiều người không nhận ra là đây là một vấn đề hiện tại do cách chúng ta nghĩ về các giải pháp AI. Chẳng hạn, khi bạn quan tâm đến việc cho phép bộ phận nhân sự của mình sàng lọc hàng trăm ứng viên, thì cách sử dụng tiềm năng đầu tiên của AI là sử dụng một thuật toán thay vì con người cho công việc đó. Xét cho cùng, đó là một nhiệm vụ dự đoán – khả năng người này với những chứng chỉ này sẽ thành công trong lĩnh vực kinh doanh này là bao nhiêu? Nhưng cách sử dụng AI này là một giải pháp điểm, có thể hiệu quả, nhưng như chúng tôi đã nhấn mạnh trong suốt cuốn sách này, thường sẽ cần thiết kể lại toàn bộ cấp độ hệ thống. Khi loại bỏ những hậu quả bất lợi của sự thiên vị, cần phải có tư duy hệ thống.

Điểm xuất phát của chúng tôi là quan điểm của chúng tôi về sự thiên vị trong AI, điều này khác biệt và, chúng tôi dám nói, một quan điểm trái ngược về việc liệu AI có duy trì sự phân biệt đối xử hay không. Khi được xem bằng cách sử dụng tư duy hệ thống, các cơ hội dành cho AI liên quan đến sự thiên vị là rất hứa hẹn. Họ cung cấp một giải pháp cho nhiều khía cạnh của sự phân biệt đối xử. Và chính vì họ đưa ra điều này nên họ phải đối mặt với sự kháng cự. Sự thật khó chịu về sự phân biệt đối xử là việc loại bỏ nó sẽ tạo ra kẻ thắng người thua khi quyền lực thay đổi. Do đó, khả năng chống lại

việc áp dụng AI có thể cao hơn chính xác khi AI có khả năng tạo ra các hệ thống mới loại bỏ nhiều khía cạnh của sự phân biệt đối xử.

CƠ HỘI CHỐNG PHÂN BIỆT ĐỐI XỬ

Bởi vì AI tạo cơ hội để hiểu các nguồn gốc của sự thiên vị và với kiến thức này có thể được sử dụng một cách thích hợp trong quá trình ra quyết định, chúng mang lại cơ hội để giảm thiểu sự phân biệt đối xử.

Hãy xem xét một ví dụ đơn giản. Người da màu báo cáo tình trạng đau đầu gối cao hơn nhiều so với người da trắng. Có hai cách giải thích khác nhau cho việc này. Đầu tiên, những người da màu có thể bị viêm xương khớp nghiêm trọng hơn ở đầu gối. Ngoài ra, các yếu tố khác bên ngoài đầu gối, chẳng hạn như căng thẳng trong cuộc sống hoặc sự cô lập với xã hội, có thể khiến người da màu báo cáo mức độ đau đầu gối cao hơn.

Những giải thích này ngụ ý các phương pháp điều trị khác nhau. Nếu vấn đề là viêm xương khớp nghiêm trọng hơn, thì vật lý trị liệu, thuốc men và phẫu thuật có thể giúp ích. Nếu vấn đề nằm ngoài đầu gối, thì phương pháp điều trị hiệu quả nhất có thể tập trung vào việc cải thiện sức khỏe tâm thần.



Nhiều bác sĩ nghi ngờ rằng các yếu tố bên ngoài đầu gối quan trọng hơn trong việc giải thích sự khác biệt về chủng tộc. Các nghiên cứu đã so sánh cơn đau do bệnh nhân báo cáo với đánh giá của bác sĩ X-quang về viêm xương khớp gối dựa trên hình ảnh y tế. Các bác sĩ X-quang đánh giá dựa trên các phương pháp như phân loại Kellgren-Lawrence (KL), qua đó các bác sĩ kiểm tra hình ảnh đầu gối của bệnh nhân và cho điểm dựa trên sự hiện diện của gai xương, biến dạng xương và các yếu tố khác. Ngay cả sau khi điều chỉnh các đánh giá này, những người da màu cho biết mức độ đau đớn cao hơn.

Nhà khoa học máy tính Emma Pierson và các đồng tác giả của cô nghi ngờ vấn đề có thể nằm ở hệ thống phân loại. Các phương pháp đo viêm xương khớp, bao gồm phân loại KL, đã được phát triển từ nhiều thập kỷ trước ở người da trắng. Họ có thể bỏ sót những nguyên nhân thể chất gây ra đau đớn ở những người không phải là người da trắng. Các bác sĩ X-quang cũng có thể thiên vị khi đánh giá bệnh nhân không phải da trắng, hạ thấp nỗi đau của họ khi đưa ra chẩn đoán.

AI có thể giúp đỡ. Pierson và các đồng tác giả của cô ấy đã chụp hàng nghìn hình ảnh về đầu gối. Đối với mỗi hình ảnh, họ có mức độ đau tự báo cáo

của bệnh nhân. Khi bác sĩ chụp X-quang cho điểm các hình ảnh, chỉ có 9% sự khác biệt về chủng tộc về cơn đau dường như được giải thích bởi các yếu tố bên trong đầu gối.

Sau đó, các tác giả đã đánh giá liệu AI có thể sử dụng hình ảnh để dự đoán cơn đau được báo cáo hay không. Trí tuệ nhân tạo của họ đã dự đoán 43% sự khác biệt về chủng tộc trong nỗi đau. AI đã xác định các yếu tố bên trong đầu gối mà con người đã bỏ sót và những yếu tố này giải thích gần gấp 5 lần sự khác biệt về cơn đau được báo cáo giữa người da màu và người da trắng.

Sự khác biệt về chủng tộc trong điều trị cho thấy rằng nhiều bệnh nhân không phải là người da trắng sẽ được điều trị bên ngoài đầu gối khi rõ ràng có thứ gì đó ở đầu gối gây ra cơn đau. Tại đây, AI đã giúp xác định sự phân biệt đối xử có hệ thống trong chăm sóc sức khỏe và đưa ra cách khắc phục.

Để giải quyết vấn đề phân biệt đối xử, cả hai điều này đều cần thiết. Bạn cần phát hiện sự phân biệt đối xử và bạn cần khắc phục nó. Điều này đúng với cả dự đoán của con người và máy móc. Nói cách khác, xóa bỏ phân biệt đối xử cần có một hệ thống.



PHÁT HIỆN PHÂN BIỆT ĐỐI XỬ

Phát hiện sự phân biệt đối xử là khó khăn. Mặc dù có rất nhiều khiếu nại pháp lý chỉ trích sự phân biệt đối xử trong công nghệ và các ngành công nghiệp khác, nhưng rất ít quyết định có lợi cho các nguyên đơn. Nhiều trường hợp nổi bật nhất mà nhân viên bị kiện vì phân biệt đối xử cuối cùng đã được quyết định có lợi cho người sử dụng lao động hoặc bị loại bỏ.

Nhiều trường hợp trong số này tập trung vào việc liệu các công ty có phân biệt đối xử về tiền lương hoặc thăng chức hay không. Giả sử một công ty công nghệ bị buộc tội phân biệt giới tính trong các chương trình khuyến mãi. Rõ ràng là công ty đã thăng chức cho một số người thay vì nguyên đơn đã làm việc ở công ty lâu hơn, nhưng câu hỏi trọng tâm của vụ kiện sẽ là tại sao.

Nguyên đơn sẽ cho rằng công ty cố ý phân biệt đối xử với cô ấy. Công ty sẽ trả lời rằng nguyên đơn cụ thể “ít phải là nạn nhân của sự phân biệt đối xử hơn là một nhân viên khó tính và lười cá đã từ chối lời khuyên để cải thiện”, như tờ *New York Times* đã mô tả cách tiếp cận của một bị đơn. Các nhà quản lý sẽ được hỏi liệu họ có phân biệt đối xử trong các đề xuất thăng chức hay không. Tất nhiên, họ sẽ nói không. Luật sư của nguyên đơn

có thể hỏi thẳng thừng: “Liệu thân chủ của tôi có được thăng chức nếu cô ấy là đàn ông không?” “KHÔNG”. Luật sư của nguyên đơn sẽ cố gắng so sánh hiệu suất của những người đã được thăng cấp cho hiệu suất của nguyên đơn, nhưng hiệu suất rất khó đo lường và có quá nhiều sự mơ hồ trong so sánh. Ngay cả khi có sự phân biệt đối xử, rất khó để chứng minh. Không có hai người hoàn toàn giống nhau. Các nhà quản lý xem xét nhiều yếu tố khi đưa ra quyết định thăng chức và tuyển dụng. Nếu không có tuyên bố rõ ràng về ý định phân biệt đối xử, thăm phán hoặc bồi thẩm đoàn khó có thể tự tin rằng quyết định của một người là phân biệt đối xử. Không thể biết những gì thực sự trong tâm trí của một ai đó.

KHÔNG CÓ HAI NGƯỜI NÀO HOÀN TOÀN GIỐNG NHAU

Sendhil Mullainathan là chuyên gia phát hiện sự phân biệt đối xử. Năm 2001, chỉ 3 năm sau khi nhận bằng tiến sĩ, Mullainathan và đồng tác giả Marianne Bertrand bắt đầu đo lường sự phân biệt đối xử trong thị trường lao động Hoa Kỳ. Họ đã gửi những bản lý lịch hư cấu để quảng cáo truy nã trên các tờ báo ở Boston và Chicago. Đối với mỗi quảng cáo, họ đã gửi 4 bản lý lịch. Hai là chất lượng cao, và hai là chất lượng thấp. Họ chỉ định ngẫu nhiên



một trong những bản lý lịch chất lượng cao có tên nghe giống người Mỹ gốc Phi (Lakisha Washington hoặc Jamal Jones) và tên còn lại nghe có vẻ da trắng (Emily Walsh hoặc Greg Baker). Tương tự như vậy, họ chỉ định ngẫu nhiên một trong những bản lý lịch chất lượng thấp tên là người Mỹ gốc Phi và tên còn lại nghe có vẻ da trắng.

Sau đó, họ chờ xem liệu những ứng viên hư cấu của họ có được gọi phỏng vấn hay không. Tên màu trắng nhận được nhiều cuộc gọi lại hơn 50%. Khoảng cách giữa những bản lý lịch chất lượng cao với những cái tên nghe có vẻ da trắng và những bản lý lịch chất lượng cao với những cái tên nghe có vẻ giống người Mỹ gốc Phi thậm chí còn lớn hơn. Rõ ràng đã có sự phân biệt đối xử trong thị trường lao động.

15 năm sau, Mullainathan lại làm điều đó. Hiện là giáo sư Đại học Chicago và là người nhận giải thưởng “Genius Grant” của MacArthur, ông và các đồng tác giả đã phát hiện ra rằng một thuật toán được sử dụng rộng rãi để xác định những bệnh nhân có nhu cầu sức khỏe phức tạp đã bị phân biệt chủng tộc. Ở một mức độ rủi ro nhất định, bệnh nhân người Mỹ gốc Phi thực sự ốm yếu hơn đáng kể so với bệnh nhân da trắng. Việc khắc phục sự chênh lệch sẽ tăng gần gấp 3 tỷ lệ bệnh nhân người

Mỹ gốc Phi nhận được các nguồn lực bổ sung để quản lý việc chăm sóc của họ.

Sự thiên vị nảy sinh vì máy được thiết kế để dự đoán chi phí chăm sóc sức khỏe như một đại diện cho bệnh tật, thay vì bản thân căn bệnh đó. Khả năng tiếp cận dịch vụ chăm sóc không bình đẳng có nghĩa là hệ thống chăm sóc sức khỏe của Hoa Kỳ dành ít tiền hơn cho việc chăm sóc bệnh nhân người Mỹ gốc Phi hơn bệnh nhân da trắng. Vì vậy, một cỗ máy dự đoán sử dụng chi tiêu chăm sóc sức khỏe như một đại diện cho bệnh tật sẽ đánh giá thấp mức độ nghiêm trọng của bệnh tật ở người Mỹ gốc Phi và các nhóm bệnh nhân khác ít được chăm sóc hơn.

Sau khi nghiên cứu xong, Mullainathan đã phản ánh về hai dự án. Cả hai nghiên cứu đều ghi nhận sự bất công về chủng tộc: Trong nghiên cứu đầu tiên, ứng viên có tên nghe có vẻ Da đen nhận được ít cuộc phỏng vấn xin việc hơn. Trong lần thứ hai, bệnh nhân Da đen được chăm sóc tệ hơn.

Nhưng họ khác nhau ở một khía cạnh quan trọng. Đầu tiên, các nhà quản lý tuyển dụng đã đưa ra những quyết định thiên vị. Trong trường hợp thứ hai, thủ phạm là một chương trình máy tính.

Là đồng tác giả của cả hai nghiên cứu, tôi xem chúng như một bài học về sự tương phản. Song song



với nhau, chúng cho thấy sự khác biệt rõ rệt giữa hai loại thành kiến: con người và thuật toán.

Nghiên cứu trước đó đòi hỏi sự sáng tạo và nỗ lực phi thường để phát hiện sự phân biệt đối xử. Ông mô tả nó như một “hoạt động bí mật phức tạp” đã diễn ra trong nhiều tháng.

Ngược lại, nghiên cứu sau này đơn giản hơn. Mullainathan mô tả nó là “một bài tập thống kê – tương đương với việc hỏi thuật toán ‘bạn sẽ làm gì với bệnh nhân này?’ hàng trăm nghìn lần và vạch ra sự khác biệt về chủng tộc. Công việc mang tính kỹ thuật và thuộc lòng, không đòi hỏi sự lén lút cũng như sự tháo vát”.

Đo lường sự phân biệt đối xử ở mọi người rất khó, đòi hỏi phải kiểm soát cẩn thận các tình huống. Đo lường sự phân biệt đối xử bằng máy đơn giản hơn. Cung cấp cho máy đúng dữ liệu và xem kết quả ra sao. Nhà nghiên cứu có thể đến gặp AI và nói, Nếu người đó như thế này thì sao? Nếu người như vậy thì sao? Có thể thử hàng nghìn giả định. Điều đó là không thể với con người. Mullainathan lưu ý: “Con người khó hiểu theo cách mà các thuật toán không thể hiểu được.

SỬA CHỮA PHÂN BIỆT ĐỐI XỬ

Phát hiện sự phân biệt đối xử chỉ là bước đầu tiên. Một khi phân biệt đối xử được phát hiện, bạn muốn sửa chữa nó. Con người khó sửa. Vì vậy, bạn cần một hệ thống cái đó không dựa vào cái đó.

Trong nghiên cứu sơ yếu lý lịch, ngay cả khi bạn có thể vượt qua thử thách tìm ra công ty nào có lỗi, thì “thay đổi trái tim và khối óc của mọi người không phải là vấn đề đơn giản”. Bằng chứng về các công cụ như đào tạo thiên vị ngầm là hỗn hợp. Chúng tôi không biết có giải pháp nào có thể giảm thiểu sự phân biệt đối xử của hàng nghìn, thậm chí hàng triệu người mỗi ngày. Hai thập kỷ sau nghiên cứu ban đầu đó, Emily và Greg vẫn có nhiều việc làm hơn Lakisha và Jamal.

Ngược lại với AI. Ngay cả trước khi họ công bố nghiên cứu về phân biệt đối xử bằng thuật toán, Mullainathan và các đồng tác giả của ông đã làm việc với công ty để khắc phục sự cố. Họ bắt đầu bằng cách liên hệ với công ty, công ty có thể sao chép kết quả của nghiên cứu bằng các mô phỏng của chính họ. Ở bước đầu tiên, nó cho thấy rằng việc bao gồm dự đoán sức khỏe với dự đoán chi phí hiện tại sẽ giảm 84% sai lệch. Các tác giả đã cung cấp dịch vụ của họ miễn phí cho một số hệ thống chăm sóc sức khỏe bằng cách



sử dụng các loại thuật toán này. Nhiều người đã chấp nhận lời đề nghị.

Bài viết học thuật của Obermeyer và các đồng nghiệp kết luận rằng “các sai lệch nhân có thể sửa chữa được... Bởi vì các nhân là yếu tố chính quyết định cả chất lượng dự đoán và sai lệch dự đoán, nên việc lựa chọn cẩn thận có thể cho phép chúng ta tận hưởng lợi ích của các dự đoán bằng thuật toán đồng thời giảm thiểu rủi ro của chúng”. Như Mullainathan đã nói, “[C]treo các thuật toán dễ hơn thay đổi con người: phần mềm trên máy tính có thể được cập nhật; ‘đồ ươi’ trong não của chúng ta cho đến nay đã được chứng minh là kém mềm dẻo hơn nhiều.

BÊN TRONG HỘP AI

AI dễ bị sai lệch. Điều này có thể có nghĩa là các nhóm yếu thế bị đối xử tệ hơn những nhóm khác. Nói cách khác, AI có thể là một nguồn phân biệt đối xử.

AI cũng có thể làm giảm sự phân biệt đối xử. Họ có thể phát hiện các hành vi phân biệt đối xử ở người, như trong việc phát hiện đau đầu gối. Họ cũng có thể xem xét kỹ lưỡng. Việc phát hiện sự phân biệt đối xử bằng thuật toán dễ dàng hơn nhiều so với sự phân biệt đối xử của con người.

Phân biệt đối xử AI cũng có thể sửa chữa được. Phần mềm có thể được điều chỉnh và các nguồn sai lệch đã xác định có thể được loại bỏ.

Khắc phục sự phân biệt đối xử như vậy là không dễ dàng. Đầu tiên, nó đòi hỏi con người muốn sửa chữa thành kiến. Nếu những người quản lý AI muốn triển khai một AI có khả năng phân biệt đối xử, họ sẽ gặp ít khó khăn khi thực hiện điều đó. Và bởi vì AI là phần mềm nên sự phân biệt đối xử của nó có thể xảy ra trên quy mô lớn. Tuy nhiên, việc bắt một AI cố ý phân biệt đối xử sẽ dễ dàng hơn một con người cố tình phân biệt đối xử. AI để lại dấu vết kiểm toán. Một cơ quan quản lý được tài trợ tốt với các kiểm toán viên được đào tạo bài bản có thể tiếp cận AI có thể chạy các mô phỏng để tìm kiếm sự phân biệt đối xử, giống như Mullainathan và các đồng tác giả của ông đã làm. Thật không may, các hệ thống pháp luật và quy định hiện tại của chúng ta phải vật lộn với những thách thức này, vì chúng được thiết kế cho thế giới của những người ra quyết định là con người, không có sự trợ giúp của các thuật toán.

Thứ hai, ngay cả khi được triển khai bởi những người có thiện chí muốn giảm bớt thành kiến, chi tiết vẫn quan trọng. Và tập trung vào các chi tiết là tốn thời gian và tốn kém. Có nhiều cách sai lệch có



thể xâm nhập vào các dự đoán của AI. Khắc phục sự thiên vị đòi hỏi phải hiểu nguồn gốc của nó. Điều này đòi hỏi đầu tư vào việc lưu trữ dữ liệu về các quyết định trong quá khứ. Nó đòi hỏi các khoản đầu tư vào việc mô phỏng các nguồn sai lệch tiềm ẩn để xem AI duy trì như thế nào. Và nỗ lực đầu tiên có thể không hoạt động. Dữ liệu mới có thể cần phải được thu thập. Các quy trình mới có thể được yêu cầu.

Thứ ba, một AI làm giảm sự thiên vị có thể thay đổi người nắm quyền ra quyết định trong một tổ chức. Nếu không có AI, có thể các nhà quản lý cá nhân sẽ đưa ra quyết định thuê ai. Những nhà quản lý này, ngay cả với ý định tốt nhất, có thể tuyển dụng thông qua các mối quan hệ xã hội của họ theo cách dẫn đến sự thiên vị ngoài ý muốn. Với AI được thiết kế để giảm bớt sự thiên vị, việc tuyển dụng thông qua các kết nối xã hội sẽ khó khăn hơn. Một giám đốc điều hành cấp cao hơn sẽ đặt ra ngưỡng để xem xét sơ yếu lý lịch. Giám đốc điều hành đó có thể nhận ra rằng nếu tất cả các nhà quản lý trong toàn công ty được tuyển dụng thông qua các mối quan hệ xã hội của họ, thì một lực lượng lao động đa dạng sẽ là điều không thể. AI làm giảm sự phân biệt đối xử, nhưng nó cũng làm giảm quyền quyết định của các nhà quản lý cá nhân trong việc tuyển dụng so với các mục tiêu do

bộ điều hành đặt ra. Vì vậy, những nhà quản lý đó có thể chống lại sự thay đổi ở cấp độ hệ thống sẽ làm giảm quyền lực của họ.

Không phải ai cũng sẽ ăn mừng vì đã giảm được thành kiến. Năm 2003, Giải bóng chày Major League đã sử dụng một công cụ mới để xác định vị trí của các cú ném bóng trên đĩa có tên là Hệ thống thông tin trọng tài QuesTec. QuesTec đã đánh giá các quả bóng và các cú đánh do trọng tài gọi. Không có gì ngạc nhiên khi các trọng tài chống lại công cụ này. Một số cầu thủ ngôi sao cũng vậy. Sandy Alderson, khi đó là phó chủ tịch của hoạt động, đã mô tả một động lực cho công cụ, tuyên bố rằng một số người chơi kỳ cựu sẽ nhận được lợi ích từ sự nghi ngờ và có những quả bóng và cú đánh có lợi cho họ. Nhiều ngôi sao lớn nhất của trò chơi đã phản nản, bao gồm cả vận động viên ném bóng từng đoạt giải thưởng Tom Glavine và nhiều người chiến thắng MVP Barry Bonds. Sau đó – vận động viên ném bóng xuất sắc của Arizona Diamondbacks Curt Schilling trở thành một Luddite thời hiện đại, đập vỡ máy ảnh sau một trận thua. Một công cụ tự động trong đó máy tính dự đoán các quả bóng và cú đánh có thể làm giảm sự thiên vị, nhưng những người hưởng lợi từ sự thiên vị có thể không thích điều đó.



CẦN MỘT HỆ THỐNG

Amazon thuê rất nhiều người. Cứ 153 công nhân ở Hoa Kỳ thì có một người là nhân viên của Amazon. Vì vậy, bạn sẽ không ngạc nhiên khi họ rất quan tâm đến việc phát triển AI để hỗ trợ tuyển dụng. Vào năm 2014, nó đã làm được điều đó. Chỉ một năm sau, nó đã loại bỏ hệ thống và không bao giờ đưa vào lĩnh vực này. Tại sao? Bởi vì nó không đánh giá các ứng viên cho phần mềm và các công việc kỹ thuật khác theo cách trung lập về giới tính. Lý do là một lý do quen thuộc. AI của Amazon được đào tạo dựa trên dữ liệu quá khứ của các ứng viên nam. AI rõ ràng đã giảm bớt các tham chiếu đến phụ nữ, bao gồm cả các trường đại học dành cho nữ. Các chỉnh sửa đơn giản không thể khôi phục tính trung lập.

Bạn đọc những câu chuyện như thế này và nghĩ rằng AI thiên vị một cách vô vọng. Nhưng theo một cách khác mà bạn có thể đọc câu chuyện này là AI đã bị thiên vị, bị đánh giá là sai và do đó đã không được triển khai. Điều tương tự cũng có thể xảy ra đối với các nhà tuyển dụng con người? Chúng tôi thực sự biết câu trả lời: AI đã được đào tạo về những nhà tuyển dụng đó ngay từ đầu.

Đồng thời, kinh nghiệm này đã dạy cho các nhà phát triển AI rằng việc đào tạo dựa trên dữ

liệu trong quá khứ thường không đủ tốt. Họ cần những nguồn dữ liệu mới, sẽ mất thời gian để phát triển. Nhưng cuối cùng, AI có thể được đánh giá. Hơn nữa, nó có thể được theo dõi liên tục về hiệu suất.

Đây là một cái tiến tiềm năng sâu sắc so với cách chúng ta đối phó với sự phân biệt đối xử hiện nay. Các biện pháp can thiệp ngày nay để giảm bớt sự phân biệt đối xử chủ yếu dựa trên kết quả: Có sự khác biệt giữa kết quả của các nhóm khác nhau không? Và các biện pháp can thiệp thường là các quy tắc trực tiếp để cố gắng khắc phục sự cân bằng và đạt được kết quả tương đương. Vấn đề là những can thiệp đó có thể gây chia rẽ.

Ngược lại, điều mọi người thường muốn là loại bỏ nguồn gốc của thành kiến – đặc biệt là động cơ của những người đang đưa ra quyết định. Họ không muốn tự khắc phục các kết quả bình đẳng (mặc dù đạt được điều đó không phải là vấn đề), nhưng thay vào đó họ muốn được đối xử bình đẳng. Tuy nhiên, khi mọi người đưa ra quyết định và chúng ta không thể nhìn thấy động lực của họ, thì làm sao chúng ta có thể tin tưởng rằng có sự đối xử bình đẳng?

Nếu dự đoán AI có thể là trọng tâm của những quyết định như vậy, thì chúng ta có thể đạt được



một tiêu chuẩn khách quan. Chúng ta có thể thấy cách AI đối xử với mọi người và vì chúng ta biết AI không có động cơ rõ ràng để đối xử khác biệt với những người giống nhau, nên chúng ta có thể nỗ lực để đảm bảo AI thực sự làm được điều đó.

Dự đoán tự động giúp tạo tiêu chuẩn dễ dàng hơn. Giống như tất cả các cầu thủ bóng chày có thể đối mặt với cùng một khu vực tấn công, tất cả các tài xế đều có thể đối mặt với các tiêu chuẩn thực thi giao thông giống nhau. Có những thành kiến được ghi chép rõ ràng trong việc thực thi giao thông; ví dụ, người lái xe da đen bị vượt nhiều hơn người da trắng. Một giải pháp dễ dàng là tự động hóa các vé tăng tốc. Chúng tôi có công nghệ cho việc này. Phát hiện tốc độ, chụp ảnh và sau đó xử phạt các tài xế. Một hệ thống tự động công bằng hơn và an toàn hơn, giảm khả năng xảy ra đụng độ bạo lực giữa cảnh sát và công chúng.

Nhưng lợi ích của việc làm đó vượt xa giải pháp điểm. Việc tin tưởng rằng mọi người đang được đối xử bình đẳng sẽ thay đổi cách mọi người tương tác với hệ thống và mức độ an toàn mà họ cảm thấy khi hành động trong đó. Nó cũng loại bỏ nhu cầu can thiệp chỉ đơn giản là cố gắng trông đẹp mắt trên sổ sách, chẳng hạn như đặt kết quả là một phần của các nhóm cố định hoặc thậm chí là hạn ngạch.

Nhưng tiến tới loại bỏ sự đối xử khác biệt sẽ không liền mạch; điều này chính xác là vì nó sẽ thay đổi kết quả trong hệ thống. Không phải ai cũng hoan nghênh một hệ thống tự động. Cũng giống như các cầu thủ bóng chày ngôi sao, những người lái xe nhận được lợi ích từ sự tùy ý của cảnh sát có thể bức bối với máy quay. Một số người sẽ đấu tranh để trả tiền phạt. Hơn nữa, một hệ thống tự động sẽ không khoan dung nếu ai đó tăng tốc vì những lý do chính đáng như cấp cứu y tế. Cảnh sát không thể sử dụng các điểm dừng để phát hiện tội phạm ngoài tốc độ.

Tuy nhiên, tốc độ giết chết. Để người lái xe ở dưới tốc độ giới hạn sẽ cứu được mạng sống. Nhưng việc thực thi không đồng đều và thường có sự phân biệt đối xử. Thực thi tự động sẽ bắt được nhiều tài xế hơn và giảm phân biệt đối xử.

Khi AI được triển khai dưới dạng giải pháp điểm, chúng có thể khuếch đại những thành kiến hiện có và tăng khả năng phân biệt đối xử. Điều này đã dẫn đến những tiêu đề tiêu cực về AI và sự phân biệt đối xử mà chúng ta thấy trên báo chí. Khi AI được coi là một điểm giải pháp, sự thiên vị của AI là một vấn đề và có thể tạo ra sự phản kháng thích hợp đối với việc áp dụng các máy dự đoán.



Khi sự thiên vị được nhìn từ tư duy hệ thống, AI có thể dẫn đến thay đổi làm giảm sự phân biệt đối xử. Trong khi những người được hưởng lợi từ những thành kiến hiện có sẽ chống lại, thì vẫn có lý do để lạc quan. Giống như Mullainathan, chúng tôi nhận thấy tiềm năng của AI trong việc giảm bớt sự thiên vị trong mọi loại quyết định. Sự lạc quan về AI này che đậy một sự bi quan rộng lớn hơn về việc ra quyết định của con người. Cả con người và AI sẽ bị sai lệch. Như nhà khoa học máy tính Marzyeh Ghassemi của MIT đã nói sau một bài giảng về những thành kiến trong máy học trong lĩnh vực chăm sóc sức khỏe, khi nói đến thành kiến, “Con người thật tồi tệ”. Xu hướng AI có thể được phát hiện và sau đó giải quyết. Các giải pháp hệ thống AI mới trên tất cả các lĩnh vực, từ giáo dục đến chăm sóc sức khỏe và từ ngân hàng đến cảnh sát, có thể được thiết kế và triển khai để giảm phân biệt đối xử. Các hệ thống AI có thể được giám sát liên tục và hồi tố để đảm bảo thành công liên tục trong việc xóa bỏ phân biệt đối xử. Giá như nó dễ dàng sửa chữa ở người.

NHỮNG ĐIỂM CHÍNH

- Câu chuyện phổ biến xung quanh AI là chúng học được thành kiến của con người và khuếch đại nó. Chúng tôi đồng ý và ủng hộ sự cảnh giác liên tục. Một câu chuyện khác là hệ thống AI không nên được đưa vào các quyết định quan trọng như tuyển dụng, vay ngân hàng, yêu cầu bảo hiểm, phán quyết pháp lý và tuyển sinh đại học vì chúng không rõ ràng – chúng ta không thể nhìn thấy bên trong hộp đen – và chúng kéo dài sự phân biệt đối xử. Chúng tôi không đồng ý. Chúng tôi nghĩ rằng chúng nên được giới thiệu chính xác trong các quyết định quan trọng bởi vì chúng có thể xem xét kỹ lưỡng theo cách mà con người không làm được. Chúng tôi không thể thẩm vấn hiệu quả một người quản lý tuyển dụng con người mà chúng tôi nghi ngờ có hành vi phân biệt đối xử với hàng nghìn câu hỏi như “Bạn có thuê người này không nếu họ giống hệt nhau, ngoại trừ người da trắng?” và mong đợi một phản ứng trung thực. Tuy nhiên, chúng ta có thể hỏi chính xác câu hỏi đó và hàng nghìn câu hỏi khác về hệ thống AI và nhận được câu trả lời nhanh chóng và chính xác.



- Giáo sư Sendhil Mullainathan của Đại học Chicago đã đối chiếu hai nghiên cứu của ông về thành kiến. Trong một lần, ông đo lường sự phân biệt đối xử của con người trong việc tuyển dụng. Mặt khác, ông ấy đo lường sự phân biệt đối xử của AI trong chăm sóc sức khỏe.

Sợ sánh cả hai, ông ấy lưu ý rằng việc phát hiện và khắc phục sự phân biệt đối xử trong hệ thống AI dễ dàng hơn nhiều so với hệ thống của con người: “[C]ần thuật toán dễ hơn thay đổi con người: phần mềm trên máy tính có thể được cập nhật; ‘đồ ւớ’ trong não của chúng ta cho đến nay đã được chứng minh là kém mềm dẻo hơn nhiều.

- Ngày nay, những cá nhân phản đối việc áp dụng các hệ thống AI nhất lại là những người lo ngại nhất về sự phân biệt đối xử. Chúng tôi dự đoán điều đó sẽ đảo ngược chính xác. Một khi mọi người nhận ra rằng sự phân biệt đối xử dễ phát hiện và khắc phục trong các hệ thống AI hơn là ở con người, thì sự phản kháng lớn nhất đối với việc áp dụng các hệ thống AI sẽ không đến từ những người muốn giảm thiểu sự phân biệt đối xử mà là từ những người được hưởng lợi nhiều nhất từ nó.

VỀ TÁC GIẢ

AJAY AGRAWAL là giáo sư về quản lý chiến lược và là Chủ tịch Geoffrey Taber về Khởi nghiệp và Đổi mới tại Trường Quản lý Rotman của Đại học Toronto. Ông là người sáng lập Phòng thí nghiệm hủy diệt sáng tạo và đồng sáng lập NEXT Canada, cả hai chương trình phi lợi nhuận đều hỗ trợ thương mại hóa khoa học thông qua tinh thần kinh doanh.

Ông là cộng tác viên nghiên cứu tại Cục Nghiên cứu Kinh tế Quốc gia, ở Cambridge, Massachusetts; một thành viên hội đồng cố vấn học thuật tại Trung tâm Quy định và Thị trường tại Brookings, ở Washington, DC; một thành viên ban cố vấn tại Trung tâm Công nghệ và Xã hội Khốĩ Đại học Carnegie Mellon, ở Pittsburgh; và một khoa liên kết tại Viện Trí tuệ nhân tạo Vector, ở Toronto. Ajay tiến hành nghiên cứu về kinh tế đổi mới và phục vụ trong ban biên tập của Khoa học Quản lý. Ông là người đồng sáng lập công ty AI/robot Sanctuary. Nhiệm vụ của công ty là tạo ra trí thông minh giống con người đầu tiên trên thế giới trong các robot hình người có mục đích chung. Ajay có bằng tiến sĩ về



kinh tế học ứng dụng tại Đại học British Columbia. Ông được bổ nhiệm làm Thành viên của Order of Canada vào năm 2022.

JOSHUA GANS là giáo sư về quản lý chiến lược và là người giữ chức Chủ tịch Jeffrey S. Skoll về Đổi mới Kỹ thuật và Tinh thần Doanh nhân tại Trường Quản lý Rotman, Đại học Toronto. Joshua cũng là nhà kinh tế trưởng tại Phòng thí nghiệm Phá hủy Sáng tạo của Đại học Toronto. Joshua có hơn 150 ấn phẩm học thuật được bình duyệt và là biên tập viên (chiến lược) của Khoa học Quản lý. Ông là cộng tác viên nghiên cứu tại Cục Nghiên cứu Kinh tế Quốc gia và có học bổng tại MIT, Viện e61, Học viện Luohan, Trung tâm Phân tích Kinh tế Quốc tế, Trường Kinh doanh Melbourne, Hiệp hội Tăng tốc và Học viện Khoa học Xã hội, Úc. Ông là tác giả của hai cuốn sách giáo khoa thành công và viết 10 cuốn sách nổi tiếng, bao gồm *Parentonomics* (2009), *Information Wants to Be Shared* (2012), *The Disruption Dilemma* (2016), *Scholarly Publishing and Its Discontents* (2017), *Innovation + Equality* (2019), *The Lỗ hổng thông tin về đại dịch* (2020) và *Giải pháp thông tin về đại dịch* (2021). Joshua có bằng tiến sĩ kinh tế tại Đại học Stanford và vào năm 2008, anh đã được trao Giải thưởng Nhà kinh tế trẻ của Hiệp hội Kinh tế Úc (tương đương với huy chương John Bates Clark của Úc).

AVI GOLDFARB là Chủ tịch Rotman về Trí tuệ nhân tạo và Chăm sóc sức khỏe, đồng thời là giáo sư tiếp thị tại Trường Quản lý Rotman, Đại học Toronto. Avi cũng là nhà khoa học dữ liệu chính tại Phòng thí nghiệm hủy diệt sáng tạo, một khoa liên kết tại Viện Vector và tại Viện Công nghệ và Xã hội Schwartz Reisman, thành viên ban cố vấn khoa học của HiParis!, cộng tác viên nghiên cứu tại Cục nghiên cứu kinh tế quốc gia, và là cựu biên tập viên cấp cao tại Khoa học Tiếp thị. Nghiên cứu của Avi tập trung vào các cơ hội và thách thức của nền kinh tế kỹ thuật số. Ông đã xuất bản các bài báo học thuật về tiếp thị, máy tính, luật, quản lý, y học, vật lý, khoa học chính trị, y tế công cộng, thống kê và kinh tế. Công trình của ông về quảng cáo trực tuyến đã giành được Giải thưởng Tác động Dài hạn của Hiệp hội Khoa học Tiếp thị INFORMS. Avi đã làm chứng trước Ủy ban Tư pháp Thượng viện Hoa Kỳ về cạnh tranh và quyền riêng tư trong quảng cáo kỹ thuật số. Ông có bằng tiến sĩ kinh tế tại Đại học Northwestern.

NHÀ XUẤT BẢN CÔNG THƯƠNG

Trụ sở: Số 655 Phạm Văn Đồng, Bắc Từ Liêm, Hà Nội

Điện thoại: 024 3934 1562 - Fax: 024 3938 7164

Website: <http://nxbconghuong.vn>

Email: nxbct@moit.gov.vn

AI 5.0 – NHANH HƠN, DỄ HƠN, RẺ HƠN, CHÍNH XÁC HƠN

Chịu trách nhiệm xuất bản:

Giám đốc - Tổng Biên tập

TRƯƠNG THU HIỀN

Biên tập	Lương Thị Ngọc Bích
Vẽ bìa	Mạnh Cường
Trình bày	Minh Ngọc
Sửa bản in	Lê Hạnh

In 2000 bản, khổ 13x20.5cm tại Công ty Cổ phần In và Thương mại Prima.
Địa chỉ: Số 35, ngõ 93 Hoàng Quốc Việt, Cầu Giấy, Hà Nội. Số xác nhận
ĐKXB: 295-2024/CXBIPH/49-20/CT. Quyết định xuất bản số 100/QĐ-NXBCT
cấp ngày 19 tháng 04 năm 2024. Mã ISBN: 978-604-481-288-5. In xong và
nộp lưu chiếu Quý II năm 2024

ĐỐI TÁC LIÊN KẾT:
CÔNG TY TNHH VĂN HÓA VÀ TRUYỀN THÔNG 1980 BOOKS

Trụ sở chính tại Hà Nội:

20-H2, ngõ 6 Trần Kim Xuyến, Yên Hòa, Cầu Giấy, Hà Nội.

Tel: 0243.7880225 / Fax: 0243.7880225

Chi nhánh tại TP. Hồ Chí Minh:

Số 42/35 đường Nguyễn Minh Hoàng, phường 12, Q.Tân Bình, Tp HCM

Tel: 0283.933.3216

Website/link đặt sách: www.1980books.com

Email: rights.1980books@gmail.com

1980
BOOKS
KHUẤT MÀU TAY TRẮC

KỸ
NĂNG

AI 5.0

NHANH HƠN, DỄ HƠN,
RẺ HƠN, CHÍNH XÁC HƠN

Sự trỗi dậy của AI đang và sẽ tác động đến nhiều lĩnh vực trên thế giới – giáo dục, nghệ thuật, tài chính ngân hàng, dược phẩm, ô tô, công nghệ y tế, sản xuất và bán lẻ. Nhưng nó chỉ mới bắt đầu cuộc phiêu lưu hướng tới những dự đoán rẻ hơn, dễ dàng hơn, tốt hơn, nhanh hơn và chính xác hơn nhằm thúc đẩy các quyết định kinh doanh chiến lược.

Nhiều ngành đang trong tình trạng yên tĩnh trước cơn bão Những ai nhanh nhạy nắm bắt AI sẽ là những người nắm trong tay quyền lực, ngược lại, có thể bị chính AI nhấn chìm.

Trong kinh tế, quyền lực mang lại lợi nhuận.

Trong xã hội, quyền lực mang lại sự kiểm soát.

Công ty TNHH Văn Hóa và Truyền Thông 1980 Books
Trụ sở chính:

20-H2, ngò 6 Trần Kim Xuyên, Yên Hòa, Cầu Giấy, HN
Tel: (024) 3 788 0225

Chi nhánh tại TP.HCM:

42/35 Đường Nguyễn Minh Hoàng, P. 12, Tân Bình, HCM
Tel: (028) 3 933 3216

Đặt mua sách:

www.1980books.com

www.facebook.com/1980BOOKS.COM.VN

AI 5.0 - NHANH HƠN, DỄ HƠN, RẺ HƠN, CHÍNH XÁC HƠN

ISBN: 978-604-481-288-5



9 786044 812885



8 936066 696050

Giá: 209.000đ