 Bản dịch cuốn sách "Artificial Intelligence For Dummies" sang tiếng Việt

# Giới thiệu

Bạn hầu như không thể tránh khỏi việc nghe nói về trí tuệ nhân tạo (AI) ngày nay. Bạn thấy AI trong phim ảnh, sách vở, các bản tin, và trên mạng. AI là một phần của robot, ô tô tự lái, drone, hệ thống y tế, trang web mua sắm trực tuyến và nhiều công nghệ khác ảnh hưởng đến cuộc sống hàng ngày của bạn. Một số người thậm chí đã tin tưởng AI đến mức ngủ thiếp đi trong khi ô tô tự lái đưa họ đến đích — điều này, dĩ nhiên, là bất hợp pháp.

Nhiều chuyên gia đang lấp đầy bạn với thông tin (và cả thông tin sai lệch) về AI. Một số coi AI là dễ thương và thân thiện; trong khi những người khác lại xem nó như một nguy cơ đe dọa nhân loại. Vấn đề ở đây là, khi bạn bị quá tải thông tin, bạn sẽ khó phân biệt điều gì là thật và điều gì chỉ là sản phẩm của trí tưởng tượng quá đà. Vậy bạn có thể tin tưởng AI đến mức nào? Phần lớn những lời thổi phồng về AI xuất phát từ kỳ vọng không thực tế của các nhà khoa học, doanh nhân và những người kinh doanh.

Artificial Intelligence For Dummies, 2nd Edition là cuốn sách bạn cần nếu bạn cảm thấy rằng mình thực sự không biết gì về công nghệ được cho là một phần không thể thiếu trong cuộc sống của mình. Cuốn sách này giúp bạn nhận ra những ứng dụng thực tế và cần thiết của AI trong cuộc sống hàng ngày, từ việc sử dụng bộ điều chỉnh nhiệt thông minh trong nhà đến những ứng dụng y tế tiên tiến.

Ngoài ra, cuốn sách còn chia sẻ những sự thật về những gì AI không thể làm. AI sẽ không bao giờ có thể thực hiện một số hoạt động nhất định cho đến khi công nghệ phát triển vượt bậc trong tương lai. Điều này giúp xua tan những ảo tưởng khiến bạn bị mù mờ về AI.

Về cuốn sách này Cuốn sách bắt đầu bằng cách giúp bạn hiểu rõ AI là gì, những gì nó cần để hoạt động, và tại sao nó từng thất bại trong quá khứ. Bạn sẽ khám phá cách AI liên quan đến các lĩnh vực như phân tích dữ liệu, học máy, và học sâu. Cuốn sách cũng cung cấp các đường dẫn đến thông tin bên ngoài để bạn có cái nhìn toàn diện hơn.

Ngoài ra, bạn còn được hướng dẫn cách hiểu các khái niệm chính về AI thông qua các biểu tượng và ghi chú, giúp việc học tập trở nên dễ dàng hơn.

# Mục lục

**Trang tiêu đề**  
**Lời giới thiệu**

**Phần 1: Giới thiệu về AI**

**Chương 1: Giới thiệu về AI**

* Định nghĩa thuật ngữ AI
* Tìm hiểu lịch sử AI
* Xem xét các ứng dụng của AI
* Tránh sự cường điệu hóa và đánh giá quá cao AI
* Mối liên kết giữa AI và công nghệ máy tính

**Chương 2: Định nghĩa vai trò của dữ liệu**

* Tìm kiếm dữ liệu như một tài nguyên phổ quát trong thời đại này
* Sử dụng dữ liệu thành công
* Chăm sóc và làm sạch dữ liệu
* Cân nhắc năm sự thật sai lệch trong dữ liệu
* Xác định giới hạn của việc thu thập dữ liệu
* Xem xét các vấn đề bảo mật dữ liệu

**Chương 3: Cân nhắc việc sử dụng thuật toán**

* Hiểu vai trò của thuật toán
* Khám phá cỗ máy học hỏi

**Chương 4: Đi đầu trong việc phát triển phần cứng chuyên dụng**

* Dựa vào phần cứng tiêu chuẩn
* Sử dụng GPU
* Làm việc với bộ xử lý học sâu (DLP)
* Tạo môi trường xử lý chuyên dụng
* Tăng cường khả năng của phần cứng
* Thêm cảm biến chuyên dụng
* Xây dựng phương pháp tương tác với môi trường

**Phần 2: Cân nhắc các ứng dụng AI trong xã hội**

**Chương 5: Nhìn nhận các ứng dụng AI trong máy tính**

* Giới thiệu các loại ứng dụng phổ biến
* Cách AI làm cho ứng dụng trở nên thân thiện hơn
* Tự động sửa lỗi
* Đưa ra gợi ý
* Cân nhắc các lỗi dựa trên AI

**Chương 6: Tự động hóa các quy trình thông thường**

* Phát triển giải pháp cho sự nhàm chán
* Làm việc trong môi trường công nghiệp
* Tạo ra môi trường an toàn

**Chương 7: Sử dụng AI để đáp ứng nhu cầu y tế**

* Áp dụng hệ thống giám sát bệnh nhân di động
* Làm cho con người trở nên hiệu quả hơn
* Giải quyết một loạt khả năng thể chất
* Hoàn thiện phân tích theo những cách mới
* Dựa vào công nghệ từ xa
* Tạo ra các kỹ thuật phẫu thuật mới
* Thực hiện nhiệm vụ bằng cách tự động hóa
* Kết hợp robot và chuyên gia y tế

**Chương 8: Dựa vào AI để cải thiện tương tác giữa con người**

* Phát triển các cách giao tiếp mới
* Trao đổi ý tưởng
* Cải thiện khả năng cảm nhận của con người

**Phần 3: Làm việc với các ứng dụng AI dựa trên phần mềm**

**Chương 9: Thực hiện phân tích dữ liệu cho AI**

* Định nghĩa phân tích dữ liệu
* Định nghĩa học máy
* Xem xét cách học hỏi từ dữ liệu

**Chương 10: Ứng dụng học máy trong AI**

* Đi theo nhiều con đường khác nhau để học hỏi
* Khám phá sự thật trong xác suất
* Phát triển các cây phân loại

**Chương 11: Cải thiện AI với học sâu**

* Tạo hình mạng nơ-ron tương tự não bộ con người
* Mô phỏng bộ não học hỏi
* Giới thiệu về học sâu
* Nhận diện các cạnh và hình dạng từ hình ảnh
* Học cách bắt chước nghệ thuật và cuộc sống

**Phần 4: Làm việc với AI trong các ứng dụng phần cứng**

**Chương 12: Phát triển robot**

* Định nghĩa vai trò của robot
* Lắp ráp một robot cơ bản

**Chương 13: Bay cùng drone**

* Ghi nhận trạng thái hiện đại
* Định nghĩa các ứng dụng của drone

**Chương 14: Sử dụng xe tự hành AI**

* Tìm hiểu lịch sử ngắn gọn
* Hiểu tương lai của khả năng di chuyển
* Sử dụng xe tự hành
* Vượt qua sự không chắc chắn trong nhận thức

**Phần 5: Cân nhắc tương lai của AI**

**Chương 15: Hiểu các ứng dụng không khả thi**

* Sử dụng AI ở nơi nó không hoạt động
* Cân nhắc tác động của "mùa đông AI"
* Tạo ra giải pháp cho các vấn đề

**Chương 16: Nhìn thấy AI trong không gian**

* Quan sát vũ trụ
* Khai thác tài nguyên không gian
* Khám phá các địa điểm mới
* Xây dựng cấu trúc trong không gian

**Chương 17: Tham gia vào các nỗ lực của con người**

* Giữ con người nổi bật
* Sống và làm việc trong không gian
* Xây dựng thành phố trong môi trường khắc nghiệt
* Làm cho con người hiệu quả hơn
* Giải quyết các vấn đề ở quy mô hành tinh

**Phần 6: Các nội dung thú vị**

**Chương 18: Mười nhóm nghề mà AI không thể thay thế**

* Thực hiện tương tác con người
* Tạo ra những điều mới
* Đưa ra quyết định trực giác

**Chương 19: Mười đóng góp đáng kể của AI cho xã hội**

* Cân nhắc các tương tác đặc trưng của con người
* Phát triển các giải pháp công nghiệp
* Tạo ra môi trường công nghệ mới
* Làm việc với AI trong không gian

**Chương 20: Mười cách AI đã thất bại**

* Hiểu biết
* Khám phá
* Đồng cảm

**Phụ lục: Mục lục tra cứu**  
**Thông tin về tác giả**  
**Trang quảng cáo**  
**Kết nối với Dummies**  
**Thỏa thuận cấp phép người dùng cuối**

# Table of Contents

Title Page

Introduction

Part 1: Introducing AI

Chapter 1: Introducing AI

Defining the Term AI

Understanding the History of AI

Considering AI Uses

Avoiding AI Hype and Overestimation

Connecting AI to the Underlying Computer

Chapter 2: Defining the Role of Data

Finding Data Ubiquitous in This Age

Using Data Successfully

Manicuring the Data

Considering the Five Mistruths in Data

Defining the Limits of Data Acquisition

Considering Data Security Issues

Chapter 3: Considering the Use of Algorithms

Understanding the Role of Algorithms

Discovering the Learning Machine

Chapter 4: Pioneering Specialized Hardware

Relying on Standard Hardware

Using GPUs

Working with Deep Learning Processors (DLPs)

Creating a Specialized Processing Environment

Increasing Hardware Capabilities

Adding Specialized Sensors

Devising Methods to Interact with the Environment

Part 2: Considering the Uses of AI in Society

Chapter 5: Seeing AI Uses in Computer Applications

Introducing Common Application Types

Seeing How AI Makes Applications Friendlier

Performing Corrections Automatically

Making Suggestions

Considering AI-based Errors

Chapter 6: Automating Common Processes

Developing Solutions for Boredom

Working in Industrial Settings

Creating a Safe Environment

Chapter 7: Using AI to Address Medical Needs

Implementing Portable Patient Monitoring

Making Humans More Capable

Addressing a Range of Physical Abilities

Completing Analysis in New Ways

Relying on Telepresence

Devising New Surgical Techniques

Performing Tasks Using Automation

Combining Robots and Medical Professionals

Chapter 8: Relying on AI to Improve Human Interaction

Developing New Ways to Communicate

Exchanging Ideas

Embellishing Human Sensory Perception

Part 3: Working with Software-Based AI Applications

Chapter 9: Performing Data Analysis for AI

Defining Data Analysis

Defining Machine Learning

Considering How to Learn from Data

Chapter 10: Employing Machine Learning in AI

Taking Many Different Roads to Learning

Exploring the Truth in Probabilities

Growing Trees that Can Classify

Chapter 11: Improving AI with Deep Learning

Shaping Neural Networks Similar to the Human Brain

Mimicking the Learning Brain

Introducing Deep Learning

Detecting Edges and Shapes from Images

Learning to Imitate Art and Life

Part 4: Working with AI in Hardware Applications

Chapter 12: Developing Robots

Defining Robot Roles

Assembling a Basic Robot

Chapter 13: Flying with Drones

Acknowledging the State of the Art

Defining Uses for Drones

Chapter 14: Utilizing the AI-Driven Car

Getting a Short History

Understanding the Future of Mobility

Getting into a Self-Driving Car

Overcoming Uncertainty of Perceptions

Part 5: Considering the Future of AI

Chapter 15: Understanding the Nonstarter Application

Using AI Where It Won’t Work

Considering the Effects of AI Winters

Creating Solutions in Search of a Problem

Chapter 16: Seeing AI in Space

Observing the Universe

Performing Space Mining

Exploring New Places

Building Structures in Space

Chapter 17: Engaging in Human Endeavors

Keeping Human Beings Popular

Living and Working in Space

Creating Cities in Hostile Environments

Making Humans More Efficient

Fixing Problems on a Planetary Scale

Part 6: The Part of Tens

Chapter 18: Ten Occupational Categories that AI Can’t

Replace

Performing Human Interaction

Creating New Things

Making Intuitive Decisions

Chapter 19: Ten Substantial Contributions of AI to Society

Considering Human-Specific Interactions

Developing Industrial Solutions

Creating New Technology Environments

Working with AI in Space

Chapter 20: Ten Ways in Which AI Has Failed

Understanding

Discovering

Empathizing

Index

About the Authors

Advertisement Page

Connect with Dummies

End User License Agreement

# Chương 1: Giới thiệu về Trí tuệ Nhân tạo (AI)

**TRONG CHƯƠNG NÀY**

* Định nghĩa AI và lịch sử của nó.
* Sử dụng AI cho các nhiệm vụ thực tế.
* Nhận diện những cường điệu về AI.
* Kết nối AI với công nghệ máy tính.

Trí tuệ nhân tạo (AI) đã có một số lần bắt đầu, phát triển và đình trệ trong nhiều năm, một phần là do mọi người không thực sự hiểu AI là gì, hoặc thậm chí là AI nên hoàn thiện ra sao. Một phần chính của vấn đề là phim ảnh, chương trình truyền hình và sách đều thông đồng để tạo ra hy vọng sai lầm về những gì AI sẽ hoàn thành.

Ngoài ra, xu hướng *nhân cách hóa* (gắn đặc điểm của con người vào) công nghệ của con người khiến cho AI có vẻ như phải làm nhiều hơn những gì nó có thể hy vọng hoàn thiện. Vì vậy, cách tốt nhất để bắt đầu cuốn sách này là định nghĩa AI thực sự là gì, AI không phải là gì và AI kết nối với máy tính như thế nào trong hiện tại.

## 1. Định nghĩa thuật ngữ AI

Trước khi bạn có thể sử dụng một thuật ngữ theo bất kỳ cách có ý nghĩa và hữu ích nào, bạn phải có định nghĩa cho nó. Xét cho cùng, nếu không ai đồng ý về một nghĩa nào đó, thì thuật ngữ đó không có nghĩa; nó chỉ là một tập hợp các ký tự. Việc định nghĩa thành ngữ (một thuật ngữ mà nghĩa của nó không rõ ràng từ nghĩa của các thành phần cấu thành của nó) đặc biệt quan trọng với các thuật ngữ kỹ thuật đã nhận được nhiều hơn một chút đưa tin trên báo chí vào nhiều thời điểm khác nhau và theo nhiều cách khác nhau.

Nói rằng AI là trí tuệ nhân tạo không thực sự cho bạn biết bất cứ điều gì có ý nghĩa, đó là lý do tại sao có rất nhiều cuộc thảo luận và bất đồng về thuật ngữ này. Đúng, bạn có thể lập luận rằng những gì xảy ra là nhân tạo, không đến từ nguồn tự nhiên. Tuy nhiên, phần trí thông minh, tốt nhất là mơ hồ. Ngay cả khi bạn không nhất thiết đồng ý với định nghĩa về AI như xuất hiện trong các phần sau, cuốn sách này sử dụng AI theo định nghĩa đó và biết điều đó sẽ giúp bạn theo dõi phần còn lại của văn bản dễ dàng hơn.

***Trí thông minh sáng suốt***

Mọi người định nghĩa trí thông minh theo nhiều cách khác nhau. Tuy nhiên, bạn có thể nói rằng trí thông minh bao gồm một số hoạt động tinh thần bao gồm các hoạt động sau:

**Học tập:** Có khả năng tiếp thu và xử lý thông tin mới

**Lý luận:** Có khả năng thao tác thông tin theo nhiều cách khác nhau

**Hiểu biết:** Xem xét kết quả của việc thao túng thông tin

**Nắm bắt sự thật:** Xác định tính hợp lệ của thông tin bị thao túng

**Nhìn thấy mối quan hệ:** Dự đoán cách dữ liệu được xác thực tương tác với dữ liệu khác

**Xem xét ý nghĩa:** Áp dụng chân lý vào những tình huống cụ thể trong một

cách thức phù hợp với mối quan hệ của họ

**Phân biệt sự thật với niềm tin:** Xác định xem dữ liệu có được hỗ trợ đầy đủ bởi các nguồn có thể chứng minh được là có giá trị nhất quán hay không

Danh sách này có thể dễ dàng trở nên khá dài, nhưng ngay cả danh sách này cũng tương đối dễ bị diễn giải bởi bất kỳ ai chấp nhận nó là khả thi. Tuy nhiên, như bạn có thể thấy từ danh sách, trí thông minh thường tuân theo một quy trình mà hệ thống máy tính có thể bắt chước như một phần của mô phỏng:

* Đặt mục tiêu dựa trên nhu cầu hoặc mong muốn.
* Đánh giá giá trị của bất kỳ thông tin hiện có nào để hỗ trợ mục tiêu.
* Thu thập thêm thông tin có thể hỗ trợ mục tiêu. Trọng tâm ở đây là thông tin có thể hỗ trợ mục tiêu, thay vì thông tin mà bạn biết sẽ hỗ trợ mục tiêu.
* Xử lý dữ liệu sao cho nó có hình thức nhất quán với thông tin hiện có.
* Xác định mối quan hệ và giá trị thực giữa thông tin hiện có và thông tin mới.
* Xác định xem mục tiêu có đạt được hay không.
* Điều chỉnh mục tiêu dựa trên dữ liệu mới và tác động của nó đến khả năng thành công.
* Lặp lại các Bước 2 đến 7 nếu cần cho đến khi đạt được mục tiêu (tìm thấy đúng) hoặc không còn khả năng đạt được mục tiêu (tìm thấy sai).

Mặc dù bạn có thể tạo ra các thuật toán và cung cấp quyền truy cập vào dữ liệu để hỗ trợ quá trình này trong máy tính, khả năng đạt được trí thông minh của máy tính bị hạn chế nghiêm trọng. Ví dụ, máy tính không có khả năng hiểu bất cứ điều gì vì nó dựa vào các quy trình máy để xử lý dữ liệu bằng toán học thuần túy theo cách hoàn toàn cơ học thời trang. Tương tự như vậy, máy tính không thể dễ dàng phân biệt sự thật với sự dối trá (như đã mô tả trong [Chương 2](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-01-introducing-ai/#page45) ). Trên thực tế, không có máy tính nào có thể thực hiện đầy đủ bất kỳ hoạt động tinh thần nào được mô tả trong danh sách mô tả trí thông minh.

Như một phần của việc quyết định trí thông minh thực sự bao gồm những gì, việc phân loại trí thông minh cũng hữu ích. Con người không chỉ sử dụng một loại trí thông minh, mà còn dựa vào nhiều loại trí thông minh để thực hiện nhiệm vụ. Howard Gardner của Harvard đã định nghĩa một số loại trí thông minh này (xem bài viết “Trí thông minh đa dạng” của Project Zero tại Đại học Harvard để biết chi tiết) và việc biết chúng giúp bạn liên hệ chúng với các loại nhiệm vụ mà máy tính có thể mô phỏng như trí thông minh (xem [Bảng 1-1](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-01-introducing-ai/#page28) để biết phiên bản đã sửa đổi của các loại trí thông minh này với mô tả bổ sung).

[**BẢNG 1-1**](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-01-introducing-ai/#page28)**Các loại trí thông minh của con người và cách AI mô phỏng chúng**

|  |  |  |  |
| --- | --- | --- | --- |
| **Kiểu** | **Tiềm năng mô phỏng** | **Công cụ của con người** | **Sự miêu tả** |
| Thị giác - không gian | Vừa phải | Mô hình, đồ họa, biểu đồ, ảnh chụp, bản vẽ, mô hình 3-D, video, truyền hình và đa phương tiện | Trí thông minh vật lý-môi trường được sử dụng bởi những người như thủy thủ và kiến ​​trúc sư. Để di chuyển, con người cần hiểu môi trường vật lý của họ — tức là, kích thước và đặc điểm của nó.  Mọi trí thông minh của rô-bốt hoặc máy tính xách tay đều yêu cầu khả năng này, nhưng khả năng này thường khó mô phỏng (như với xe tự lái) hoặc không chính xác (như với máy hút bụi dựa nhiều vào va chạm như chúng di chuyển thông minh). |
| Cơ thể - vận động | Trung bình đến Cao | Thiết bị chuyên dụng và vật thể thực tế | Các chuyển động cơ thể, chẳng hạn như các chuyển động được sử dụng bởi một bác sĩ phẫu thuật hoặc một vũ công, đòi hỏi sự chính xác và nhận thức về cơ thể. Robot thường sử dụng loại trí thông minh này để thực hiện các nhiệm vụ lặp đi lặp lại, thường có độ chính xác cao hơn con người, nhưng đôi khi kém duyên dáng hơn.  Điều cần thiết là phải phân biệt giữa sự tăng cường của con người, chẳng hạn như một thiết bị phẫu thuật cung cấp cho bác sĩ phẫu thuật khả năng thể chất được tăng cường, và chuyển động độc lập thực sự. Cái trước chỉ đơn giản là một minh chứng cho khả năng toán học ở chỗ nó phụ thuộc vào đầu vào của bác sĩ phẫu thuật. |
| Sáng tạo | Không có | Sản phẩm nghệ thuật, mô hình tư duy mới, phát minh, thể loại sáng tác âm nhạc mới | Sáng tạo là hành động phát triển một mô hình tư duy mới dẫn đến kết quả đầu ra độc đáo dưới dạng nghệ thuật, âm nhạc và văn bản. Một loại sản phẩm thực sự mới là kết quả của sự sáng tạo. Một AI có thể mô phỏng các mô hình tư duy hiện có và thậm chí kết hợp chúng để tạo ra thứ có vẻ là một bản trình bày độc đáo nhưng thực chất chỉ là phiên bản dựa trên toán học của một mô hình hiện có. Để sáng tạo, một AI sẽ cần phải có nhận thức về bản thân, điều này đòi hỏi trí thông minh nội tâm. |
| Quan hệ giữa các cá nhân | Thấp đến trung bình | Điện thoại, hội nghị truyền hình, hội nghị truyền thanh, viết, hội nghị máy tính, email | Tương tác với người khác diễn ra ở nhiều cấp độ. Mục tiêu của dạng trí thông minh này là thu thập, trao đổi, cung cấp và thao tác thông tin dựa trên kinh nghiệm của người khác. Máy tính có thể trả lời các câu hỏi cơ bản nhờ từ khóa nhập vào, không phải vì chúng hiểu câu hỏi.  Trí thông minh diễn ra trong khi thu thập thông tin, xác định các từ khóa phù hợp, sau đó cung cấp thông tin dựa trên các từ khóa đó. Việc tham chiếu chéo các thuật ngữ trong bảng tra cứu rồi hành động theo các hướng dẫn do bảng cung cấp thể hiện trí thông minh logic, không phải trí thông minh giao tiếp. |
| Nội tâm | Không có | Sách, tài liệu sáng tạo, nhật ký, sự riêng tư và thời gian | Nhìn vào bên trong để hiểu được sở thích của bản thân và sau đó đặt ra mục tiêu dựa trên những sở thích đó hiện là loại trí thông minh chỉ có ở con người. Là máy móc, máy tính không có ham muốn, sở thích, nhu cầu hay khả năng sáng tạo. AI xử lý dữ liệu đầu vào số bằng một tập hợp các thuật toán và cung cấp đầu ra; nó không nhận thức được bất cứ điều gì mà nó làm, cũng không hiểu bất cứ điều gì mà nó làm. |
| Ngôn ngữ (thường được chia thành lời nói, âm thanh và văn bản) | Thấp đối với lời nói và âm thanh Không có đối với văn bản | Trò chơi, đa phương tiện, sách, máy ghi âm và lời nói | Làm việc với các từ ngữ là một công cụ thiết yếu để giao tiếp vì trao đổi thông tin bằng lời nói và văn bản nhanh hơn nhiều so với bất kỳ hình thức nào khác.  Hình thức trí thông minh này bao gồm hiểu đầu vào bằng miệng, bằng thính giác và bằng văn bản, quản lý đầu vào để phát triển câu trả lời và cung cấp câu trả lời dễ hiểu dưới dạng đầu ra.  Trong nhiều trường hợp, máy tính hầu như không thể phân tích đầu vào thành các từ khóa, thực sự không thể hiểu được yêu cầu và đưa ra các phản hồi có thể không thể hiểu được. Ở con người, trí thông minh ngôn ngữ bằng miệng, bằng thính giác và bằng văn bản đến từ các vùng khác nhau của não (xem “Say What? How the Brain Separates Our Ability to Talk and Write” của Đại học John Hopkins), điều đó có nghĩa là ngay cả với con người, một người có trí thông minh ngôn ngữ bằng văn bản cao có thể không có trí thông minh ngôn ngữ bằng lời nói cao tương tự. Máy tính hiện không tách biệt khả năng ngôn ngữ bằng thính giác và bằng lời nói — một chỉ đơn giản là đầu vào và đầu ra còn lại. Máy tính không thể mô phỏng khả năng ngôn ngữ bằng văn bản vì khả năng này đòi hỏi sự sáng tạo. |
| Logic- toán học | Cao (có khả năng cao hơn con người) | Trò chơi logic, điều tra, bí ẩn và câu đố trí tuệ | Tính toán kết quả, thực hiện so sánh, khám phá các mẫu và xem xét các mối quan hệ là tất cả các lĩnh vực mà máy tính hiện đang vượt trội. Khi bạn thấy máy tính đánh bại con người trong một chương trình trò chơi, đây là dạng trí thông minh duy nhất mà bạn thực sự nhìn thấy khả năng vượt trội của máy tính trong số bảy loại trí thông minh.  Đúng, bạn có thể thấy những phần nhỏ của các loại trí thông minh khác , nhưng đây là trọng tâm. Việc dựa vào một lĩnh vực để đánh giá trí thông minh của con người so với máy tính không phải là một ý kiến ​​hay. |

***Khám phá bốn cách để định nghĩa AI***

Như đã mô tả trong phần trước, khái niệm đầu tiên quan trọng cần hiểu là AI thực sự không liên quan gì đến trí thông minh của con người. Đúng, một số AI được mô hình hóa để mô phỏng trí thông minh của con người, nhưng đó chính là bản chất của nó: một mô phỏng. Khi nghĩ về AI, hãy chú ý đến sự tương tác giữa việc tìm kiếm mục tiêu, xử lý dữ liệu được sử dụng để đạt được mục tiêu đó và thu thập dữ liệu được sử dụng để hiểu rõ hơn về mục tiêu. AI dựa vào các thuật toán để đạt được kết quả có thể hoặc không liên quan gì đến con người

mục tiêu hoặc phương pháp để đạt được những mục tiêu đó. Với điều này trong tâm trí, bạn có thể phân loại AI theo bốn cách:

– **Hành động giống con người:** Khi máy tính hành động như con người, nó phản ánh tốt nhất Bài kiểm tra Turing, trong đó máy tính thành công khi không thể phân biệt được máy tính và con người (xem “Bài kiểm tra Turing” tại Sổ lưu niệm Internet Alan Turing để biết chi tiết). Thể loại này cũng phản ánh những gì phương tiện truyền thông muốn bạn tin rằng AI là gì. Bạn thấy nó được sử dụng cho các công nghệ như xử lý ngôn ngữ tự nhiên, biểu diễn kiến ​​thức, suy luận tự động và học máy (cả bốn đều phải có để vượt qua bài kiểm tra). Để vượt qua bài kiểm tra Turing, AI phải có cả bốn công nghệ trước đó và có thể tích hợp các giải pháp khác (chẳng hạn như hệ thống chuyên gia). Mitsuku (tìm thấy tại <https://chat.kuki.ai/> và <http://www.square-bear.co.uk/mitsuku/home.htm> ), một chatbot đã giành Giải thưởng Loebner năm lần cho trí tuệ nhân tạo giống con người nhất, là một ví dụ về sự tích hợp như vậy.

*Bài kiểm tra Turing ban đầu không bao gồm bất kỳ tiếp xúc vật lý nào. Bài kiểm tra Turing toàn diện của Harnad có bao gồm tiếp xúc vật lý, dưới dạng thẩm vấn khả năng nhận thức, nghĩa là máy tính cũng phải sử dụng cả thị giác máy tính và robot để thành công. Sau đây là tổng quan nhanh về các phương án thay thế khác cho Bài kiểm tra Turing:*

+Bài kiểm tra Turing đảo ngược: Một con người cố gắng thuyết phục máy tính rằng người đó không phải là máy tính (ví dụ: Bài kiểm tra Turing công khai hoàn toàn tự động để phân biệt máy tính và con người, hay CAPTCHA).

+Bài kiểm tra tín hiệu thông minh tối thiểu: Chỉ đưa ra các câu hỏi đúng/sai và có/không.

+ Bài kiểm tra Marcus: Một chương trình máy tính mô phỏng việc xem một chương trình truyền hình và chương trình được kiểm tra bằng những câu hỏi có ý nghĩa về nội dung của chương trình.

+Bài kiểm tra Lovelace 2.0: Bài kiểm tra phát hiện AI thông qua việc kiểm tra khả năng sáng tạo nghệ thuật của nó.

+ Winograd Schema Challenge: Bài kiểm tra này yêu cầu các câu hỏi trắc nghiệm theo một định dạng cụ thể.

Các kỹ thuật hiện đại bao gồm ý tưởng đạt được mục tiêu thay vì bắt chước hoàn toàn con người. Ví dụ, anh em nhà Wright đã không thành công trong việc tạo ra máy bay bằng cách sao chép chính xác cách bay của loài chim; thay vào đó, loài chim đã đưa ra những ý tưởng dẫn đến khí động học, cuối cùng dẫn đến việc con người bay. Mục tiêu là bay. Cả loài chim và con người đều đạt được mục tiêu này, nhưng họ sử dụng các phương pháp tiếp cận khác nhau.

– **Suy nghĩ như con người:** Khi máy tính suy nghĩ như con người, nó thực hiện các nhiệm vụ đòi hỏi trí thông minh (trái ngược với các quy trình máy móc) từ con người để thành công, chẳng hạn như lái xe. Để xác định xem một chương trình có suy nghĩ như con người hay không, bạn phải có một số phương pháp xác định cách con người suy nghĩ, mà phương pháp mô hình hóa nhận thức định nghĩa. Mô hình này dựa trên ba kỹ thuật:

Nội quan: Phát hiện và ghi chép các kỹ thuật được sử dụng để đạt được mục tiêu bằng cách theo dõi quá trình suy nghĩ của chính mình.

Kiểm tra tâm lý: Quan sát hành vi của một người và thêm vào cơ sở dữ liệu về các hành vi tương tự của những người khác trong cùng hoàn cảnh, mục tiêu, nguồn lực và điều kiện môi trường (cùng nhiều yếu tố khác).

– **Chụp ảnh não** : Theo dõi hoạt động của não trực tiếp thông qua nhiều phương tiện cơ học khác nhau, chẳng hạn như Chụp cắt lớp trục vi tính (CAT), Chụp cắt lớp phát xạ positron (PET), Chụp cộng hưởng từ (MRI) và Chụp từ não đồ (MEG).

Sau khi tạo mô hình, bạn có thể viết chương trình mô phỏng mô hình. Với lượng biến thiên giữa các quá trình suy nghĩ của con người và khó khăn trong việc thể hiện chính xác các quá trình suy nghĩ này như một phần của chương trình, kết quả tốt nhất chỉ là thử nghiệm. Thể loại suy nghĩ theo cách của con người này thường được sử dụng trong tâm lý học và các lĩnh vực khác, trong đó việc mô hình hóa quá trình suy nghĩ của con người để tạo ra các mô phỏng thực tế là điều cần thiết.

– **Suy nghĩ hợp lý** : Nghiên cứu cách con người suy nghĩ bằng một số tiêu chuẩn cho phép tạo ra các hướng dẫn mô tả hành vi điển hình của con người. Một người được coi là hợp lý khi tuân theo các hành vi này trong một số mức độ sai lệch nhất định. Một máy tính suy nghĩ hợp lý dựa vào các hành vi được ghi lại để tạo ra một hướng dẫn về cách tương tác với môi trường dựa trên dữ liệu có trong tay. Mục tiêu của cách tiếp cận này là giải quyết các vấn đề một cách hợp lý, khi có thể. Trong nhiều trường hợp, cách tiếp cận này sẽ cho phép tạo ra một kỹ thuật cơ sở để giải quyết một vấn đề, sau đó sẽ được sửa đổi để thực sự giải quyết vấn đề. Nói cách khác, việc giải quyết một vấn đề về nguyên tắc thường khác với việc giải quyết nó trong thực tế, nhưng bạn vẫn cần một điểm khởi đầu.

– **Hành động hợp lý** : Nghiên cứu cách con người hành động trong những tình huống nhất định dưới những ràng buộc cụ thể giúp bạn xác định được kỹ thuật nào vừa hiệu quả vừa hiệu suất. Một máy tính hoạt động hợp lý dựa vào các hành động được ghi lại để tương tác với môi trường dựa trên các điều kiện, yếu tố môi trường và dữ liệu hiện có. Cũng giống như tư duy hợp lý, các hành động hợp lý phụ thuộc vào một giải pháp về nguyên tắc, có thể không hữu ích trong thực tế. Tuy nhiên, các hành động hợp lý cung cấp một đường cơ sở mà máy tính có thể bắt đầu đàm phán để hoàn thành thành công một mục tiêu.

Các danh mục được sử dụng để định nghĩa AI cung cấp một cách để xem xét các mục đích sử dụng khác nhau hoặc các cách để áp dụng AI. Một số hệ thống được sử dụng để phân loại AI theo loại là tùy ý và không riêng biệt. Ví dụ, một số nhóm xem AI là mạnh (trí thông minh tổng quát có thể thích ứng với nhiều tình huống khác nhau) hoặc yếu (trí thông minh cụ thể được thiết kế để thực hiện tốt một nhiệm vụ cụ thể). Vấn đề với AI mạnh là nó không thực hiện tốt bất kỳ nhiệm vụ nào, trong khi AI yếu quá cụ thể để thực hiện các nhiệm vụ một cách độc lập. Mặc dù vậy, chỉ có hai phân loại loại sẽ không hoàn thành công việc ngay cả theo nghĩa chung. Bốn loại phân loại do Arend Hintze thúc đẩy (xem “Hiểu bốn loại AI, từ rô-bốt phản ứng đến các sinh vật có ý thức về bản thân” tại Conversation.com để biết chi tiết) tạo thành cơ sở tốt hơn để hiểu AI:

**Máy phản ứng:** Những cỗ máy bạn thấy đánh bại con người trong trò chơi cờ vua hoặc chơi trên các chương trình trò chơi là ví dụ về máy phản ứng. Một cỗ máy phản ứng không có trí nhớ hoặc kinh nghiệm để đưa ra quyết định. Thay vào đó, nó dựa vào sức mạnh tính toán thuần túy và các thuật toán thông minh để tạo lại mọi quyết định mọi lúc. Đây là ví dụ về AI yếu được sử dụng cho một mục đích cụ thể. (Phần “Xem xét lập luận Phòng Trung Quốc” của Chương 5 giải thích ý nghĩa của AI yếu.)

**Bộ nhớ hạn chế:** Một chiếc xe SD hoặc rô-bốt tự động không đủ thời gian để đưa ra mọi quyết định từ đầu. Những cỗ máy này dựa vào một lượng nhỏ bộ nhớ để cung cấp kiến ​​thức kinh nghiệm về nhiều tình huống khác nhau. Khi máy nhìn thấy cùng một tình huống, nó có thể dựa vào kinh nghiệm để giảm thời gian phản ứng và cung cấp nhiều tài nguyên hơn để đưa ra các quyết định mới chưa được đưa ra. Đây là một ví dụ về mức độ AI mạnh hiện tại.

**Lý thuyết về tâm trí:** Một cỗ máy có thể đánh giá cả mục tiêu bắt buộc của nó và mục tiêu tiềm năng của các thực thể khác trong cùng một môi trường có một loại hiểu biết khả thi ở một mức độ nào đó hiện nay, nhưng không có ở bất kỳ hình thức thương mại nào. Tuy nhiên, để xe SD thực sự trở nên tự động, mức độ AI này phải được phát triển đầy đủ. Một chiếc xe SD không chỉ cần biết rằng nó phải đi từ điểm này đến điểm khác, mà còn phải trực giác được các mục tiêu có khả năng xung đột của những người lái xe xung quanh nó và phản ứng phù hợp. (Robot soccer,  
<http://www.cs.cmu.edu/~robosoccer/main/> và <https://www.robocup.org/> , là một ví dụ khác về loại hiểu biết này, nhưng ở mức độ đơn giản.)

**Tự nhận thức:** Đây là loại AI mà bạn thấy trong phim. Tuy nhiên, nó đòi hỏi những công nghệ thậm chí không thể thực hiện được ngay bây giờ vì một cỗ máy như vậy sẽ có cả ý thức về bản thân và ý thức. Ngoài ra, thay vì chỉ trực giác mục tiêu của người khác dựa trên môi trường và các phản ứng của thực thể khác, loại máy này sẽ có thể suy ra ý định của người khác dựa trên kiến ​​thức kinh nghiệm.

## 2. Hiểu về lịch sử của AI

Các phần trước của chương này giúp bạn hiểu được trí thông minh theo quan điểm của con người và thấy được máy tính hiện đại không đủ khả năng mô phỏng trí thông minh như vậy, chứ đừng nói đến việc thực sự trở nên thông minh. Tuy nhiên, mong muốn tạo ra những cỗ máy thông minh (hoặc, trong thời cổ đại, là thần tượng) đã có từ lâu đời như con người. Mong muốn không đơn độc trong vũ trụ, có thứ gì đó để giao tiếp mà không có sự bất nhất của những con người khác, là một mong muốn mạnh mẽ. Tất nhiên, một cuốn sách duy nhất không thể bao quát toàn bộ lịch sử loài người, vì vậy các phần sau đây cung cấp một bản tóm tắt ngắn gọn, có liên quan về lịch sử của những nỗ lực AI hiện đại.

***Bắt đầu với logic biểu tượng tại Dartmouth***

Những chiếc máy tính đầu tiên chỉ là như vậy: thiết bị tính toán. Chúng mô phỏng khả năng thao tác các ký hiệu của con người để thực hiện các nhiệm vụ toán học cơ bản, chẳng hạn như phép cộng. Suy luận logic sau đó bổ sung khả năng thực hiện suy luận toán học thông qua các phép so sánh (chẳng hạn như xác định xem một giá trị có lớn hơn giá trị khác không). Tuy nhiên, con người vẫn cần phải xác định thuật toán được sử dụng để thực hiện phép tính, cung cấp dữ liệu cần thiết theo đúng định dạng và sau đó diễn giải kết quả. Vào mùa hè năm 1956, nhiều nhà khoa học đã tham dự một hội thảo được tổ chức tại khuôn viên trường Cao đẳng Dartmouth để làm điều gì đó nhiều hơn. Họ dự đoán rằng những cỗ máy có thể suy luận hiệu quả như con người sẽ cần, nhiều nhất, một thế hệ để xuất hiện. Họ đã sai. Chỉ đến bây giờ, chúng ta mới nhận ra những cỗ máy có thể thực hiện suy luận toán học và logic hiệu quả như con người (điều đó có nghĩa là máy tính phải nắm vững ít nhất sáu loại trí thông minh nữa trước khi đạt đến bất kỳ thứ gì gần với trí thông minh của con người).

Vấn đề được nêu ra với Dartmouth College và các nỗ lực khác vào thời điểm đó liên quan đến phần cứng — khả năng xử lý để thực hiện các phép tính đủ nhanh để tạo ra một mô phỏng. Tuy nhiên, đó không thực sự là toàn bộ vấn đề. Đúng, phần cứng có ảnh hưởng đến bức tranh, nhưng bạn không thể mô phỏng các quy trình mà bạn không hiểu. Mặc dù vậy, lý do khiến AI có phần hiệu quả ngày nay là vì phần cứng cuối cùng đã trở nên đủ mạnh để hỗ trợ số lượng phép tính cần thiết.

Vấn đề lớn nhất với những nỗ lực ban đầu này (và vẫn là một vấn đề đáng kể cho đến ngày nay) là chúng ta không hiểu con người lý luận đủ tốt để tạo ra bất kỳ loại mô phỏng nào — giả sử rằng mô phỏng trực tiếp là khả thi. Hãy xem xét lại các vấn đề xung quanh chuyến bay có người lái được mô tả trước đó trong chương. Anh em nhà Wright đã thành công không phải bằng cách mô phỏng các loài chim mà bằng cách hiểu các quá trình mà các loài chim sử dụng, do đó tạo ra lĩnh vực khí động học. Do đó, khi ai đó nói rằng cải tiến AI lớn tiếp theo sắp xuất hiện nhưng lại không có luận văn cụ thể nào về các quá trình liên quan, thì cải tiến đó không hề sắp xuất hiện.

***Tiếp tục với hệ thống chuyên gia***

Hệ thống chuyên gia lần đầu tiên xuất hiện vào những năm 1970 và một lần nữa vào những năm 1980 như một nỗ lực nhằm giảm các yêu cầu tính toán do AI đặt ra bằng cách sử dụng kiến ​​thức của các chuyên gia. Một số biểu diễn hệ thống chuyên gia đã xuất hiện, bao gồm dựa trên quy tắc (sử dụng các câu lệnh if…then để đưa ra quyết định dựa trên các quy tắc chung), dựa trên khung (sử dụng các cơ sở dữ liệu được tổ chức thành các hệ thống phân cấp thông tin chung có liên quan được gọi là khung) và dựa trên logic (dựa trên lý thuyết tập hợp để thiết lập các mối quan hệ). Sự ra đời của các hệ thống chuyên gia rất quan trọng vì chúng trình bày các triển khai AI thực sự hữu ích và thành công đầu tiên.

Bạn vẫn thấy các hệ thống chuyên gia được sử dụng ngày nay (mặc dù chúng không còn được gọi như vậy nữa). Ví dụ, trình kiểm tra chính tả và ngữ pháp trong ứng dụng của bạn là một dạng hệ thống chuyên gia. Đặc biệt, trình kiểm tra ngữ pháp dựa trên quy tắc chặt chẽ. Bạn nên tìm hiểu xung quanh để xem những nơi khác mà hệ thống chuyên gia vẫn có thể được sử dụng thực tế trong các ứng dụng hàng ngày.

Một vấn đề với các hệ thống chuyên gia là chúng có thể khó tạo và duy trì. Những người dùng đầu tiên phải học các ngôn ngữ lập trình chuyên biệt như List Processing (Lisp) hoặc Prolog. Một số nhà cung cấp đã nhìn thấy cơ hội để đưa các hệ thống chuyên gia vào tay những lập trình viên ít kinh nghiệm hoặc mới vào nghề bằng cách sử dụng các sản phẩm như VP-Expert (xem *The Illustrated VP-Expert* tại Amazon.com), dựa trên phương pháp tiếp cận dựa trên quy tắc. Tuy nhiên, các sản phẩm này thường cung cấp chức năng cực kỳ hạn chế khi sử dụng các cơ sở kiến ​​thức nhỏ.

Vào những năm 1990, cụm từ *hệ thống chuyên gia* bắt đầu biến mất. Ý tưởng cho rằng hệ thống chuyên gia là một sự thất bại đã xuất hiện, nhưng thực tế là hệ thống chuyên gia đơn giản là thành công đến mức chúng đã ăn sâu vào các ứng dụng mà chúng được thiết kế để hỗ trợ. Lấy ví dụ về trình xử lý văn bản, tại một thời điểm nào đó, bạn cần phải mua một ứng dụng kiểm tra ngữ pháp riêng như RightWriter. Tuy nhiên, hiện nay, trình xử lý văn bản có trình kiểm tra ngữ pháp tích hợp sẵn vì chúng tỏ ra rất hữu ích (nếu không phải lúc nào cũng chính xác; hãy xem bài viết *của Washington Post* “Xin chào, anh Chips PCS Learn English” để biết chi tiết).

***Vượt qua ‘mùa đông AI’***

Thuật ngữ *AI winter* ám chỉ giai đoạn cắt giảm tài trợ trong quá trình phát triển AI. Nhìn chung, AI đã đi theo con đường mà những người ủng hộ phóng đại những gì có thể, khuyến khích những người không có kiến ​​thức về công nghệ nhưng lại có nhiều tiền đầu tư. Sau đó là giai đoạn chỉ trích khi AI không đáp ứng được kỳ vọng và cuối cùng, việc cắt giảm tài trợ diễn ra. Một số chu kỳ như vậy đã xảy ra trong những năm qua — tất cả đều tàn phá sự tiến bộ thực sự.

AI hiện đang trong giai đoạn cường điệu mới vì *máy học (machine learning),* một công nghệ giúp máy tính học từ dữ liệu. Để máy tính học từ dữ liệu có nghĩa là không phụ thuộc vào một lập trình viên con người để thiết lập các hoạt động (nhiệm vụ), mà là suy ra chúng trực tiếp từ các ví dụ cho thấy máy tính nên hoạt động như thế nào. Giống như giáo dục trẻ sơ sinh bằng cách chỉ cho trẻ cách cư xử thông qua ví dụ. Máy học có những cạm bẫy vì máy tính có thể học cách làm mọi thứ không đúng thông qua việc dạy bất cẩn.

Năm nhóm nhà khoa học đang nghiên cứu các thuật toán học máy, mỗi [nhóm theo một góc nhìn khác nhau (xem phần "Tránh cường điệu và đánh giá quá cao AI" ở phần sau của chương này để biết chi tiết). Hiện tại](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-01-introducing-ai/#page40) , giải pháp thành công nhất là *học sâu (deep learning),* một công nghệ cố gắng mô phỏng bộ não con người. Học sâu có thể thực hiện được nhờ sự sẵn có của các máy tính mạnh mẽ, thuật toán thông minh hơn, các tập dữ liệu lớn được tạo ra bởi quá trình số hóa xã hội của chúng ta và các khoản đầu tư lớn từ các doanh nghiệp như Google, Facebook, Amazon và các công ty khác tận dụng sự hồi sinh của AI này cho doanh nghiệp của riêng họ. Mọi người đang nói rằng mùa đông AI đã kết thúc vì học sâu và điều đó đúng cho đến bây giờ. Tuy nhiên, khi bạn nhìn vào cách mọi người đang xem xét AI, bạn có thể dễ dàng nhận ra rằng một giai đoạn chỉ trích khác cuối cùng sẽ xảy ra trừ khi những người ủng hộ giảm bớt giọng điệu. AI có thể làm được những điều tuyệt vời, nhưng chúng chỉ là những điều tuyệt vời tầm thường (chẳng hạn như thực hiện công việc lặp đi lặp lại để tìm ra vắc-xin phòng ngừa Covid-19; xem “AI đang được sử dụng như thế nào để tạo ra và phân phối vắc-xin phòng ngừa COVID-19” tại TechRepublic.com). Phần tiếp theo sẽ mô tả cách AI đang được sử dụng hiện nay.

## 3. Xem xét việc sử dụng AI

Bạn thấy AI được sử dụng trong rất nhiều ứng dụng ngày nay. Vấn đề duy nhất là công nghệ này hoạt động tốt đến mức bạn thậm chí không biết nó tồn tại. Trên thực tế, bạn có thể ngạc nhiên khi thấy rằng nhiều thiết bị gia đình đã sử dụng AI. Ví dụ, một số bộ điều nhiệt thông minh tự động tạo lịch trình cho bạn dựa trên cách bạn kiểm soát nhiệt độ theo cách thủ công. Tương tự như vậy, đầu vào giọng nói được sử dụng để điều khiển một số thiết bị sẽ học cách bạn nói để có thể tương tác tốt hơn với bạn. AI chắc chắn xuất hiện trong ô tô của bạn và đặc biệt là tại nơi làm việc. Trên thực tế, số lượng ứng dụng của AI lên tới hàng triệu — tất cả đều an toàn và không bị phát hiện ngay cả khi chúng có bản chất khá ấn tượng. Sau đây chỉ là một số cách mà bạn có thể thấy AI được sử dụng:

**Phát hiện gian lận:** Bạn nhận được cuộc gọi từ công ty thẻ tín dụng hỏi xem bạn có thực hiện giao dịch mua cụ thể nào không. Công ty thẻ tín dụng không tò mò; họ chỉ cảnh báo bạn về việc có người khác có thể đang thực hiện giao dịch mua bằng thẻ của bạn. AI được nhúng trong mã của công ty thẻ tín dụng đã phát hiện ra một kiểu chi tiêu lạ và cảnh báo ai đó về điều đó.

**Lên lịch nguồn lực:** Nhiều tổ chức cần lên lịch sử dụng nguồn lực một cách hiệu quả. Ví dụ, bệnh viện có thể phải xác định nơi đặt bệnh nhân dựa trên nhu cầu của bệnh nhân, sự sẵn có của các chuyên gia lành nghề và thời gian bác sĩ mong đợi bệnh nhân ở lại bệnh viện.

**Phân tích phức tạp:** Con người thường cần trợ giúp với phân tích phức tạp vì có quá nhiều yếu tố cần xem xét. Ví dụ, cùng một tập hợp các triệu chứng có thể chỉ ra nhiều hơn một vấn đề. Bác sĩ hoặc chuyên gia khác có thể cần trợ giúp để đưa ra chẩn đoán kịp thời để cứu sống bệnh nhân.

**Tự động hóa:** Bất kỳ hình thức tự động hóa nào cũng có thể được hưởng lợi từ việc bổ sung AI để xử lý các thay đổi hoặc sự kiện bất ngờ. Một vấn đề với một số loại tự động hóa hiện nay là một sự kiện bất ngờ, chẳng hạn như một vật thể ở sai vị trí, thực sự có thể khiến quá trình tự động hóa dừng lại. Việc bổ sung AI vào quá trình tự động hóa có thể cho phép quá trình tự động hóa xử lý các sự kiện bất ngờ và tiếp tục như thể không có gì xảy ra.

**Dịch vụ khách hàng:** Đường dây dịch vụ khách hàng mà bạn gọi đến hôm nay thậm chí có thể không có người trực. Tự động hóa đủ tốt để tuân theo các tập lệnh và sử dụng nhiều nguồn lực khác nhau để xử lý phần lớn các câu hỏi của bạn. Với ngữ điệu giọng nói tốt (cũng do AI cung cấp), bạn thậm chí có thể không biết mình đang nói chuyện với máy tính.

**Hệ thống an toàn:** Nhiều hệ thống an toàn có trong các loại máy móc khác nhau hiện nay dựa vào AI để tiếp quản phương tiện trong thời điểm khủng hoảng. Ví dụ, nhiều hệ thống phanh tự động (ABS) dựa vào AI để dừng xe dựa trên tất cả các dữ liệu đầu vào mà phương tiện có thể cung cấp, chẳng hạn như hướng trượt. ABS vi tính thực sự đã tương đối cũ ở mức 40 năm theo góc độ công nghệ (xem “ABS (Hệ thống chống bó cứng phanh) — Lược sử về một chiếc xe cứu sinh 40 năm tuổi” tại DriveSpark.com để biết chi tiết).

**Hiệu quả của máy móc:** AI có thể giúp kiểm soát máy móc theo cách đạt được hiệu quả tối đa. AI kiểm soát việc sử dụng tài nguyên để hệ thống không vượt quá tốc độ hoặc các mục tiêu khác. Mỗi ounce năng lượng được sử dụng chính xác theo nhu cầu để cung cấp các dịch vụ mong muốn.

## 4.Tránh cường điệu và đánh giá quá cao AI

Chương này đề cập khá nhiều đến sự cường điệu về AI. Thật không may, chương này thậm chí còn không đề cập đến bề nổi của tất cả những sự cường điệu ngoài kia. Nếu bạn xem những bộ phim như *Her* và *Ex Machina,* bạn có thể bị dẫn dắt để tin rằng AI đã tiến xa hơn thực tế. Vấn đề là AI thực sự vẫn còn trong giai đoạn trứng nước, và bất kỳ loại ứng dụng nào như những ứng dụng được thể hiện trong phim đều là sản phẩm sáng tạo của trí tưởng tượng quá mức. Các phần sau đây giúp bạn hiểu cách sự cường điệu và đánh giá quá cao đang làm lệch hướng các mục tiêu mà bạn thực sự có thể đạt được bằng cách sử dụng AI ngày nay.

***Xác định năm nhóm học tập và thuật toán chính***

Bạn có thể đã nghe về một thứ gọi là điểm kỳ dị, chịu trách nhiệm cho các tuyên bố tiềm năng được trình bày trên phương tiện truyền thông và phim ảnh. Điểm *kỳ dị* về cơ bản là một thuật toán chính bao gồm tất cả năm nhóm học tập được sử dụng trong học máy. Để đạt được những gì các nguồn này đang nói với bạn, máy phải có khả năng học như con người — như được chỉ định bởi bảy loại trí thông minh được thảo luận trong phần " [Trí thông minh phân biệt](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-01-introducing-ai/#page26) ", ở đầu chương. Sau đây là năm nhóm học tập:

**Nhà biểu tượng học:** Nguồn gốc của bộ tộc này nằm ở logic và triết học. Nhóm này dựa vào suy luận ngược để giải quyết vấn đề.

**Những người theo thuyết kết nối:** Nguồn gốc của nhóm này xuất phát từ khoa học thần kinh và nhóm này dựa vào phương pháp lan truyền ngược (back-propagation) để giải quyết vấn đề.

**Những người theo chủ nghĩa tiến hóa:** Những người theo chủ nghĩa tiến hóa có nguồn gốc từ sinh học tiến hóa, dựa vào lập trình di truyền để giải quyết vấn đề.

**Người theo thuyết Bayes:** Nguồn gốc của thuyết này là thống kê và dựa vào suy luận xác suất để giải quyết vấn đề.

**Nhà loại suy học - Analogizers:** Nguồn gốc của bộ tộc này là tâm lý học. Nhóm này dựa vào máy ảo hóa (kernel machines) để giải quyết vấn đề.

Mục tiêu cuối cùng của máy học là kết hợp các công nghệ và chiến lược được năm bộ lạc áp dụng để tạo ra một thuật toán duy nhất (*thuật toán chính*) có thể học bất cứ điều gì. Tất nhiên, đạt được mục tiêu đó còn rất xa. Mặc dù vậy, các nhà khoa học như Pedro Domingos tại Đại học Washington hiện đang nỗ lực hướng tới mục tiêu đó.

Để làm cho mọi thứ trở nên ít rõ ràng hơn, năm bộ lạc có thể không cung cấp đủ thông tin để thực sự giải quyết vấn đề về trí thông minh của con người, vì vậy việc tạo ra các thuật toán chính cho cả năm bộ lạc vẫn có thể không mang lại sự kỳ dị. Tại thời điểm này, bạn nên ngạc nhiên về việc mọi người không biết nhiều về cách họ suy nghĩ hoặc lý do tại sao họ suy nghĩ theo một cách nhất định. Bất kỳ tin đồn nào bạn nghe được về việc AI chiếm lĩnh thế giới hoặc trở nên vượt trội hơn con người đều hoàn toàn sai sự thật.

***Xem xét các nguồn gây ra sự cường điệu***

Có nhiều nguồn cường điệu về AI. Khá nhiều sự cường điệu đến từ phương tiện truyền thông và được trình bày bởi những người không biết AI là gì, ngoại trừ có lẽ từ một cuốn tiểu thuyết khoa học viễn tưởng mà họ đã đọc một lần. Vì vậy, không chỉ phim ảnh hay truyền hình gây ra vấn đề với sự cường điệu về AI; mà còn là tất cả các nguồn phương tiện truyền thông khác. Bạn thường có thể tìm thấy các bản tin trình bày AI có khả năng làm điều gì đó mà nó không thể làm được vì phóng viên không hiểu công nghệ. Thật kỳ lạ, nhiều dịch vụ tin tức hiện sử dụng AI để ít nhất là bắt đầu các bài viết cho phóng viên (xem “Có phải một con robot đã viết điều này không? AI đang tác động đến báo chí như thế nào” tại Forbes.com để biết chi tiết).

Một số sản phẩm cần được thử nghiệm nhiều hơn nữa trước khi đưa ra thị trường. Bài viết "Nhìn lại năm 2020: 10 thất bại của AI" tại SyncedReview.com thảo luận về mười sản phẩm được nhà phát triển của chúng thổi phồng nhưng lại thất bại thảm hại. Một số thất bại này rất lớn và phản ánh không tốt về khả năng thực hiện nhiệm vụ của AI nói chung. Tuy nhiên, điều cần cân nhắc với một số thất bại này là con người có thể đã can thiệp vào thiết bị bằng AI. Rõ ràng, các quy trình thử nghiệm cần bắt đầu xem xét khả năng con người cố tình can thiệp vào AI như một nguồn lỗi tiềm ẩn. Cho đến khi điều đó xảy ra, AI sẽ không hoạt động như mong đợi vì mọi người sẽ tiếp tục mày mò phần mềm để cố gắng khiến nó thất bại theo cách hài hước.

Một nguyên nhân khác gây ra vấn đề là hỏi nhầm người về AI. Không phải mọi nhà khoa học, bất kể thông minh đến đâu, đều biết đủ về AI để đưa ra ý kiến ​​có thẩm quyền về công nghệ và hướng đi của nó trong tương lai. Hỏi một nhà sinh vật học về tương lai của AI nói chung cũng giống như yêu cầu nha sĩ của bạn thực hiện phẫu thuật não — điều đó đơn giản là không phải là một ý kiến ​​hay. Tuy nhiên, nhiều câu chuyện xuất hiện với những người như thế này làm nguồn thông tin. Để khám phá hướng đi trong tương lai của AI, tốt nhất là hỏi một nhà khoa học máy tính hoặc nhà khoa học dữ liệu có nền tảng vững chắc về nghiên cứu AI.

***Hiểu được sự đánh giá quá cao của người dùng***

Do sự cường điệu (và đôi khi là lười biếng hoặc mệt mỏi), người dùng liên tục đánh giá quá cao khả năng thực hiện nhiệm vụ của AI. Ví dụ, một chủ xe Tesla gần đây đã bị phát hiện ngủ trong xe khi chiếc xe phóng nhanh trên đường cao tốc với tốc độ 90 dặm/giờ (xem “Chủ xe Tesla ở Canada bị buộc tội 'ngủ' khi lái xe với tốc độ hơn 90 dặm/giờ”). Tuy nhiên, ngay cả khi người dùng đánh giá quá cao khả năng điều khiển ô tô của công nghệ, thì rõ ràng là nó vẫn hoạt động đủ tốt (ít nhất là đối với người lái xe này) để tránh hỏng hóc hoàn toàn. Tuy nhiên, bạn không cần phải tăng tốc trên đường cao tốc với tốc độ 90 dặm/giờ để gặp phải sự đánh giá quá cao của người dùng. Robot hút bụi cũng có thể không đáp ứng được kỳ vọng, thường là do người dùng tin rằng họ chỉ cần cắm thiết bị vào và sau đó không bao giờ nghĩ đến việc hút bụi nữa. Rốt cuộc, các bộ phim mô tả các thiết bị hoạt động chính xác theo cách này. Bài viết “Cách giải quyết các sự cố khó chịu nhất của máy hút bụi robot” tại RobotsInMyHome.com thảo luận về các kỹ thuật khắc phục sự cố cho nhiều loại máy hút bụi robot khác nhau vì một lý do chính đáng — robot vẫn cần sự can thiệp của con người. Vấn đề là hầu hết các robot đều cần sự can thiệp của con người vào một thời điểm nào đó vì chúng đơn giản là không có đủ kiến ​​thức để tự mình hoạt động.

## 5. Kết nối AI với máy tính cơ bản

Để thấy AI hoạt động, bạn cần có một số loại hệ thống máy tính, một ứng dụng chứa phần mềm cần thiết và một cơ sở kiến ​​thức. Hệ thống máy tính có thể là bất kỳ thứ gì có chip bên trong; trên thực tế, điện thoại thông minh cũng hoạt động tốt như máy tính để bàn đối với một số ứng dụng. Tất nhiên, nếu bạn là Amazon và muốn cung cấp lời khuyên về quyết định mua hàng tiếp theo của một người cụ thể, điện thoại thông minh sẽ không làm được — bạn cần một hệ thống máy tính thực sự lớn cho ứng dụng đó. Kích thước của hệ thống máy tính tỷ lệ thuận với khối lượng công việc mà bạn mong đợi AI thực hiện.

Ứng dụng cũng có thể khác nhau về kích thước, độ phức tạp và thậm chí là vị trí. Ví dụ, nếu bạn là một doanh nghiệp và muốn phân tích dữ liệu khách hàng để xác định cách tốt nhất để đưa ra lời chào hàng, bạn có thể dựa vào ứng dụng dựa trên máy chủ để thực hiện nhiệm vụ. Mặt khác, nếu bạn là khách hàng và muốn tìm sản phẩm trên Amazon để đi kèm với các mặt hàng mua hiện tại của mình, ứng dụng thậm chí không nằm trên máy tính của bạn; bạn truy cập ứng dụng thông qua ứng dụng dựa trên web nằm trên máy chủ của Amazon.

Cơ sở tri thức cũng khác nhau về vị trí và kích thước. Dữ liệu càng phức tạp, bạn càng có thể khai thác được nhiều hơn từ dữ liệu đó, nhưng bạn cũng cần phải thao tác nhiều hơn. Nghĩa là ‘không có bữa trưa miễn phí’ khi nói đến quản lý tri thức. Sự tương tác giữa vị trí và thời gian cũng rất quan trọng. Kết nối mạng cho phép bạn truy cập vào cơ sở tri thức trực tuyến lớn nhưng tốn thời gian vì độ trễ của kết nối mạng. Tuy nhiên, cơ sở dữ liệu cục bộ, mặc dù nhanh, nhưng có xu hướng thiếu chi tiết trong nhiều trường hợp.

# Chương 2: Định nghĩa vai trò của dữ liệu

**TRONG CHƯƠNG NÀY**

* Nhìn nhận dữ liệu như một tài nguyên phổ quát.
* Thu thập và xử lý dữ liệu.
* Tìm kiếm những sai lệch trong dữ liệu.
* Xác định giới hạn thu thập dữ liệu.
* Xem xét vấn đề bảo mật dữ liệu.

Không có gì mới về dữ liệu. Mọi ứng dụng thú vị từng được viết cho máy tính đều có dữ liệu liên quan. Dữ liệu có nhiều dạng — một số được sắp xếp, một số thì không. Điều đã thay đổi là lượng dữ liệu. Một số người thấy gần như đáng sợ khi chúng ta hiện có quyền truy cập vào quá nhiều dữ liệu chi tiết gần như mọi khía cạnh trong cuộc sống của hầu hết mọi người, đôi khi ở mức độ mà ngay cả người đó cũng không nhận ra. Ngoài ra, việc sử dụng phần cứng tiên tiến và cải tiến trong thuật toán khiến dữ liệu trở thành nguồn tài nguyên chung cho AI ngày nay.

Để làm việc với dữ liệu, trước tiên bạn phải có được dữ liệu. Ngày nay, các ứng dụng thu thập dữ liệu theo cách thủ công, như đã làm trong quá khứ, và cũng tự động, bằng các phương pháp mới. Tuy nhiên, vấn đề không chỉ là một hoặc hai kỹ thuật thu thập dữ liệu; các phương pháp thu thập diễn ra trên một chuỗi liên tục từ hoàn toàn thủ công đến hoàn toàn tự động. Ngày nay, bạn cũng thấy trọng tâm là thu thập dữ liệu này một cách có đạo đức — ví dụ, không thu thập dữ liệu mà một người chưa cấp phép. Chương này khám phá các vấn đề xung quanh việc thu thập dữ liệu.

Dữ liệu thô thường không hiệu quả cho mục đích phân tích. Chương này cũng giúp bạn hiểu nhu cầu thao tác và định hình dữ liệu để đáp ứng các yêu cầu cụ thể. Bạn cũng khám phá ra nhu cầu xác định giá trị thực của dữ liệu để đảm bảo rằng kết quả phân tích phù hợp với các mục tiêu đặt ra cho các ứng dụng ngay từ đầu.

Điều thú vị là bạn cũng phải giải quyết giới hạn thu thập dữ liệu. Không

công nghệ hiện tại có thể nắm bắt suy nghĩ từ tâm trí của ai đó thông qua phương tiện thần giao cách cảm. Tất nhiên, cũng có những giới hạn khác nữa — hầu hết trong số đó có thể bạn đã biết nhưng có thể chưa cân nhắc. Việc thu thập dữ liệu theo cách không an toàn cũng không có lợi. Dữ liệu phải không có sự thiên vị, không bị hỏng và đến từ nguồn mà bạn biết. Bạn sẽ tìm hiểu thêm về giới hạn thu thập và bảo mật dữ liệu trong chương này.

## 1. Vấn đề tìm kiếm dữ liệu trong thời đại này

*Dữ liệu lớn* không chỉ là một cụm từ thông dụng được các nhà cung cấp sử dụng để đề xuất những cách mới để lưu trữ và phân tích dữ liệu. Cuộc cách mạng dữ liệu lớn là một thực tế hàng ngày và là động lực thúc đẩy thời đại của chúng ta. Bạn có thể đã nghe dữ liệu lớn được đề cập trong nhiều ấn phẩm khoa học và kinh doanh chuyên ngành, và bạn thậm chí có thể tự hỏi thuật ngữ này thực sự có nghĩa là gì. Theo quan điểm kỹ thuật, *dữ liệu lớn* đề cập đến lượng dữ liệu máy tính lớn và phức tạp, lớn và phức tạp đến mức các ứng dụng không thể xử lý dữ liệu chỉ bằng cách sử dụng thêm dung lượng lưu trữ hoặc tăng công suất máy tính.

Dữ liệu lớn ngụ ý một cuộc cách mạng trong lưu trữ và thao tác dữ liệu. Nó ảnh hưởng đến những gì bạn có thể đạt được với dữ liệu theo các thuật ngữ định tính hơn (có nghĩa là ngoài việc làm được nhiều hơn, bạn có thể thực hiện các tác vụ tốt hơn). Theo quan điểm của con người, máy tính lưu trữ dữ liệu lớn ở các định dạng dữ liệu khác nhau (chẳng hạn như tệp cơ sở dữ liệu và tệp .csv), nhưng bất kể loại lưu trữ nào, máy tính vẫn coi dữ liệu là một luồng số một và số không (ngôn ngữ cốt lõi của máy tính). Bạn có thể xem dữ liệu là một trong hai loại, có cấu trúc và không có cấu trúc, tùy thuộc vào cách bạn tạo ra và sử dụng dữ liệu đó. Một số dữ liệu có cấu trúc rõ ràng (bạn biết chính xác dữ liệu đó chứa gì và tìm thấy mọi phần dữ liệu ở đâu), trong khi dữ liệu khác không có cấu trúc (bạn biết dữ liệu đó chứa gì, nhưng không biết chính xác cách sắp xếp dữ liệu đó).

Ví dụ điển hình về dữ liệu có cấu trúc là các bảng cơ sở dữ liệu, trong đó thông tin được sắp xếp thành các cột và mỗi cột chứa một loại thông tin cụ thể. Dữ liệu thường được cấu trúc theo thiết kế. Bạn thu thập dữ liệu một cách có chọn lọc và ghi lại đúng vị trí. Ví dụ, bạn có thể muốn đặt số lượng người mua một sản phẩm nhất định vào một cột cụ thể, trong một bảng cụ thể, trong một cơ sở dữ liệu cụ thể. Giống như thư viện, nếu bạn biết dữ liệu mình cần, bạn có thể tìm thấy dữ liệu đó ngay lập tức.

Dữ liệu phi cấu trúc bao gồm hình ảnh, video và bản ghi âm. Bạn có thể sử dụng dạng phi cấu trúc cho văn bản để có thể gắn thẻ với các đặc điểm như kích thước, ngày hoặc loại nội dung. Thông thường, bạn không biết chính xác dữ liệu xuất hiện ở đâu trong tập dữ liệu phi cấu trúc vì dữ liệu xuất hiện dưới dạng chuỗi số 1 và số 0 mà ứng dụng phải diễn giải hoặc trực quan hóa.

\*Chuyển đổi dữ liệu phi cấu trúc thành dạng có cấu trúc có thể tốn nhiều thời gian và công sức và có thể liên quan đến công sức của nhiều người. Hầu hết dữ liệu của cuộc cách mạng dữ liệu lớn đều không có cấu trúc và được lưu trữ nguyên trạng, trừ khi có ai đó biến nó thành có cấu trúc.

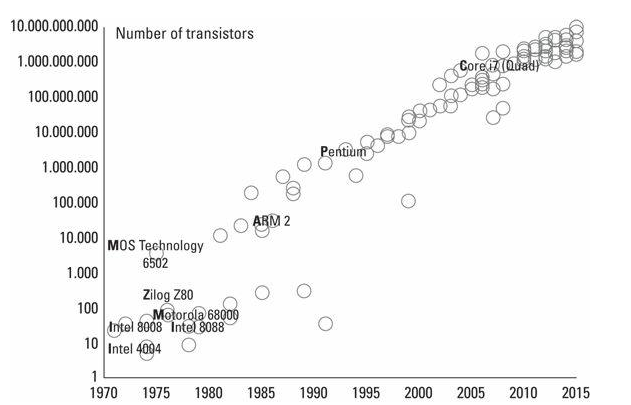
Kho dữ liệu phong phú và tinh vi này không xuất hiện đột ngột chỉ sau một đêm. Phải mất thời gian để phát triển công nghệ lưu trữ lượng dữ liệu này. Ngoài ra, phải mất thời gian để phổ biến công nghệ tạo và phân phối dữ liệu, cụ thể là máy tính, cảm biến, điện thoại di động thông minh và Internet cùng các dịch vụ World Wide Web. Các phần sau đây giúp bạn hiểu được điều gì khiến dữ liệu trở thành nguồn tài nguyên phổ biến hiện nay.

***Hiểu được ý nghĩa của Moore***

Năm 1965, Gordon Moore, người đồng sáng lập Intel và Fairchild Semiconductor, đã viết trong một bài báo có tựa đề "Nhồi nhét nhiều linh kiện hơn vào mạch tích hợp" trên IEEE.org rằng số lượng linh kiện được tìm thấy trong mạch tích hợp sẽ tăng gấp đôi mỗi năm trong thập kỷ tiếp theo. Vào thời điểm đó, bóng bán dẫn thống trị ngành điện tử. Có thể nhồi nhét nhiều bóng bán dẫn hơn vào Mạch tích hợp (IC) có nghĩa là có thể làm cho các thiết bị điện tử có khả năng và hữu ích hơn. Quá trình này được gọi là *tích hợp* và ngụ ý một quá trình thu nhỏ điện tử mạnh mẽ (làm cho cùng một mạch nhỏ hơn nhiều). Máy tính ngày nay không nhỏ hơn nhiều so với máy tính của một thập kỷ trước, nhưng chúng mạnh hơn hẳn. Điện thoại di động cũng vậy. Mặc dù chúng có cùng kích thước với thế hệ trước (và đôi khi thậm chí còn nhỏ hơn), nhưng chúng đã có thể thực hiện nhiều tác vụ hơn.

Những gì Moore nêu trong bài viết đó thực ra đã đúng trong nhiều năm. Ngành công nghiệp bán dẫn gọi đó là Định luật Moore (xem <http://www.mooreslaw.org/> để biết chi tiết). Sự nhân đôi đã xảy ra trong mười năm đầu tiên, như đã dự đoán. Năm 1975, Moore đã sửa lại tuyên bố của mình, dự đoán sự nhân đôi sau mỗi hai năm. [Hình 2-1](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-02-defining-the-role-of-data/#page50) cho thấy tác động của sự nhân đôi này.

\*Bắt đầu từ năm 2012, sự không phù hợp bắt đầu xảy ra giữa tốc độ tăng dự kiến ​​và những gì các công ty bán dẫn có thể đạt được liên quan đến việc thu nhỏ. Nhiều kỹ sư hiện đang nói rằng Định luật Moore đã chết. Tuy nhiên, một số ít, chẳng hạn như Jim Keller, giám đốc kỹ thuật bán dẫn của Intel, cho rằng vẫn còn nhiều chỗ để cải thiện chip. Charles Leiserson và Neil Thompson có quan điểm khác, cho rằng những cải tiến sẽ đến từ phần mềm, thuật toán và kiến ​​trúc chip chuyên dụng tốt hơn.



[**HÌNH 2-1:**](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-02-defining-the-role-of-data/#page49)Nhồi ngày càng nhiều bóng bán dẫn vào CPU.

Trong tương lai, Định luật Moore có thể không còn áp dụng nữa vì ngành công nghiệp sẽ chuyển sang công nghệ mới, chẳng hạn như sản xuất linh kiện bằng cách sử dụng tia laser quang học thay vì bóng bán dẫn (xem bài viết “Máy tính quang học: Giải quyết vấn đề với tốc độ ánh sáng” tại BBVAOpenMind.com để biết thông tin chi tiết về máy tính quang học). Một giải pháp khác dựa vào máy tính lượng tử, chẳng hạn như máy tính đang được Google và IBM phát triển (bạn có thể đọc thêm về công nghệ này trong bài viết “Google xác nhận đột phá về 'ưu thế lượng tử'” tại The Verge.com). Điều quan trọng là kể từ năm 1965, việc tăng gấp đôi số lượng linh kiện sau mỗi hai năm đã mở ra những tiến bộ to lớn trong lĩnh vực điện tử kỹ thuật số, có những hậu quả sâu rộng trong việc thu thập, lưu trữ, thao tác và quản lý dữ liệu.

Định luật Moore có tác động trực tiếp đến dữ liệu. Nó bắt đầu với các thiết bị thông minh hơn. Các thiết bị càng thông minh thì sự khuếch tán càng lớn (bằng chứng là thiết bị điện tử hiện diện ở khắp mọi nơi ngày nay). Sự khuếch tán càng lớn thì giá càng thấp, tạo ra một vòng lặp vô tận thúc đẩy việc sử dụng các máy tính mạnh mẽ và các cảm biến nhỏ ở khắp mọi nơi. Với lượng bộ nhớ máy tính lớn và các đĩa lưu trữ dữ liệu lớn hơn, hậu quả là sự mở rộng tính khả dụng của dữ liệu, chẳng hạn như các trang web, hồ sơ giao dịch, phép đo, hình ảnh kỹ thuật số và các loại dữ liệu khác.

***Sử dụng dữ liệu ở mọi nơi***

Các nhà khoa học cần những chiếc máy tính mạnh hơn người bình thường vì các thí nghiệm khoa học của họ. Họ bắt đầu xử lý lượng dữ liệu ấn tượng nhiều năm trước khi bất kỳ ai nghĩ ra thuật ngữ *dữ liệu lớn.* Vào thời điểm đó, Internet không tạo ra lượng dữ liệu khổng lồ như ngày nay. Hãy nhớ rằng dữ liệu lớn không phải là mốt nhất thời do các nhà cung cấp phần mềm và phần cứng tạo ra mà có cơ sở trong nhiều lĩnh vực khoa học, chẳng hạn như thiên văn học (sứ mệnh không gian), vệ tinh (giám sát và theo dõi), khí tượng học, vật lý (máy gia tốc hạt) và hệ gen học (trình tự DNA).

Mặc dù một ứng dụng AI có thể chuyên về một lĩnh vực khoa học, chẳng hạn như Watson của IBM, tự hào về khả năng chẩn đoán y khoa ấn tượng vì nó có thể học thông tin từ hàng triệu bài báo khoa học về bệnh tật và y học, nhưng trình điều khiển ứng dụng AI thực tế thường có nhiều khía cạnh trần tục hơn. Các ứng dụng AI thực tế chủ yếu được đánh giá cao vì có thể nhận dạng các vật thể, di chuyển dọc theo các đường dẫn hoặc hiểu những gì mọi người nói và nói chuyện với họ. Đóng góp dữ liệu cho sự phục hưng AI thực tế đã định hình nó theo cách như vậy không bắt nguồn từ các nguồn dữ liệu khoa học cổ điển.

Internet hiện nay tạo ra và phân phối dữ liệu mới với số lượng lớn.

Sản lượng dữ liệu hàng ngày hiện tại ước tính lên tới khoảng 2,5 nghìn tỷ (một con số có 18 số không) byte, với phần lớn dữ liệu phi cấu trúc như video và âm thanh. Tất cả dữ liệu này đều liên quan đến các hoạt động, cảm xúc, trải nghiệm và mối quan hệ thông thường của con người. Khi lướt qua dữ liệu này, AI có thể dễ dàng học được cách lý luận và hành động giống con người hơn. Sau đây là một số ví dụ về dữ liệu thú vị hơn mà bạn có thể tìm thấy:

 Các kho lưu trữ lớn về khuôn mặt và biểu cảm từ các bức ảnh và video được đăng trên các trang web truyền thông xã hội như Facebook, YouTube và Google cung cấp thông tin về giới tính, độ tuổi, cảm xúc và có thể là sở thích tình dục, khuynh hướng chính trị hoặc IQ (xem “Giáo sư cho biết AI đọc khuôn mặt sẽ có thể phát hiện quan điểm chính trị và IQ của bạn” tại The Guardian.com).

 Thông tin y tế và dữ liệu sinh trắc học được lưu giữ riêng tư từ đồng hồ thông minh, đo dữ liệu cơ thể như nhiệt độ và nhịp tim trong cả thời gian ốm đau và khỏe mạnh. Thật thú vị, dữ liệu từ đồng hồ thông minh được coi là phương pháp phát hiện sớm các bệnh nghiêm trọng, chẳng hạn như Covid-19 (xem “Sử dụng dữ liệu đồng hồ thông minh để phát hiện sớm các trường hợp COVID-19” tại JAMA Network.com để biết chi tiết).

 Bộ dữ liệu về cách mọi người liên hệ với nhau và điều gì thúc đẩy sự quan tâm của họ từ các nguồn như phương tiện truyền thông xã hội và công cụ tìm kiếm. Ví dụ, một nghiên cứu từ Trung tâm Tâm lý học của Đại học Cambridge tuyên bố rằng các tương tác trên Facebook chứa rất nhiều dữ liệu về các mối quan hệ thân mật (xem “Máy tính của bạn hiểu bạn hơn bạn bè của bạn, các nhà nghiên cứu nói” tại The Guardian.com).

 Thông tin về cách chúng ta nói được ghi lại bằng điện thoại di động. Ví dụ, OK Google, một chức năng có trên điện thoại di động Android, thường xuyên ghi lại các câu hỏi và đôi khi thậm chí còn nhiều hơn thế nữa, như đã giải thích trong “Google's been quietly recording your voice; here's how to listen to— and delete—the archive” tại Quartz.com.

Mỗi ngày, người dùng kết nối nhiều thiết bị hơn với Internet để bắt đầu lưu trữ dữ liệu cá nhân mới. Hiện nay có những trợ lý cá nhân được đặt trong nhà, chẳng hạn như Amazon Echo và các thiết bị nhà thông minh tích hợp khác

cung cấp các cách để điều chỉnh và tạo điều kiện thuận lợi cho môi trường trong nước. Đây chỉ là phần nổi của tảng băng chìm vì nhiều công cụ phổ biến khác trong cuộc sống hàng ngày đang được kết nối với nhau (từ tủ lạnh đến bàn chải đánh răng) và có thể xử lý, ghi lại và truyền dữ liệu. Internet vạn vật (IoT) đang trở thành hiện thực. Năm 2015, cơ sở thiết bị IoT được lắp đặt là 3,6 tỷ. Các chuyên gia ước tính rằng đến năm 2025, cơ sở thiết bị IoT được lắp đặt sẽ đạt 30,9 tỷ, tăng 858 phần trăm (xem “Có bao nhiêu thiết bị IoT vào năm 2021?” tại TechJury.net và “Kết nối thiết bị hoạt động Internet vạn vật (IoT) và không phải IoT trên toàn thế giới từ năm 2010 đến năm 2025 (tính bằng tỷ)” tại Statista.com.

***Đưa thuật toán vào hoạt động***

Loài người hiện đang ở ngã ba đường đáng kinh ngạc của khối lượng dữ liệu chưa từng có, được tạo ra bởi phần cứng ngày càng nhỏ hơn và mạnh hơn. Dữ liệu cũng ngày càng được xử lý và phân tích bởi chính những máy tính mà quá trình này giúp lan truyền và phát triển. Câu nói này có vẻ hiển nhiên, nhưng dữ liệu đã trở nên phổ biến đến mức giá trị của nó không còn chỉ nằm ở thông tin mà nó chứa (chẳng hạn như trường hợp dữ liệu được lưu trữ trong cơ sở dữ liệu của một công ty cho phép hoạt động hàng ngày của công ty), mà còn nằm ở việc sử dụng dữ liệu như một phương tiện để tạo ra các giá trị mới. Một số người gọi dữ liệu đó là "dầu mới". Những giá trị mới này chủ yếu tồn tại trong cách các ứng dụng xử lý, lưu trữ và truy xuất dữ liệu, cũng như cách bạn thực sự sử dụng dữ liệu đó thông qua các thuật toán thông minh.

Thuật toán và AI đã thay đổi trò chơi dữ liệu. Như đã đề cập trong chương trước, thuật toán AI đã thử nhiều cách tiếp cận khác nhau theo thứ tự sau:

* Thuật toán đơn giản
* Lý luận tượng trưng dựa trên logic
* Hệ thống chuyên gia

Trong những năm gần đây, các thuật toán AI đã chuyển sang mạng nơ-ron và ở dạng trưởng thành nhất của chúng là học sâu. Khi quá trình chuyển đổi phương pháp này diễn ra, dữ liệu đã chuyển từ thông tin được xử lý bởi các thuật toán được xác định trước sang trở thành thứ định hình nên thuật toán thành

một cái gì đó hữu ích cho nhiệm vụ. Dữ liệu chuyển từ chỉ là nguyên liệu thô cung cấp nhiên liệu cho giải pháp thành nghệ nhân của chính giải pháp đó, như thể hiện trong [Hình 2-2](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-02-defining-the-role-of-data/#page54) .

A graph showing the value of deep learning

Description automatically generated

[**HÌNH 2-2:**](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-02-defining-the-role-of-data/#page54)Với các giải pháp AI hiện tại, nhiều dữ liệu hơn đồng nghĩa với nhiều trí thông minh hơn.

Do đó, ảnh chụp một số chú mèo con của bạn ngày càng trở nên hữu ích không chỉ vì giá trị tình cảm của nó - miêu tả những chú mèo nhỏ dễ thương của bạn - mà còn vì nó có thể trở thành một phần trong quá trình học tập của AI để khám phá những khái niệm chung hơn, chẳng hạn như đặc điểm nào biểu thị một chú mèo hoặc hiểu thế nào là dễ thương.

Ở quy mô lớn hơn, một công ty như Google cung cấp thuật toán của mình từ dữ liệu có sẵn miễn phí, chẳng hạn như nội dung của các trang web hoặc văn bản có trong các văn bản và sách có sẵn công khai. Phần mềm nhện của Google thu thập dữ liệu trên web, nhảy từ trang web này sang trang web khác, truy xuất các trang web có nội dung văn bản và hình ảnh của chúng. Ngay cả khi Google trả lại một phần dữ liệu cho người dùng dưới dạng kết quả tìm kiếm, thì nó vẫn trích xuất các loại thông tin khác từ dữ liệu bằng các thuật toán AI của mình, học hỏi từ dữ liệu đó cách đạt được các mục tiêu khác.

Các thuật toán xử lý từ ngữ có thể giúp hệ thống AI của Google hiểu và dự đoán nhu cầu của bạn ngay cả khi bạn không diễn đạt chúng bằng một tập hợp từ khóa mà bằng ngôn ngữ tự nhiên rõ ràng, không rõ ràng, ngôn ngữ mà chúng ta nói hàng ngày (và đúng vậy, ngôn ngữ hàng ngày thường không rõ ràng). Nếu hiện tại bạn đang cố gắng đặt câu hỏi, không chỉ là chuỗi từ khóa, cho công cụ tìm kiếm Google, bạn sẽ nhận thấy rằng nó có xu hướng trả lời đúng. Kể từ năm 2012, với sự ra mắt của bản cập nhật Hummingbird (đọc thông tin chi tiết trong “Câu hỏi thường gặp: Tất cả về thuật toán 'Hummingbird' mới của Google” tại Search Engine Land.com), Google đã dần có khả năng hiểu các từ đồng nghĩa và khái niệm tốt hơn, điều này vượt ra ngoài dữ liệu ban đầu mà nó thu thập được và đây là kết quả của một quy trình AI.

Vài năm sau Hummingbird, Google đã triển khai một thuật toán thậm chí còn tiên tiến hơn có tên là RankBrain (“Câu hỏi thường gặp: Tất cả về thuật toán Google RankBrain” tại Search Engine Land.com), thuật toán này học trực tiếp từ hàng triệu truy vấn mỗi ngày và có thể trả lời các truy vấn tìm kiếm mơ hồ hoặc không rõ ràng, thậm chí được diễn đạt bằng tiếng lóng hoặc thuật ngữ thông tục hoặc chỉ đơn giản là đầy lỗi. RankBrain không phục vụ tất cả các truy vấn, nhưng nó học từ dữ liệu về cách trả lời truy vấn tốt hơn. Sau khi giới thiệu vào năm 2015, nó nhanh chóng bắt đầu xử lý 15 phần trăm các truy vấn của công cụ tìm kiếm, nhưng nó sẽ không thay thế Hummingbird (xem “7 điều bạn có thể chưa biết về RankBrain của Google” tại Act On.com và “Sự kết hợp giữa Rankbrain, Hummingbird và Tối ưu hóa công cụ tìm kiếm” tại Connectica.com để biết chi tiết).

## 2. Vấn đề thu thập và sử dụng dữ liệu

Có dữ liệu dồi dào là không đủ để tạo ra một AI thành công. Hiện tại, thuật toán AI không thể trích xuất thông tin trực tiếp từ dữ liệu thô. Hầu hết các thuật toán đều dựa vào việc thu thập và xử lý bên ngoài trước khi phân tích. Khi một thuật toán thu thập thông tin hữu ích, nó có thể không biểu diễn đúng thông tin. Các phần sau đây giúp bạn hiểu cách thu thập, xử lý và tự động hóa việc thu thập dữ liệu theo góc nhìn tổng quan.

***Xem xét các nguồn dữ liệu***

Dữ liệu bạn sử dụng đến từ nhiều nguồn. Nguồn dữ liệu phổ biến nhất là từ thông tin do con người nhập vào tại một thời điểm nào đó. Ngay cả khi hệ thống tự động thu thập dữ liệu trang web mua sắm, con người vẫn là người đầu tiên nhập thông tin. Con người nhấp vào nhiều mặt hàng khác nhau, thêm chúng vào giỏ hàng, chỉ định các đặc điểm (như kích thước và số lượng), rồi thanh toán. Sau đó, sau khi bán hàng, con người sẽ đánh giá trải nghiệm mua sắm, sản phẩm và phương thức giao hàng và đưa ra nhận xét. Tóm lại, mọi trải nghiệm mua sắm cũng trở thành một bài tập thu thập dữ liệu.

Nhiều nguồn dữ liệu ngày nay dựa vào dữ liệu đầu vào thu thập được từ nguồn con người. Con người cũng cung cấp dữ liệu đầu vào thủ công. Bạn gọi điện hoặc đến một văn phòng nào đó để đặt lịch hẹn với một chuyên gia. Sau đó, một nhân viên lễ tân sẽ thu thập thông tin từ bạn cần thiết cho cuộc hẹn. Dữ liệu được thu thập thủ công này cuối cùng sẽ nằm trong một tập dữ liệu ở đâu đó để phục vụ mục đích phân tích.

Dữ liệu cũng được thu thập từ các cảm biến và các cảm biến này có thể có hầu như bất kỳ hình thức nào. Ví dụ, nhiều tổ chức dựa trên việc thu thập dữ liệu vật lý, chẳng hạn như số lượng người xem một vật thể trong cửa sổ, vào việc phát hiện điện thoại di động. Phần mềm nhận dạng khuôn mặt có khả năng phát hiện ra khách hàng quay lại.

Tuy nhiên, các cảm biến có thể tạo ra các tập dữ liệu từ hầu như bất cứ thứ gì. Dịch vụ thời tiết dựa vào các tập dữ liệu được tạo ra bởi các cảm biến theo dõi môi trường

các điều kiện như mưa, nhiệt độ, độ ẩm, mây che phủ, v.v. Hệ thống giám sát bằng rô-bốt giúp sửa những lỗi nhỏ trong hoạt động của rô-bốt bằng cách liên tục phân tích dữ liệu thu thập được bởi các cảm biến giám sát. Một cảm biến, kết hợp với một ứng dụng AI nhỏ, có thể cho bạn biết khi nào bữa tối của bạn được nấu chín hoàn hảo vào tối nay. Cảm biến thu thập dữ liệu, nhưng ứng dụng AI sử dụng các quy tắc để giúp xác định thời điểm thức ăn được nấu chín đúng cách.

***Thu thập dữ liệu đáng tin cậy***

Từ *đáng tin cậy* có vẻ dễ định nghĩa, nhưng lại rất khó thực hiện. Một thứ gì đó đáng tin cậy khi kết quả mà nó tạo ra vừa được mong đợi vừa nhất quán. Một nguồn dữ liệu đáng tin cậy tạo ra dữ liệu tầm thường không chứa bất kỳ điều bất ngờ nào; không ai bị sốc chút nào bởi kết quả. Tùy thuộc vào quan điểm của bạn, thực tế có thể là một điều tốt khi hầu hết mọi người không ngáp dài rồi ngủ thiếp đi khi xem xét dữ liệu. Những điều bất ngờ khiến dữ liệu đáng để phân tích và xem xét. Do đó, dữ liệu có khía cạnh tính hai mặt. Chúng ta muốn dữ liệu đáng tin cậy, tầm thường, được dự đoán đầy đủ, chỉ xác nhận những gì chúng ta đã biết, nhưng điều bất ngờ chính là điều khiến việc thu thập dữ liệu trở nên hữu ích ngay từ đầu.

Tuy nhiên, bạn không muốn dữ liệu quá khác thường đến mức gần như trở nên đáng sợ khi xem xét. Cần duy trì sự cân bằng khi thu thập dữ liệu. Dữ liệu phải nằm trong một số giới hạn nhất định (như được mô tả trong phần “ [Manicuring the Data](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-02-defining-the-role-of-data/#page62) ” (Xử lý dữ liệu), sau trong chương này). Dữ liệu cũng phải đáp ứng [các tiêu chí cụ thể về giá trị chân lý (như được mô tả trong phần “Xem xét năm điều không đúng trong dữ liệu”, sau trong chương này). Dữ liệu cũng phải đến](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-02-defining-the-role-of-data/#page66) theo các khoảng thời gian mong đợi và tất cả các trường của bản ghi dữ liệu đến phải đầy đủ.

\*Ở một mức độ nào đó, bảo mật dữ liệu cũng ảnh hưởng đến độ tin cậy của dữ liệu. Tính nhất quán của dữ liệu có nhiều dạng. Khi dữ liệu đến, bạn có thể đảm bảo rằng dữ liệu nằm trong phạm vi mong đợi và xuất hiện ở một dạng cụ thể. Tuy nhiên, sau khi bạn lưu trữ dữ liệu, độ tin cậy có thể giảm trừ khi bạn đảm bảo rằng dữ liệu vẫn ở dạng mong đợi. Một thực thể can thiệp vào dữ liệu sẽ ảnh hưởng đến độ tin cậy, khiến dữ liệu

nghi ngờ và có khả năng không sử dụng được để phân tích sau này. Đảm bảo độ tin cậy của dữ liệu có nghĩa là sau khi dữ liệu đến, không ai can thiệp vào dữ liệu để làm cho nó phù hợp với phạm vi mong đợi (làm cho nó trở nên tầm thường).

***Làm cho đầu vào của con người đáng tin cậy hơn***

Con người mắc lỗi — đó là một phần của con người. Trên thực tế, việc mong đợi rằng con người sẽ không mắc lỗi là điều không hợp lý. Tuy nhiên, nhiều thiết kế ứng dụng cho rằng con người bằng cách nào đó sẽ không mắc lỗi nào cả. Thiết kế này mong đợi rằng mọi người sẽ chỉ đơn giản là tuân theo các quy tắc. Thật không may, phần lớn người dùng chắc chắn sẽ không đọc các quy tắc vì hầu hết con người cũng lười biếng hoặc quá vội vã khi phải làm những việc không thực sự giúp ích trực tiếp cho họ.

Hãy xem xét việc nhập một tiểu bang vào một biểu mẫu. Nếu bạn chỉ cung cấp một trường văn bản, một số người dùng có thể nhập toàn bộ tên tiểu bang, chẳng hạn như Kansas. Tất nhiên, một số người dùng sẽ mắc lỗi đánh máy hoặc lỗi viết hoa và đưa ra Kanzuz, Kansus hoặc kANSAS. Bỏ qua những lỗi này, mọi người và tổ chức cũng có nhiều cách tiếp cận khác nhau để thực hiện nhiệm vụ. Một người nào đó trong ngành xuất bản có thể sử dụng hướng dẫn về phong cách của Associated Press (AP) và nhập Kan. Một người nào đó lớn tuổi hơn và quen với các hướng dẫn của Văn phòng In ấn Chính phủ (GPO) có thể nhập Kans. thay thế. Các chữ viết tắt khác cũng được sử dụng. Bưu điện Hoa Kỳ (USPS) sử dụng KS, nhưng Cảnh sát biển Hoa Kỳ sử dụng KA. Trong khi đó, biểu mẫu của Tổ chức Tiêu chuẩn Quốc tế (ISO) sử dụng US-KS. Lưu ý, đây chỉ là mục nhập tiểu bang, khá đơn giản — hoặc ít nhất là bạn đã nghĩ như vậy trước khi đọc phần này. Rõ ràng là vì tiểu bang sẽ không sớm đổi tên nên bạn có thể chỉ cần cung cấp hộp danh sách thả xuống trên biểu mẫu để chọn tiểu bang theo định dạng bắt buộc, do đó loại bỏ sự khác biệt trong cách viết tắt, lỗi đánh máy và lỗi viết hoa chỉ trong một lần.

\*Hộp danh sách thả xuống hoạt động tốt với một mảng dữ liệu đầu vào tuyệt vời và việc sử dụng chúng đảm bảo rằng dữ liệu đầu vào của con người vào các trường đó trở nên cực kỳ đáng tin cậy vì con người không có lựa chọn nào khác ngoài việc sử dụng một trong các mục nhập mặc định. Tất nhiên, con người luôn có thể chọn mục nhập không chính xác, đó là lúc kiểm tra lại phát huy tác dụng. Một số ứng dụng mới hơn so sánh mã bưu chính với các mục nhập thành phố và tiểu bang để xem chúng có khớp không. Khi chúng không khớp (đôi khi chỉ là vấn đề viết hoa), người dùng lại được yêu cầu cung cấp dữ liệu đầu vào chính xác. Việc kiểm tra lại này gần như gây khó chịu, nhưng người dùng không có khả năng nhìn thấy nó thường xuyên, vì vậy nó không nên trở nên quá khó chịu.

Ngay cả với các kiểm tra chéo và mục nhập tĩnh, con người vẫn có nhiều chỗ để mắc lỗi. Ví dụ, nhập số có thể gây ra vấn đề. Khi người dùng cần nhập 2.00, bạn có thể thấy 2, hoặc 2.0, hoặc 2., hoặc bất kỳ mục nhập nào khác. May mắn thay, việc phân tích cú pháp mục nhập và định dạng lại sẽ khắc phục được sự cố và bạn có thể thực hiện tác vụ này tự động mà không cần sự trợ giúp của người dùng. (Thật không may, một số trang web trực tuyến muốn bạn nhập thông tin như thẻ tín dụng có dấu gạch ngang, một số có khoảng trắng và một số không có khoảng trắng nào cả, điều này khiến phiên làm việc trở nên rất khó hiểu khi ứng dụng không tự động sửa mục nhập.)

Thật không may, việc định dạng lại sẽ không sửa được lỗi nhập số. Bạn có thể giảm thiểu một phần các lỗi như vậy bằng cách bao gồm các kiểm tra phạm vi. Khách hàng không thể mua –5 thanh xà phòng. Và, trừ khi khách hàng thực sự bẩn hoặc sở hữu một trang trại gấu túi, việc nhập 50.000 thanh xà phòng cũng có thể là một lỗi. Cách hợp pháp để chứng minh rằng khách hàng đang trả lại các thanh xà phòng là xử lý việc trả lại, không phải là bán hàng. Tuy nhiên, người dùng có thể chỉ đơn giản là đã mắc lỗi và bạn có thể cung cấp thông báo nêu rõ phạm vi nhập phù hợp cho giá trị.

***Sử dụng thu thập dữ liệu tự động***

Một số người cho rằng thu thập dữ liệu tự động giải quyết được mọi vấn đề đầu vào của con người liên quan đến tập dữ liệu. Trên thực tế, thu thập dữ liệu tự động mang lại một số lợi ích:

 Tính nhất quán tốt hơn

 Độ tin cậy được cải thiện

 Xác suất thiếu dữ liệu thấp hơn

 Độ chính xác được nâng cao

 Giảm độ biến thiên cho những thứ như đầu vào được định thời gian

Thật không may, nói rằng thu thập dữ liệu tự động giải quyết mọi vấn đề là hoàn toàn không đúng. Thu thập dữ liệu tự động vẫn dựa vào các cảm biến, ứng dụng và phần cứng máy tính do con người thiết kế, chỉ cung cấp quyền truy cập vào dữ liệu mà con người quyết định cho phép. Do những giới hạn mà con người đặt ra đối với các đặc điểm của thu thập dữ liệu tự động, kết quả thường cung cấp ít thông tin hữu ích hơn so với mong đợi của các nhà thiết kế. Do đó, thu thập dữ liệu tự động luôn trong tình trạng thay đổi liên tục khi các nhà thiết kế cố gắng giải quyết các vấn đề đầu vào.

Thu thập dữ liệu tự động cũng gặp phải lỗi phần mềm và phần cứng có trong bất kỳ hệ thống máy tính nào, nhưng có khả năng xảy ra *các vấn đề mềm* cao hơn (phát sinh khi hệ thống rõ ràng đang hoạt động nhưng không cung cấp kết quả mong muốn) so với các loại thiết lập dựa trên máy tính khác. Khi hệ thống hoạt động, độ tin cậy của dữ liệu đầu vào vượt xa khả năng của con người. Tuy nhiên, khi xảy ra các vấn đề mềm, hệ thống thường không nhận ra, như con người có thể, rằng có vấn đề và do đó, tập dữ liệu có thể chứa nhiều dữ liệu tầm thường hoặc thậm chí là dữ liệu xấu.

***Thu thập dữ liệu cá nhân một cách có đạo đức***

Đối với một số người, bất kỳ thứ gì xuất hiện trên Internet đều tự động được coi là phạm vi công cộng — bao gồm khuôn mặt của mọi người và tất cả thông tin cá nhân của họ. Thực tế là bạn nên coi mọi thứ là có bản quyền và không được phép sử dụng theo cách phạm vi công cộng để sử dụng dữ liệu một cách an toàn. Ngay cả những người nhận ra rằng tài liệu có bản quyền cũng thường sẽ quay lại với các nguyên tắc sử dụng hợp lý (như đã giải thích trong “Sử dụng tài liệu có bản quyền” tại Baylor University.edu). Sử dụng hợp lý có thể là một chủ đề rất khó hiểu, như đã chứng kiến ​​trong vụ kiện của Author's Guild kiện Google (xem “Quyết định quan trọng nhất của tòa án đối với khoa học dữ liệu và học máy” tại Towards Data Science.com) cuối cùng đã được quyết định có lợi cho Google, nhưng chỉ vì Google đã đáp ứng một số yêu cầu rất nghiêm ngặt. Trong

Ngoài ra, loại sử dụng hợp lý này chỉ áp dụng cho sách chứ không phải cho con người.

Vấn đề khi chỉ xem xét sử dụng hợp lý là việc xem xét quyền riêng tư của một người cũng rất cần thiết (bạn có thể đọc về nhiều luật khác nhau trong “Luật riêng tư trên Internet được tiết lộ — thông tin cá nhân của bạn được bảo vệ như thế nào trực tuyến” tại Thomson Reuters Legal.com). Do đó, không có gì ngạc nhiên khi một cuộc náo động lớn nổ ra khi các công ty bắt đầu thu thập hình ảnh những người đeo khẩu trang trên Internet mà không xin phép bất kỳ ai (xem “Ảnh tự sướng đeo khẩu trang của bạn có thể đào tạo công cụ nhận dạng khuôn mặt tiếp theo” tại Cnet.com). Trên thực tế, Facebook đang bị kiện vì sử dụng sai dữ liệu người dùng (xem “Vụ kiện dữ liệu khuôn mặt trị giá 35 tỷ đô la chống lại Facebook sẽ được tiến hành” tại TechCrunch.com).

Quyền riêng tư cũng đã tạo ra một ngành công nghiệp mới để khiến khuôn mặt của một người trở nên ít hữu ích hơn đối với các công ty quyết tâm lấy dữ liệu miễn phí mà không được phép bằng mọi cách có thể (xem bài viết "Công cụ này có thể bảo vệ ảnh của bạn khỏi nhận dạng khuôn mặt" *của tờ New York Times* ). Thực tế là, bất kể bạn đứng ở đâu về vấn đề sử dụng miễn phí, bạn vẫn cần cân nhắc đến việc sử dụng dữ liệu có đạo đức mà bạn có được bất kể nguồn dữ liệu đó là gì. Sau đây là một số cân nhắc cần lưu ý khi bạn thu thập dữ liệu cá nhân một cách có đạo đức:

**Xin phép:** Một số nghiên cứu sẽ yêu cầu bạn phải có khả năng xác định những người được sử dụng trong một tập dữ liệu. Việc ra ngoài và lấy Thông tin nhận dạng cá nhân (PII) không phải là cách tốt để thu thập dữ liệu. Trước hết, bạn không thể chắc chắn rằng thông tin đó đầy đủ hoặc chính xác, vì vậy bất kỳ phân tích nào bạn thực hiện đều đáng ngờ. Mặt khác, bạn có thể gặp phải hậu quả lộn xộn và tốn kém của các hành động pháp lý. Cách tốt nhất để lấy dữ liệu bằng PII là xin phép. Bạn có thể tìm thấy một số tài nguyên trực tuyến để xin phép, ví dụ như ở cấp chính phủ, bằng cách tìm đúng tài nguyên, chẳng hạn như "Cách xin phép người tiêu dùng trước khi truy cập vào Thông tin nhận dạng cá nhân (PII)".

**Sử dụng các kỹ thuật vệ sinh:***Vệ sinh dữ liệu* bao gồm việc xóa thông tin cá nhân, chẳng hạn như tên, địa chỉ, số điện thoại, ID, v.v. khỏi một tập dữ liệu để việc xác định một cá nhân cụ thể trong một tập dữ liệu trở nên không thể. Ngoài các biến văn bản và tập dữ liệu,

bạn phải cân nhắc mọi loại dữ liệu. Ví dụ, nếu bạn đang làm việc với các bộ sưu tập ảnh, điều tối quan trọng là bạn phải thực hiện các bước để làm mờ khuôn mặt và xóa biển số xe khỏi ảnh. Thật kỳ lạ, nếu bạn thực hiện tìm kiếm trên Google bằng cách sử dụng *vệ sinh* làm thuật ngữ chính, bạn vẫn nhận được nhiều liên kết liên quan đến việc vệ sinh. Thêm *quyền riêng tư* làm thuật ngữ chính khác giúp tìm ra loại bài viết bạn thực sự cần. Mặc dù vậy, bạn có thể thấy rằng bạn không phải là người duy nhất bối rối về quy trình này.

**Tránh suy luận dữ liệu:** Khi thu thập dữ liệu, một số người dùng sẽ từ chối chia sẻ thông tin nhận dạng cá nhân, chẳng hạn như giới tính và tuổi tác. Một khuyến nghị là suy luận thông tin này khi có ảnh hoặc thông tin khác của người dùng. Thật không may, tên được liên kết với một giới tính trong một nền văn hóa cụ thể có thể được gán cho giới tính khác trong các nền văn hóa khác. Vấn đề với suy luận tuổi tác thậm chí còn sâu sắc hơn. Ví dụ, thuật toán học máy có thể suy ra độ tuổi sai của người bạch tạng, có thể ảnh hưởng đến một trong 3.000 cá nhân, tùy thuộc vào nơi dữ liệu đến (xem chi tiết trong “Bản tin thông tin – Bạch tạng là gì?” tại NOAH.com.)

**Tránh khái quát hóa:** Nhiều lĩnh vực nghiên cứu ngày nay cố gắng áp dụng sai số liệu thống kê và kết quả học máy, dẫn đến việc một cá nhân bị đối xử tệ theo cách nào đó. Điều quan trọng cần nhớ là số liệu thống kê áp dụng cho nhóm, không phải cho cá nhân.

## 3. Vấn đề xử lý dữ liệu

Một số người sử dụng thuật ngữ *thao túng* khi nói về dữ liệu, tạo ấn tượng rằng dữ liệu bằng cách nào đó bị thay đổi theo cách vô đạo đức hoặc gian xảo. Có lẽ một thuật ngữ tốt hơn sẽ là *làm móng,* làm cho dữ liệu được định hình tốt và đẹp. Tuy nhiên, bất kể bạn sử dụng thuật ngữ nào, dữ liệu thô hiếm khi đáp ứng các yêu cầu để xử lý và phân tích. Để có được thứ gì đó từ dữ liệu, bạn phải làm móng để đáp ứng các nhu cầu cụ thể. Các phần sau đây thảo luận về nhu cầu làm móng dữ liệu.

***Xử lý dữ liệu bị thiếu***

Để trả lời đúng một câu hỏi nhất định, bạn phải có tất cả các dữ kiện. Bạn có thể đoán câu trả lời cho một câu hỏi mà không cần tất cả các dữ kiện, nhưng sau đó, câu trả lời đó có khả năng sai cũng như đúng. Thông thường, một người đưa ra quyết định, về cơ bản là trả lời một câu hỏi, mà không có tất cả các dữ kiện được cho là đã vội vàng đưa ra kết luận. Khi phân tích dữ liệu, có lẽ bạn đã vội vàng đưa ra nhiều kết luận hơn bạn nghĩ vì thiếu dữ liệu. Một *bản ghi dữ liệu,* một mục nhập trong *tập dữ liệu* (là tất cả dữ liệu), bao gồm *các trường* chứa các dữ kiện được sử dụng để trả lời một câu hỏi. Mỗi trường chứa một loại dữ liệu duy nhất giải quyết một dữ kiện duy nhất. Nếu trường đó trống, bạn không có dữ liệu cần thiết để trả lời câu hỏi bằng bản ghi dữ liệu cụ thể đó.

\*Là một phần của quá trình xử lý dữ liệu bị thiếu, bạn phải biết rằng dữ liệu bị thiếu. Việc xác định rằng tập dữ liệu của bạn bị thiếu thông tin thực sự có thể khá khó khăn vì nó đòi hỏi bạn phải xem xét dữ liệu ở cấp độ thấp — điều mà hầu hết mọi người không chuẩn bị để làm và tốn thời gian ngay cả khi bạn có các kỹ năng cần thiết. Thông thường, manh mối đầu tiên của bạn về việc dữ liệu bị thiếu là các câu trả lời vô lý mà các câu hỏi của bạn nhận được từ thuật toán và tập dữ liệu liên quan. Khi thuật toán là thuật toán phù hợp để sử dụng, thì tập dữ liệu phải có lỗi. Sau đây là một số vấn đề cần cân nhắc:

**Thiếu dữ liệu cần thiết:** Có thể xảy ra sự cố khi thu thập dữ liệu

quá trình không bao gồm tất cả dữ liệu cần thiết để trả lời một câu hỏi cụ thể. Đôi khi bạn nên bỏ một sự thật thay vì sử dụng một sự thật bị hư hỏng đáng kể.

**Một số dữ liệu bị thiếu:** Các trường ít bị hỏng hơn có thể bị mất dữ liệu theo một trong hai cách, ngẫu nhiên hoặc tuần tự, như được mô tả tại đây:

**Dữ liệu bị thiếu ngẫu nhiên thường là kết quả của lỗi của con người hoặc cảm biến.** Sửa dữ liệu bị thiếu ngẫu nhiên là cách dễ nhất. Bạn có thể sử dụng giá trị trung bình hoặc trung vị đơn giản để thay thế. Không, tập dữ liệu không hoàn toàn chính xác, nhưng có khả năng sẽ hoạt động đủ tốt để có được câu trả lời hợp lý.

**Dữ liệu bị mất tuần tự xảy ra trong một số loại lỗi tổng quát.** Việc sửa dữ liệu bị mất tuần tự khó hơn đáng kể, nếu không muốn nói là không thể, vì bạn không có bất kỳ dữ liệu xung quanh nào để dựa vào đó để đưa ra bất kỳ phỏng đoán nào. Nếu bạn có thể tìm ra nguyên nhân của dữ liệu bị mất, đôi khi bạn có thể tái tạo lại nó.

***Xem xét sự không khớp dữ liệu***

Dữ liệu có thể tồn tại cho từng bản ghi dữ liệu trong một tập dữ liệu, nhưng nó có thể không phù hợp với dữ liệu khác trong các tập dữ liệu khác mà bạn sở hữu. Ví dụ, dữ liệu số trong một trường trong một tập dữ liệu có thể là kiểu dấu phẩy động (có dấu thập phân), nhưng là kiểu số nguyên trong tập dữ liệu khác. Trước khi bạn có thể kết hợp hai tập dữ liệu, các trường phải chứa cùng một kiểu dữ liệu.

Có thể xảy ra nhiều loại sai lệch khác. Ví dụ, các trường ngày tháng nổi tiếng là được định dạng theo nhiều cách khác nhau. Để so sánh các ngày tháng, định dạng dữ liệu phải giống nhau. Tuy nhiên, ngày tháng cũng rất nguy hiểm vì chúng có xu hướng trông giống nhau nhưng thực tế lại không giống nhau. Ví dụ, ngày tháng trong một tập dữ liệu có thể sử dụng Giờ trung bình Greenwich (GMT) làm cơ sở, trong khi ngày tháng trong một tập dữ liệu khác có thể sử dụng một múi giờ khác. Trước khi bạn có thể so sánh thời gian, bạn phải căn chỉnh chúng theo cùng một múi giờ. Thậm chí còn kỳ lạ hơn khi ngày tháng trong một tập dữ liệu đến từ một vị trí sử dụng Giờ tiết kiệm ánh sáng ban ngày (DST), nhưng ngày tháng từ một vị trí khác thì không.

Ngay cả khi kiểu dữ liệu và định dạng giống nhau, vẫn có thể xảy ra tình trạng không căn chỉnh dữ liệu khác. Ví dụ, các trường trong một tập dữ liệu có thể không khớp với các trường trong tập dữ liệu khác. Trong một số trường hợp, những khác biệt này rất dễ sửa. Một tập dữ liệu có thể coi tên và họ là một trường duy nhất, trong khi một tập dữ liệu khác có thể sử dụng các trường riêng biệt cho tên và họ. Câu trả lời là thay đổi tất cả các tập dữ liệu để sử dụng một trường duy nhất hoặc thay đổi tất cả chúng để sử dụng các trường riêng biệt cho tên và họ. Thật không may, nhiều tình trạng không căn chỉnh trong nội dung dữ liệu khó phát hiện hơn. Trên thực tế, hoàn toàn có khả năng bạn không thể phát hiện ra chúng. Tuy nhiên, trước khi từ bỏ, hãy cân nhắc các giải pháp tiềm năng sau cho vấn đề này:

 Tính toán dữ liệu còn thiếu từ dữ liệu khác mà bạn có thể truy cập.

 Xác định vị trí dữ liệu bị thiếu trong tập dữ liệu khác.

 Kết hợp các tập dữ liệu để tạo thành một tổng thể cung cấp các trường nhất quán.

 Thu thập dữ liệu bổ sung từ nhiều nguồn khác nhau để bổ sung dữ liệu còn thiếu.

 Xác định lại câu hỏi của bạn để không còn cần đến dữ liệu còn thiếu nữa.

***Tách dữ liệu hữu ích khỏi dữ liệu khác***

Một số tổ chức cho rằng họ không bao giờ có thể có quá nhiều dữ liệu, nhưng việc dư thừa dữ liệu cũng trở thành vấn đề như không đủ dữ liệu. Để giải quyết vấn đề một cách hiệu quả, AI chỉ cần đủ dữ liệu. Việc xác định câu hỏi mà bạn muốn trả lời một cách ngắn gọn và rõ ràng sẽ giúp ích, cũng như sử dụng thuật toán chính xác (hoặc nhóm thuật toán). Tất nhiên, vấn đề chính khi có quá nhiều dữ liệu là việc tìm ra giải pháp (sau khi lội qua tất cả dữ liệu bổ sung đó) mất nhiều thời gian hơn và đôi khi bạn nhận được kết quả khó hiểu vì bạn không thể nhìn thấy toàn cảnh vì chỉ thấy cây.

\*Là một phần của việc tạo tập dữ liệu bạn cần để phân tích, bạn tạo một bản sao của dữ liệu gốc thay vì sửa đổi nó. Luôn giữ nguyên dữ liệu gốc, thô để bạn có thể sử dụng nó cho các phân tích khác sau này. Ngoài ra, việc tạo đầu ra dữ liệu phù hợp để phân tích có thể yêu cầu một số lần thử vì bạn có thể thấy rằng đầu ra không đáp ứng

nhu cầu của bạn. Vấn đề là tạo một tập dữ liệu chỉ chứa dữ liệu cần thiết để phân tích, nhưng hãy nhớ rằng dữ liệu có thể cần các loại cắt tỉa cụ thể để đảm bảo đầu ra mong muốn.

## 4. Tìm hiểu về 05 sai lầm trong dữ liệu

Con người thường nhìn nhận dữ liệu theo đúng bản chất của nó trong nhiều trường hợp: một ý kiến. Trên thực tế, trong một số trường hợp, mọi người bóp méo dữ liệu đến mức nó trở nên vô dụng, một *sự sai lệch.* Máy tính không thể phân biệt được dữ liệu đúng và không đúng — tất cả những gì nó nhìn thấy là dữ liệu. Một trong những vấn đề khiến việc tạo ra một AI thực sự suy nghĩ như con người trở nên khó khăn, nếu không muốn nói là không thể, là con người có thể làm việc với sự sai lệch còn máy tính thì không. Điều tốt nhất bạn có thể hy vọng đạt được là xem dữ liệu sai lệch như dữ liệu ngoại lệ rồi lọc ra, nhưng kỹ thuật đó không nhất thiết giải quyết được vấn đề vì con người vẫn sẽ sử dụng dữ liệu và cố gắng xác định sự thật dựa trên những sự sai lệch có ở đó.

\* Một suy nghĩ phổ biến về việc tạo ra các tập dữ liệu ít bị ô nhiễm hơn là thay vì cho phép con người nhập dữ liệu, việc thu thập dữ liệu thông qua cảm biến hoặc các phương tiện khác nên khả thi. Thật không may, cảm biến và các phương pháp nhập dữ liệu cơ học khác phản ánh mục tiêu của những người phát minh ra chúng và giới hạn của những gì công nghệ cụ thể có thể phát hiện. Do đó, ngay cả dữ liệu có nguồn gốc từ máy móc hoặc cảm biến cũng có thể tạo ra sự sai lệch mà AI rất khó phát hiện và khắc phục.

Các phần sau đây sử dụng một vụ tai nạn xe hơi làm ví dụ chính để minh họa năm loại thông tin sai lệch có thể xuất hiện trong dữ liệu. Các khái niệm mà vụ tai nạn đang cố gắng mô tả có thể không phải lúc nào cũng xuất hiện trong dữ liệu và chúng có thể xuất hiện theo những cách khác với những gì đã thảo luận. Thực tế vẫn là bạn thường cần xử lý những loại thông tin này khi xem dữ liệu.

***Nhiệm vụ***

Sự sai lầm của ủy ban là những sự sai lầm phản ánh nỗ lực trắng trợn nhằm thay thế thông tin đúng sự thật cho thông tin không đúng sự thật. Ví dụ, khi điền vào báo cáo tai nạn, một người nào đó có thể nói rằng mặt trời làm họ bị lóa mắt trong giây lát, khiến họ không thể nhìn thấy người mà họ đâm phải. Trên thực tế, có lẽ người đó bị phân tâm bởi điều gì đó khác hoặc thực sự không nghĩ đến việc lái xe (có thể đang cân nhắc đến một bữa tối ngon lành). Nếu không ai có thể bác bỏ lý thuyết này, người đó có thể được giảm phí. Tuy nhiên, vấn đề là dữ liệu cũng sẽ bị ô nhiễm. Hậu quả là bây giờ một công ty bảo hiểm sẽ căn cứ vào dữ liệu sai lệch để tính phí bảo hiểm.

\* Mặc dù có vẻ như việc nói dối trắng trợn có thể tránh được hoàn toàn, nhưng thường thì không phải vậy. Con người nói "những lời nói dối trắng trợn" để tránh làm người khác xấu hổ hoặc để giải quyết vấn đề với ít nỗ lực cá nhân nhất. Đôi khi việc nói dối trắng trợn dựa trên thông tin đầu vào sai lệch hoặc tin đồn. Trên thực tế, có rất nhiều nguồn gây ra lỗi nói dối trắng trợn đến mức thực sự khó có thể đưa ra một kịch bản mà ai đó có thể tránh được hoàn toàn. Tuy nhiên, nói dối trắng trợn là một loại nói dối trắng trợn mà ai đó có thể tránh được thường xuyên hơn là không.

***Bỏ sót***

Sự thiếu sót hoặc sự thật là khi một người nói sự thật trong mọi sự kiện đã nêu, nhưng lại bỏ sót một sự kiện quan trọng có thể thay đổi nhận thức về toàn bộ sự việc. Nghĩ lại về báo cáo tai nạn, giả sử bạn đâm phải một con nai, gây ra thiệt hại đáng kể cho chiếc xe của bạn. Bạn nói sự thật rằng đường ướt; trời gần chạng vạng nên ánh sáng không tốt như mong đợi; bạn hơi chậm khi nhấn phanh; và con nai chỉ chạy ra từ một bụi cây bên lề đường. Kết luận sẽ là sự cố chỉ đơn giản là một tai nạn.

Tuy nhiên, bạn đã bỏ sót một sự thật quan trọng: Bạn đã nhắn tin vào thời điểm đó. Nếu cơ quan thực thi pháp luật biết về tin nhắn, họ sẽ thay đổi lý do gây tai nạn thành lái xe mất tập trung. Bạn có thể bị phạt và người điều chỉnh bảo hiểm sẽ sử dụng một lý do khác khi nhập sự cố vào cơ sở dữ liệu. Cũng giống như sự gian dối của hoa hồng, dữ liệu sai lệch kết quả sẽ thay đổi cách công ty bảo hiểm điều chỉnh phí bảo hiểm.

\* Việc tránh sự sai lệch hoặc thiếu sót là gần như không thể. Đúng, mọi người có thể cố tình bỏ sót sự thật trong báo cáo, nhưng khả năng họ quên đưa tất cả sự thật vào cũng cao. Xét cho cùng, hầu hết mọi người đều khá bàng hoàng sau một vụ tai nạn, vì vậy rất dễ mất tập trung và chỉ báo cáo những sự thật để lại ấn tượng quan trọng nhất. Ngay cả khi sau đó một người nhớ thêm các chi tiết và báo cáo chúng, thì cơ sở dữ liệu cũng không bao giờ có thể chứa một tập hợp đầy đủ các sự thật.

***Sai lệch về góc nhìn***

Sự sai lệch về góc nhìn xảy ra khi nhiều bên xem xét một sự cố từ nhiều góc nhìn. Ví dụ, khi xem xét một vụ tai nạn liên quan đến người đi bộ bị tông, người lái xe, người bị xe đâm và người ngoài cuộc chứng kiến ​​sự việc đều sẽ có góc nhìn khác nhau. Một cảnh sát tiếp nhận báo cáo từ mỗi người sẽ dễ hiểu là nhận được các sự kiện khác nhau từ mỗi người, ngay cả khi giả định rằng mỗi người nói sự thật theo cách họ biết. Trên thực tế, kinh nghiệm cho thấy điều này hầu như luôn xảy ra và những gì cảnh sát nộp dưới dạng báo cáo là mức trung gian giữa những gì mỗi người liên quan nêu, được bổ sung bằng kinh nghiệm cá nhân. Nói cách khác, báo cáo sẽ gần với sự thật, nhưng không đủ gần với sự thật đối với AI.

Khi xử lý góc nhìn, điều quan trọng là phải xem xét điểm thuận lợi. Người lái xe có thể nhìn thấy bảng điều khiển và biết tình trạng của xe tại thời điểm xảy ra tai nạn. Đây là thông tin mà hai bên còn lại không có. Tương tự như vậy, người bị xe đâm có điểm thuận lợi nhất để nhìn thấy biểu cảm khuôn mặt của người lái xe (ý định). Người ngoài cuộc có thể ở vị trí tốt nhất để xem liệu người lái xe có cố gắng dừng lại hay không và đánh giá các vấn đề như liệu người lái xe có cố gắng đánh lái hay không. Mỗi bên sẽ phải lập báo cáo dựa trên dữ liệu đã nhìn thấy mà không có lợi ích của dữ liệu ẩn.

\* Quan điểm có lẽ là sự dối trá nguy hiểm nhất

bởi vì bất kỳ ai cố gắng tìm ra sự thật trong tình huống này, tốt nhất là, sẽ kết thúc với một giá trị trung bình của nhiều câu chuyện khác nhau, mà sẽ không bao giờ hoàn toàn chính xác. Một con người xem thông tin có thể dựa vào trực giác và bản năng để có thể có được sự xấp xỉ tốt hơn về sự thật, nhưng AI sẽ luôn chỉ sử dụng giá trị trung bình, điều đó có nghĩa là AI luôn ở thế bất lợi đáng kể. Thật không may, việc tránh những quan điểm sai lệch là điều không thể bởi vì bất kể bạn có bao nhiêu nhân chứng cho sự kiện đó, thì điều tốt nhất bạn có thể hy vọng đạt được là sự xấp xỉ về sự thật, chứ không phải sự thật thực sự.

Cũng có một loại sai lầm khác cần xem xét, đó là sai lầm về góc nhìn. Hãy nghĩ về tình huống này: Bạn là một người khiếm thính vào năm 1927. Mỗi tuần, bạn đến rạp để xem một bộ phim câm, và trong một giờ hoặc hơn, bạn cảm thấy giống như mọi người khác. Bạn có thể trải nghiệm bộ phim theo cùng một cách mà mọi người khác đều trải nghiệm; không có sự khác biệt nào. Vào tháng 10 năm đó, bạn thấy một biển báo cho biết rạp đang nâng cấp để hỗ trợ hệ thống âm thanh để có thể chiếu *phim nói* — phim có nhạc nền. Biển báo cho biết đó là điều tuyệt vời nhất từ ​​trước đến nay, và hầu như mọi người đều đồng ý, ngoại trừ bạn, người khiếm thính, người giờ đây cảm thấy mình như một công dân hạng hai, khác biệt với mọi người khác và thậm chí gần như bị loại khỏi rạp. Trong mắt người khiếm thính, biển báo đó là sai lầm; thêm hệ thống âm thanh là điều tồi tệ nhất có thể, chứ không phải là điều tốt nhất có thể. Vấn đề là những gì có vẻ đúng nói chung thực ra không đúng với tất cả mọi người. Ý tưởng về một chân lý chung — một chân lý đúng với tất cả mọi người — là một huyền thoại. Nó không tồn tại.

***Sự thiên vị/thành kiến***

Sự sai lệch của thành kiến ​​xảy ra khi một người có thể nhìn thấy sự thật nhưng vì những mối quan tâm hoặc niềm tin cá nhân nên không thể thực sự nhìn thấy nó. Ví dụ, khi nghĩ về một vụ tai nạn, người lái xe có thể tập trung sự chú ý hoàn toàn vào giữa đường đến nỗi con nai ở lề đường trở nên vô hình. Do đó, người lái xe không có thời gian để phản ứng khi con nai đột nhiên quyết định lao ra giữa đường để cố gắng băng qua.

Một vấn đề với sự thiên vị là nó có thể cực kỳ khó để phân loại. Ví dụ, một người lái xe không nhìn thấy con nai có thể gặp tai nạn thực sự,

nghĩa là con nai bị che khuất khỏi tầm nhìn bởi bụi cây. Tuy nhiên, người lái xe cũng có thể phạm tội lái xe mất tập trung do không tập trung đúng cách. Người lái xe cũng có thể bị mất tập trung trong giây lát. Tóm lại, vấn đề không phải là người lái xe không nhìn thấy con nai; thay vào đó, vấn đề là tại sao người lái xe không nhìn thấy con nai. Trong nhiều trường hợp, việc xác nhận nguồn gốc của sự thiên vị trở nên quan trọng khi tạo ra một thuật toán được thiết kế để tránh nguồn thiên vị.

\* Về mặt lý thuyết, luôn có thể tránh được sự sai lệch của thiên kiến. Tuy nhiên, trên thực tế, tất cả con người đều có những thiên kiến ​​ở nhiều dạng khác nhau và những thiên kiến ​​đó sẽ luôn dẫn đến những sự sai lệch làm lệch lạc các tập dữ liệu. Chỉ cần khiến ai đó thực sự nhìn và sau đó thấy một thứ gì đó — để nó được ghi nhận trong não của người đó — là một nhiệm vụ khó khăn. Con người dựa vào các bộ lọc để tránh quá tải thông tin và các bộ lọc này cũng là nguồn gốc của sự sai lệch vì chúng ngăn cản mọi người thực sự nhìn thấy mọi thứ.

***Khung tham chiếu***

Trong năm sự nhầm lẫn, khung tham chiếu thực sự không cần phải là kết quả của bất kỳ loại lỗi nào, mà là kết quả của sự hiểu biết. Sự nhầm lẫn khung tham chiếu xảy ra khi một bên mô tả điều gì đó, chẳng hạn như một sự kiện như tai nạn, và vì bên thứ hai thiếu kinh nghiệm với sự kiện đó, nên các chi tiết trở nên lộn xộn hoặc hoàn toàn bị hiểu sai. Có rất nhiều tiết mục hài kịch dựa trên các lỗi khung tham chiếu. Một ví dụ nổi tiếng là từ Abbott và Costello, *Who's On First* , mà bạn có thể tìm thấy trên YouTube.com. Việc khiến một người hiểu được người thứ hai đang nói gì có thể là điều không thể khi người thứ nhất thiếu kiến ​​thức kinh nghiệm — khung tham chiếu.

Một ví dụ khác về sự không trung thực trong khuôn khổ tham chiếu xảy ra khi một bên không thể hiểu được bên kia. Ví dụ, một thủy thủ trải qua một cơn bão trên biển. Có lẽ đó là gió mùa, nhưng hãy giả sử trong một khoảnh khắc rằng cơn bão là đáng kể — có thể đe dọa đến tính mạng. Ngay cả khi sử dụng video, phỏng vấn và trình mô phỏng, trải nghiệm ở trên biển trong một cuộc sống

cơn bão đe dọa sẽ không thể truyền đạt cho một người chưa từng trải qua cơn bão như vậy; người đó không có đủ khả năng để hình dung.

\* Cách tốt nhất để tránh sự sai lệch về khung tham chiếu là đảm bảo rằng tất cả các bên liên quan có thể phát triển các khung tham chiếu tương tự. Để hoàn thành nhiệm vụ này, các bên khác nhau cần có kiến ​​thức kinh nghiệm tương tự để đảm bảo việc truyền dữ liệu chính xác từ người này sang người khác. Tuy nhiên, khi làm việc với một tập dữ liệu, vốn là dữ liệu tĩnh được ghi lại, các lỗi về khung tham chiếu vẫn sẽ xảy ra khi người xem tiềm năng thiếu kiến ​​thức kinh nghiệm cần thiết.

AI sẽ luôn gặp phải các vấn đề về khung tham chiếu vì AI không nhất thiết phải có khả năng tạo ra trải nghiệm. Một ngân hàng dữ liệu về kiến ​​thức thu được không hoàn toàn giống vậy. Ngân hàng dữ liệu sẽ chứa các sự kiện, nhưng trải nghiệm không chỉ dựa trên các sự kiện mà còn dựa trên các kết luận mà công nghệ hiện tại không thể sao chép.

## 5. Xác định giới hạn của việc thu thập dữ liệu

Nếu bạn có cảm giác rằng mọi người đang lấy dữ liệu của bạn mà không suy nghĩ hay lý do, thì bạn đúng rồi. Trên thực tế, các tổ chức thu thập, phân loại và lưu trữ dữ liệu của mọi người — dường như không có mục tiêu hay ý định. Theo “Data Never Sleeps” tại Domo.com, thế giới đang thu thập dữ liệu với tốc độ phi thường mỗi phút. Sau đây chỉ là một số ví dụ:

 Zoom có ​​208.333 người tham gia họp.

 Người dùng đăng 347.222 bài viết trên Instagram.

 Microsoft Teams kết nối 52.083 người dùng.

 Người dùng chia sẻ 41.666.667 tin nhắn trên WhatsApp.

 Mọi người thực hiện 1.388.889 cuộc gọi video/thoại.

Việc thu thập dữ liệu đã trở thành một chất gây nghiện đối với các tổ chức trên toàn thế giới và một số người nghĩ rằng tổ chức nào thu thập được nhiều nhất bằng cách nào đó sẽ giành được giải thưởng. Tuy nhiên, việc thu thập dữ liệu, tự nó, không mang lại kết quả gì. Cuốn sách *The Hitchhiker's Guide to the Galaxy* của Douglas Adams minh họa rõ ràng vấn đề này. Trong cuốn sách này, một chủng loài siêu sinh vật đã chế tạo một máy tính khổng lồ để tính toán ý nghĩa của "cuộc sống, vũ trụ và mọi thứ". Câu trả lời là 42 thực sự không giải quyết được bất cứ điều gì, vì vậy một số sinh vật phàn nàn rằng việc thu thập, phân loại và phân tích tất cả dữ liệu được sử dụng cho câu trả lời không tạo ra kết quả hữu ích. Máy tính — một máy tính có tri giác, không hơn không kém — nói với những người nhận được câu trả lời rằng câu trả lời thực sự đúng, nhưng họ cần biết câu hỏi để câu trả lời có ý nghĩa. Việc thu thập dữ liệu có thể diễn ra với số lượng không giới hạn, nhưng việc tìm ra câu hỏi phù hợp để hỏi có thể rất khó khăn, nếu không muốn nói là không thể.

\* Vấn đề chính mà bất kỳ tổ chức nào cũng cần giải quyết liên quan đến việc thu thập dữ liệu là những câu hỏi nào cần hỏi và tại sao những câu hỏi đó lại quan trọng. Việc điều chỉnh việc thu thập dữ liệu để trả lời những câu hỏi bạn cần trả lời là rất quan trọng. Ví dụ, nếu bạn đang điều hành một cửa hàng trong thị trấn, bạn có thể cần trả lời những câu hỏi như thế này:

 Mỗi ngày có bao nhiêu người đi bộ trước cửa hàng?

 Có bao nhiêu người dừng lại để nhìn vào cửa sổ?

 Họ nhìn trong bao lâu?

 Họ đang tìm kiếm vào thời điểm nào trong ngày?

 Có phải màn hình nào cũng mang lại kết quả tốt hơn không?

 Trong số những cách trưng bày này, cách nào thực sự thu hút mọi người đến cửa hàng và mua sắm?

Danh sách có thể dài thêm, nhưng ý tưởng là việc tạo ra một danh sách các câu hỏi giải quyết các nhu cầu kinh doanh cụ thể là điều cần thiết. Sau khi bạn tạo một danh sách, bạn phải xác minh rằng mỗi câu hỏi thực sự quan trọng — nghĩa là giải quyết một nhu cầu — và sau đó xác định loại thông tin nào bạn cần để trả lời câu hỏi.

\* Tất nhiên, việc cố gắng thu thập tất cả dữ liệu này bằng tay là điều không thể, đó là lúc tự động hóa phát huy tác dụng. Có vẻ như, tự động hóa sẽ tạo ra dữ liệu đầu vào đáng tin cậy, có thể lặp lại và nhất quán. Tuy nhiên, nhiều yếu tố trong quá trình tự động hóa thu thập dữ liệu có thể tạo ra dữ liệu không thực sự hữu ích. Ví dụ, hãy xem xét các vấn đề sau:

 Cảm biến chỉ có thể thu thập dữ liệu mà chúng được thiết kế để thu thập, do đó bạn có thể bỏ lỡ dữ liệu khi cảm biến được sử dụng không được thiết kế cho mục đích đó.

[Mọi người tạo ra dữ liệu sai lệch theo nhiều cách khác nhau (xem phần “Xem xét năm thông tin sai lệch trong dữ liệu” ở đầu chương này để biết chi tiết),](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-02-defining-the-role-of-data/#page66) điều đó có nghĩa là dữ liệu bạn nhận được có thể là sai.

 Dữ liệu có thể bị sai lệch khi các điều kiện thu thập dữ liệu không được xác định chính xác.

 Giải thích dữ liệu không chính xác có nghĩa là kết quả cũng sẽ không chính xác.

 Việc chuyển đổi một câu hỏi thực tế thành một thuật toán mà máy tính có thể hiểu được là một quá trình dễ xảy ra lỗi.

Nhiều vấn đề khác (đủ để viết thành một cuốn sách) cần được xem xét. Khi bạn kết hợp dữ liệu thu thập kém, không đúng định dạng với các thuật toán không thực sự trả lời được câu hỏi của bạn, bạn sẽ nhận được kết quả có thể thực sự dẫn doanh nghiệp của bạn đi sai hướng, đó là lý do tại sao AI thường bị đổ lỗi cho kết quả không nhất quán hoặc không đáng tin cậy. Đặt câu hỏi đúng, thu thập dữ liệu đúng, thực hiện xử lý đúng và sau đó phân tích dữ liệu chính xác đều là những yêu cầu cần thiết để biến việc thu thập dữ liệu thành loại công cụ mà bạn có thể tin cậy.

## 6. Vấn đề bảo mật dữ liệu

Phần này thảo luận về bảo mật dữ liệu theo quan điểm bảo vệ tính toàn vẹn của dữ liệu, thay vì ngăn chặn ai đó đánh cắp dữ liệu hoặc bảo vệ quyền riêng tư. Bảo mật dữ liệu không có nghĩa là đặt dữ liệu vào két an toàn — giả sử rằng việc này thậm chí có thể thực hiện được với dữ liệu ngày nay. Dữ liệu chỉ hữu ích khi có thể truy cập được. Tất nhiên, nhu cầu làm cho dữ liệu có thể truy cập được có nghĩa là chấp nhận rủi ro rằng ai đó sẽ làm điều gì đó mà bạn không muốn thực hiện với dữ liệu. Các phần sau đây thảo luận về một số vấn đề bảo mật dữ liệu mà bạn cần cân nhắc.

***Hiểu dữ liệu có chủ đích thiên vị***

Sự thiên vị xuất hiện trong hầu hết mọi tập dữ liệu hiện có ngày nay, ngay cả các tập dữ liệu được tạo tùy chỉnh. Tập dữ liệu thường bị thiên vị vì các phương pháp thu thập bị thiên vị, các phương pháp phân tích bị thiên vị và bản thân dữ liệu cũng bị thiên vị. Bạn thường thấy các bài viết trực tuyến có tiêu đề như "8 loại thiên vị trong phân tích dữ liệu và cách tránh chúng", điều này có nghĩa là mọi người nhận ra sự tồn tại của thiên vị và muốn giảm thiểu nó càng nhiều càng tốt. Tuy nhiên, đôi khi bạn thấy điều ngược lại mới đúng; những người sử dụng tập dữ liệu cố tình thiên vị theo một cách nào đó. Sau đây là một số lĩnh vực mà dữ liệu bị thiên vị một cách cố ý:

**Chính trị:** Các động thái chính trị có thể trở thành nguồn gốc của sự thiên vị dữ liệu. Hai nhóm có quan điểm đối lập sẽ sử dụng cùng một tập dữ liệu và thu được hai kết quả hoàn toàn khác nhau hỗ trợ cho quan điểm cụ thể của họ. Vấn đề là các bản ghi được chọn và các tính năng của tập dữ liệu được sử dụng để tạo ra kết quả. Trong các trường hợp khác, một nhóm sẽ dùng đến các kỹ thuật như sử dụng người trả lời giả mạo trong các cuộc thăm dò (xem “Đánh giá rủi ro đối với các cuộc thăm dò trực tuyến từ những người trả lời giả mạo” từ Pew Research.org để biết chi tiết).

**Y khoa:** Khi các nhóm y khoa quảng cáo cho mọi người tham gia thử nghiệm thuốc men, thủ thuật và các nhu cầu khác, nhóm mà họ tiếp cận thường không đại diện cho toàn bộ dân số, do đó dữ liệu bị thiên vị. Ví dụ, bài viết “Người lớn tuổi, nhóm thiểu số không được đại diện đầy đủ trong

“Thử nghiệm vắc-xin COVID-19” tại AARP.com chỉ ra rằng các thử nghiệm vắc-xin không có đủ nhóm thiểu số và người lớn tuổi, dẫn đến sai lệch dữ liệu.

**Pháp lý:** Việc sử dụng COMPAS để dự đoán khả năng tái phạm là một ví dụ khác về sự thiên vị của dữ liệu và thuật toán, như đã giải thích trong “Injustice Ex Machina: Predictive Algorithms in Criminal Sentencing,” tại UCLA Law Review.org. Bài viết chỉ ra rất nhiều lỗi với COMPAS đến mức ý tưởng tốt nhất có thể là bắt đầu lại từ đầu, vì phần mềm này đang hủy hoại cuộc sống của mọi người ở mức độ chưa từng có.

**Tuyển dụng:** Việc sử dụng các tập dữ liệu và thuật toán toàn diện được cho là làm giảm nguy cơ thiên vị trong việc tuyển dụng và thăng chức cho cá nhân trong một tổ chức. Theo “All the Ways Hiring Algorithms Can Introduce Bias” tại Harvard Business Review.org, điều ngược lại thường đúng. Các tập dữ liệu trở thành sự khuếch đại của các hoạt động tuyển dụng thiên vị trong một tổ chức hoặc trong toàn xã hội.

**Khác:** Bất cứ khi nào một tập dữ liệu và các thuật toán liên quan của nó bị ảnh hưởng bởi sự thiên vị, kết quả sẽ không lý tưởng. Thuật *ngữ công bằng trong học máy* thể hiện ý tưởng rằng kết quả của bất kỳ phân tích nào cũng phải thể hiện đúng các điều kiện thực tế trong xã hội (xem “Hướng dẫn về công bằng trong học máy” tại Towards Data Science.com để biết chi tiết). Nếu kết quả của một phân tích không khớp với kết quả nhận được sau đó, thì phân tích đó bị lỗi và dữ liệu thường phải chịu phần lớn trách nhiệm.

***Xử lý tình trạng hỏng hóc nguồn dữ liệu***

Ngay cả khi mọi người không chọn lọc dữ liệu hoặc sử dụng các nguồn dữ liệu không phản ánh được các điều kiện thực tế trên thế giới, như đã mô tả ở phần trước, thì các nguồn dữ liệu vẫn có thể bị hỏng. Ví dụ, khi xem các đánh giá sản phẩm trên một trang web, bạn không thể chắc chắn rằng

 Đánh giá được tạo ra bởi người thật.

 Một số người chưa bỏ phiếu nhiều hơn một lần.

 Người đó không chỉ đang có một ngày đặc biệt tồi tệ (hoặc ít có khả năng là tốt đẹp).

 Các đánh giá phản ánh một bộ phận không nhỏ của xã hội.

Trên thực tế, các đánh giá có thể thiên vị và sai lệch đến mức gần như không thể tin được. Thật không may, sự sai lệch nguồn dữ liệu đến từ nhiều nguồn khác:

 Cảm biến có thể bị hỏng, tạo ra kết quả không chính xác.

 Một cuộc tấn công của virus có thể gây ra lỗi dữ liệu.

 Cơ sở dữ liệu hoặc phần mềm khác có lỗi.

 Con người nhập dữ liệu không chính xác vào cơ sở dữ liệu.

 Các hiện tượng tự nhiên như sét đánh có thể gây ra sự cố tạm thời trong quá trình thu thập dữ liệu.

Bạn có thể dựa vào một số cách tiếp cận để giải quyết mọi loại hỏng dữ liệu. Lưu trữ dữ liệu trên đám mây có xu hướng làm giảm các vấn đề liên quan đến phần cứng, thời tiết hoặc các vấn đề khác gây mất dữ liệu. Đảm bảo rằng bạn có các quy trình và đào tạo tại chỗ, cùng với việc giám sát liên tục, có thể giúp giảm lỗi của con người. Sự tham gia tích cực của quản trị viên và sử dụng tường lửa có thể làm giảm các loại hỏng dữ liệu nguồn khác.

\* Tất cả các biện pháp này phản ánh những gì bạn có thể làm tại địa phương. Khi thực hiện quét màn hình và các kỹ thuật khác để lấy dữ liệu từ các nguồn trực tuyến, các nhà khoa học dữ liệu phải sử dụng các biện pháp khác để đảm bảo dữ liệu vẫn tinh khiết. Không thể xác thực cho một nguồn trực tuyến trừ khi nguồn đó được kiểm tra mỗi lần sử dụng.

***Hủy bỏ botnet bằng sinkholing***

*Botnet* là nhóm máy tính được phối hợp tập trung vào việc thực hiện các nhiệm vụ cụ thể, hầu hết là các nhiệm vụ bất chính. Phần ngắn này tập trung vào các botnet cung cấp dữ liệu sai lệch cho một tập dữ liệu hoặc chiếm quyền kiểm soát tài khoản để sửa đổi thông tin tài khoản theo những cách nhất định. Bất kể sử dụng phương tiện nào, bất kể mục đích là gì, botnet thường làm hỏng hoặc làm sai lệch dữ liệu theo những cách khiến bất kỳ loại phân tích nào cũng không thành công. Một trong những phương pháp tốt nhất để xử lý  
với các botnet này là làm chìm chúng — nghĩa là chuyển hướng chúng đến một vị trí mà chúng không thể gây hại. Bài viết “Hacker Lexicon: Sinkholing là gì?” *của Wired cung cấp các kỹ thuật để thực hiện nhiệm vụ này.*

# Chương 3: Xem xét việc sử dụng thuật toán

**NỘI DUNG CHƯƠNG NÀY**

* Khám phá vai trò của thuật toán trong AI.
* Chiến thắng trò chơi bằng tìm kiếm không gian trạng thái và thuật toán min-max.
* Phân tích cách hoạt động của hệ thống chuyên gia.
* Xem rằng máy học (machine learning) và học sâu (deep learning) là một phần của AI.

Dữ liệu là một yếu tố thay đổi cuộc chơi trong AI. Những tiến bộ trong AI cho thấy rằng đối với một số vấn đề, việc chọn đúng lượng dữ liệu quan trọng hơn thuật toán phù hợp. Ví dụ, vào năm 2001, hai nhà nghiên cứu từ Microsoft, Banko và Brill, trong bài báo đáng nhớ của họ, "Scaling to Very Very Large Corpora for Natural Language Disambiguation" (Mở rộng đến các tập hợp dữ liệu rất rất lớn để giải quyết sự mơ hồ của ngôn ngữ tự nhiên), đã chứng minh rằng nếu bạn muốn máy tính tạo ra mô hình ngôn ngữ, bạn không cần thuật toán thông minh nhất trong thị trấn. Sau khi đưa hơn một tỷ từ vào ngữ cảnh để giải quyết vấn đề, bất kỳ thuật toán nào cũng sẽ bắt đầu hoạt động cực kỳ tốt. Chương này giúp bạn hiểu mối quan hệ giữa các thuật toán và dữ liệu được sử dụng để khiến chúng thực hiện công việc hữu ích.

Tuy nhiên, bất kể bạn có bao nhiêu dữ liệu, bạn vẫn cần một thuật toán để làm cho chúng hữu ích. Ngoài ra, bạn phải thực hiện *phân tích dữ liệu* (một loạt các bước có thể xác định được) để làm cho dữ liệu hoạt động chính xác với các thuật toán đã chọn. Bạn không được phép sử dụng bất kỳ lối tắt nào. Mặc dù AI là tự động hóa thông minh, nhưng đôi khi tự động hóa phải lùi lại phía sau phân tích. Máy móc tự học là tương lai xa. Bạn sẽ không tìm thấy máy móc nào biết điều gì là phù hợp và có thể cắt giảm hoàn toàn mọi sự can thiệp của con người ngày nay. Nửa sau của chương này giúp bạn hiểu vai trò của các hệ thống chuyên gia, máy học, học sâu và các ứng dụng như AlphaGo trong việc đưa các khả năng trong tương lai đến gần hơn với thực tế.

## 1. Hiểu vai trò của thuật toán

Thuật *toán* là một quy trình bao gồm một chuỗi các phép toán. Thông thường, máy tính xử lý các phép toán này bằng cách tìm ra giải pháp đúng cho một bài toán trong thời gian hữu hạn hoặc cho bạn biết rằng không có giải pháp nào tồn tại. Mặc dù con người đã giải các thuật toán theo cách thủ công trong hàng nghìn năm, nhưng việc này có thể tiêu tốn rất nhiều thời gian và đòi hỏi nhiều phép tính số, tùy thuộc vào độ phức tạp của bài toán bạn muốn giải. Thuật toán là tất cả về việc tìm ra giải pháp và càng nhanh và càng dễ thì càng tốt. Thuật toán đã trở thành mã hóa cứng trong trí thông minh của con người đã nghĩ ra chúng và bất kỳ máy móc nào hoạt động trên thuật toán đều không thể không phản ánh trí thông minh được nhúng vào các quy trình thuật toán như vậy. AI cung cấp phương tiện để mô phỏng con người trong quá trình xử lý và giải các thuật toán hiện có, nhưng AI không thể thay thế con người hoặc bắt chước sự sáng tạo của con người trong việc nghĩ ra các thuật toán mới.

Mọi người có xu hướng nhận ra AI khi một công cụ đưa ra một cách tiếp cận mới lạ và tương tác với người dùng theo cách giống con người. Ví dụ bao gồm các trợ lý kỹ thuật số như Siri, Alexa, Cortana và Google Assistant. Tuy nhiên, một số công cụ phổ biến khác, chẳng hạn như bộ định tuyến GPS và các công cụ lập kế hoạch chuyên dụng (như những công cụ được sử dụng để tránh va chạm ô tô, máy bay tự lái và sắp xếp kế hoạch sản xuất) thậm chí không giống như AI vì chúng quá phổ biến và bị coi là điều hiển nhiên khi chúng hoạt động ở chế độ ẩn. Ngoài ra, điều quan trọng là phải xem xét các hình thức AI thay thế, chẳng hạn như bộ điều nhiệt thông minh kiểm soát môi trường dựa trên mức sử dụng trước đây và dữ liệu môi trường hiện tại, và bộ mở cửa nhà để xe thông minh tự động phát hiện khi bạn vô tình để cửa mở sau khi đi làm.

Đây rõ ràng là *hiệu ứng AI,* như được Pamela McCorduck, một tác giả người Mỹ đã viết một lịch sử đáng chú ý về AI, *Machines Who Think,* đặt tên và mô tả vào năm 1979. (Phiên bản trên Amazon.com là phiên bản cập nhật.) Hiệu ứng AI nêu rằng mọi người sớm quên đi các chương trình máy tính thông minh, thành công, trở thành những diễn viên thầm lặng trong khi sự chú ý chuyển sang các vấn đề AI vẫn cần giải quyết. Tầm quan trọng của các thuật toán cổ điển đối với AI bị bỏ qua và mọi người bắt đầu mơ tưởng về AI được tạo ra từ công nghệ bí truyền hoặc họ đánh đồng nó với những tiến bộ gần đây, chẳng hạn như học máy và học sâu.

***Hiểu thuật toán có nghĩa là gì***

Một thuật toán luôn trình bày một loạt các bước, nhưng nó không nhất thiết phải thực hiện tất cả các bước này để giải quyết một vấn đề (một số bước là tùy chọn hoặc chỉ được thực hiện trong các điều kiện cụ thể). Một nhóm các bước liên quan là một *hoạt động* , chẳng hạn như hoạt động pha trà bao gồm các bước sau:

* Đổ nước vào ấm trà.
* Bật lửa để đun nóng nước trong ấm trà.
* Khi nước nóng, đổ nước vào cốc.
* Cho một túi trà vào cốc và ngâm trà trong thời gian khuyến nghị.
* Bỏ túi trà ra.
* (Tùy chọn) Thêm đường vào trà.
* (Tùy chọn) Thêm sữa vào trà.
* Uống trà.
* (Tùy chọn) Đổ trà vào bồn rửa khi không uống được. Xem bồn rửa tan chảy.

Phạm vi của thuật toán cực kỳ rộng lớn. Các hoạt động có thể liên quan đến việc lưu trữ dữ liệu, khám phá dữ liệu và sắp xếp hoặc sắp xếp dữ liệu thành các cấu trúc dữ liệu. Bạn có thể tìm thấy các thuật toán giải quyết các vấn đề trong khoa học, y học, tài chính, sản xuất và cung ứng công nghiệp và truyền thông.

Tất cả các thuật toán đều chứa các chuỗi hoạt động để tìm ra giải pháp đúng cho một vấn đề trong thời gian hợp lý (hoặc báo cáo lại nếu không tìm thấy giải pháp nào). Một phân lớp thuật toán, *phương pháp tìm kiếm,* tạo ra các giải pháp tốt nhưng không nhất thiết phải hoàn hảo khi thời gian quan trọng hơn việc tìm ra giải pháp hoàn hảo. Các thuật toán AI tự phân biệt mình với các thuật toán chung bằng cách giải quyết các vấn đề mà độ phân giải thường được coi là (hoặc thậm chí là độc quyền) sản phẩm của hành vi thông minh của con người. Các thuật toán AI có xu hướng giải quyết các vấn đề phức tạp, thường là một phần của lớp vấn đề *NP-hoàn chỉnh* (trong đó NP là thời gian đa thức không xác định) mà con người thường xuyên giải quyết bằng cách kết hợp phương pháp tiếp cận hợp lý và trực giác. Sau đây chỉ là một vài ví dụ:

 Lên lịch các vấn đề và phân bổ các nguồn lực khan hiếm  Tìm kiếm các tuyến đường trong không gian vật lý hoặc tượng hình phức tạp

 Nhận ra các mô hình trong tầm nhìn hình ảnh (so với thứ gì đó như phục hồi hình ảnh hoặc xử lý hình ảnh) hoặc nhận thức âm thanh

 Xử lý ngôn ngữ (cả hiểu văn bản và dịch ngôn ngữ)

 Chơi (và chiến thắng) các trò chơi cạnh tranh

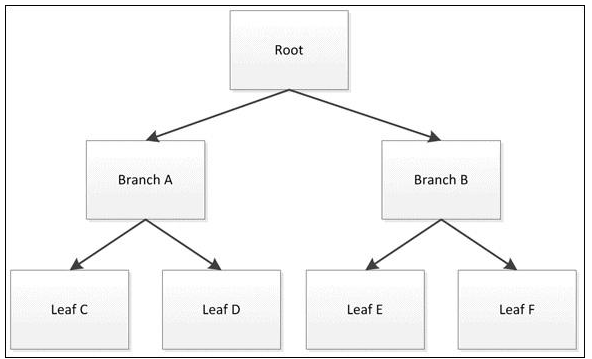
\*Các bài toán NP-complete phân biệt chúng với các bài toán thuật toán khác vì việc tìm ra lời giải cho chúng trong một khung thời gian hợp lý vẫn chưa khả thi. NP-complete không phải là loại bài toán mà bạn giải bằng cách thử tất cả các kết hợp hoặc khả năng có thể. Ngay cả khi bạn có máy tính mạnh hơn những máy tính hiện nay, việc tìm kiếm lời giải sẽ kéo dài gần như mãi mãi. Tương tự như vậy, trong AI, loại bài toán này được gọi là *AI-complete* .

***Lập kế hoạch và phân nhánh: Cây và nút***

Lập kế hoạch giúp bạn xác định trình tự các hành động cần thực hiện để đạt được một mục tiêu nhất định. Quyết định về kế hoạch là một vấn đề AI cổ điển và bạn tìm thấy các ví dụ về lập kế hoạch trong sản xuất công nghiệp, phân bổ tài nguyên và di chuyển rô-bốt bên trong phòng. Bắt đầu từ trạng thái hiện tại, trước tiên AI xác định tất cả các hành động có thể từ trạng thái đó. Về mặt kỹ thuật, nó *mở rộng* trạng thái hiện tại thành một số trạng thái tương lai. Sau đó, nó mở rộng tất cả các trạng thái tương lai thành trạng thái tương lai của riêng chúng, v.v. Khi bạn không thể mở rộng các trạng thái nữa và AI dừng việc mở rộng, AI đã tạo ra một *không gian trạng thái,* bao gồm bất kỳ điều gì có thể xảy ra trong tương lai. AI có thể tận dụng không gian trạng thái không chỉ như một dự đoán khả thi (thực tế, nó dự đoán mọi thứ, mặc dù một số trạng thái tương lai có nhiều khả năng xảy ra hơn những trạng thái khác) mà còn vì AI có thể sử dụng không gian trạng thái đó để khám phá các quyết định mà nó có thể đưa ra để đạt được mục tiêu theo cách tốt nhất. Quá trình này được gọi là *tìm kiếm không gian trạng thái.*

Làm việc với không gian trạng thái đòi hỏi phải sử dụng cả cấu trúc dữ liệu và thuật toán cụ thể. Các cấu trúc dữ liệu cốt lõi thường được sử dụng là cây và đồ thị. Các thuật toán được ưa chuộng để khám phá đồ thị hiệu quả bao gồm tìm kiếm theo chiều rộng hoặc tìm kiếm theo chiều sâu.

Xây dựng một cây hoạt động theo cách tương tự như cách một cây phát triển trong thế giới vật lý. Mỗi mục bạn thêm vào cây là một *nút.* Các nút kết nối với nhau bằng các liên kết. Sự kết hợp của các nút và liên kết tạo thành một cấu trúc trông giống như một cây, như thể hiện trong [Hình 3-1](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-03-considering-the-use-of-algorithms/#page85) .



[**HÌNH 3-1:**](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-03-considering-the-use-of-algorithms/#page85)Một cái cây có thể trông giống như cây thật hoặc có rễ hướng lên trên.

\*Cây có một nút gốc, giống như cây vật lý. *Nút gốc* là điểm bắt đầu cho quá trình xử lý mà bạn thực hiện. Kết nối với gốc là các nhánh hoặc lá. *Nút lá* là điểm kết thúc của cây. *Các nút nhánh* hỗ trợ các nhánh hoặc lá khác. Kiểu cây được hiển thị trong [Hình 3-1](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-03-considering-the-use-of-algorithms/#page85) là cây nhị phân vì mỗi nút có tối đa hai kết nối (nhưng cây biểu diễn không gian trạng thái có thể có nhiều nhánh).

Khi xem xét cây, Nhánh B là *con* của nút Gốc. Đó là vì nút Gốc xuất hiện đầu tiên trong cây. Lá E và Lá F đều là con của Nhánh B, khiến Nhánh B trở thành *cha* của Lá E và Lá F. Mối quan hệ giữa các nút rất quan trọng vì các cuộc thảo luận về cây thường xem xét mối quan hệ con/cha giữa các nút. Nếu không có các thuật ngữ này, các cuộc thảo luận về cây có thể trở nên khá khó hiểu.

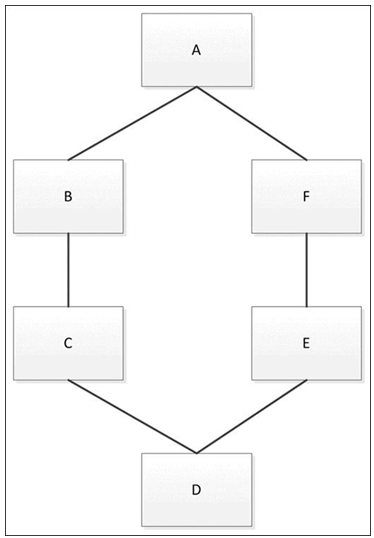
***Mở rộng cây bằng cách sử dụng các nút đồ thị***

Biểu *đồ* là một dạng mở rộng của cây. Giống như cây, bạn có các nút kết nối với nhau để tạo mối quan hệ. Tuy nhiên, không giống như cây nhị phân, một nút biểu đồ có thể có nhiều hơn một hoặc hai kết nối. Trên thực tế, các nút biểu đồ thường có vô số kết nối và quan trọng nhất là các nút có thể kết nối theo bất kỳ hướng nào, không chỉ từ nút cha đến nút con. Tuy nhiên, để đơn giản, hãy xem xét biểu đồ được hiển thị trong [Hình 3-2](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-03-considering-the-use-of-algorithms/#page87) .

Đồ thị là những cấu trúc thể hiện một số nút (hoặc đỉnh) được kết nối bởi một số cạnh hoặc cung (tùy thuộc vào cách biểu diễn). Khi bạn nghĩ về đồ thị, hãy nghĩ về một cấu trúc như bản đồ, trong đó mỗi vị trí trên bản đồ là một nút và các đường phố là các cạnh. Cách biểu diễn này khác với cây, trong đó mỗi đường dẫn kết thúc ở một nút lá. Tham khảo [Hình 3-2](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-03-considering-the-use-of-algorithms/#page87) để xem biểu đồ được biểu diễn. Đồ thị đặc biệt hữu ích khi tìm ra các trạng thái biểu diễn một loại không gian vật lý. Ví dụ, GPS sử dụng đồ thị để biểu diễn các địa điểm và đường phố.

Biểu đồ cũng thêm một vài điểm mới mà bạn có thể chưa từng nghĩ đến. Ví dụ, biểu đồ có thể bao gồm khái niệm về hướng. Không giống như cây, có mối quan hệ cha/con, một nút biểu đồ có thể kết nối với bất kỳ nút nào khác có hướng cụ thể trong đầu. Hãy nghĩ về những con phố trong thành phố. Hầu hết các con phố đều là hai chiều, nhưng một số là đường một chiều chỉ cho phép di chuyển theo một hướng.

Việc trình bày kết nối đồ thị có thể không thực sự phản ánh thực tế của đồ thị. Một đồ thị có thể chỉ định *trọng số* cho một kết nối cụ thể. Trọng số có thể xác định khoảng cách giữa hai điểm, xác định thời gian cần thiết để đi qua tuyến đường, chỉ định lượng nhiên liệu sử dụng để đi hết tuyến đường hoặc cung cấp các loại thông tin khác.



[**HÌNH 3-2:**](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-03-considering-the-use-of-algorithms/#page86)Các nút đồ thị có thể kết nối với nhau theo vô số cách.

\* Một cây không gì khác hơn là một đồ thị trong đó bất kỳ hai đỉnh nào được kết nối bằng đúng một đường đi và cây không cho phép chu trình (để có thể quay lại đỉnh cha từ bất kỳ đỉnh con nào). Nhiều thuật toán đồ thị chỉ áp dụng cho cây.

***Đi qua đồ thị***

*Duyệt đồ thị* có nghĩa là tìm kiếm (ghé thăm) từng đỉnh (nút) theo một thứ tự cụ thể. Quá trình ghé thăm một đỉnh có thể bao gồm cả việc đọc và cập nhật nó. Bạn khám phá ra các đỉnh chưa được ghé thăm khi bạn duyệt đồ thị. Đỉnh được khám phá (vì bạn vừa ghé thăm nó) hoặc được xử lý (vì thuật toán đã thử tất cả các cạnh rời khỏi nó) sau khi tìm kiếm. Thứ tự tìm kiếm xác định loại tìm kiếm được thực hiện:

**Không có thông tin (tìm kiếm mù):** AI khám phá không gian trạng thái mà không có thông tin bổ sung ngoại trừ cấu trúc đồ thị mà nó phát hiện khi duyệt qua. Sau đây là hai thuật toán tìm kiếm mù phổ biến, được thảo luận trong các phần sau:

**Tìm kiếm theo chiều rộng (BFS):** Bắt đầu từ gốc đồ thị và khám phá mọi nút gắn với gốc. Sau đó, nó tìm kiếm ở cấp độ tiếp theo, khám phá từng cấp độ theo lượt cho đến khi đến cuối. Do đó, trong đồ thị ví dụ, tìm kiếm khám phá từ A đến B và C trước khi chuyển sang khám phá D. BFS khám phá đồ thị theo cách có hệ thống, khám phá các đỉnh xung quanh đỉnh bắt đầu theo cách vòng tròn. Nó bắt đầu bằng cách truy cập tất cả các đỉnh cách đỉnh bắt đầu một bước; sau đó nó di chuyển ra hai bước, rồi ba bước, v.v.

**Tìm kiếm theo chiều sâu (DFS):** Bắt đầu từ gốc đồ thị và sau đó khám phá mọi nút từ gốc đó xuống một đường dẫn duy nhất đến cuối. Sau đó, nó quay lại và bắt đầu khám phá các đường dẫn không được thực hiện trong đường dẫn tìm kiếm hiện tại cho đến khi nó đạt đến gốc một lần nữa. Tại thời điểm đó, nếu có các đường dẫn khác để thực hiện từ gốc,

Thuật toán chọn một và bắt đầu tìm kiếm tương tự một lần nữa. Ý tưởng là khám phá từng đường dẫn hoàn toàn trước khi khám phá bất kỳ đường dẫn nào khác.

**Có thông tin (phương pháp tìm kiếm):** Phương pháp tìm kiếm tìm hoặc khám phá ra phương pháp tốt nhất để duyệt đồ thị dựa trên các quy tắc chung (chẳng hạn như hệ thống chuyên gia) hoặc các thuật toán sử dụng thời gian đa thức bậc thấp. Đó là một phỏng đoán có căn cứ về một giải pháp chỉ ra hướng của kết quả mong muốn nhưng không thể cho biết chính xác cách để đạt được kết quả đó. Giống như bị lạc trong một thành phố xa lạ và có người chỉ cho bạn một cách nhất định để đến khách sạn của bạn (nhưng không có hướng dẫn chính xác). Vì tìm kiếm này có thông tin (mặc dù không chính xác), nên nó cũng có thể ước tính chi *phí* còn lại (thời gian, tài nguyên hoặc giá trị khác xác định tuyến đường nào tốt hơn trong một trường hợp cụ thể) để đi từ một trạng thái cụ thể đến một giải pháp. Sau đây là ba thuật toán tìm kiếm theo phương pháp tìm kiếm thông thường (xem [phần "Sử dụng tìm kiếm cục bộ và phương pháp tìm kiếm" của chương này và Chương](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-03-considering-the-use-of-algorithms/#page235)

[10 để biết thêm chi tiết):](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-03-considering-the-use-of-algorithms/#page235)

**Tìm kiếm theo phương pháp tốt nhất trước:** Một hàm đánh giá hỗ trợ tìm kiếm bằng cách xác định tính mong muốn của việc mở rộng một nút cụ thể dựa trên chi phí của các nút tiếp theo. Chi phí của mỗi nút được lưu trữ trong hàng đợi hoặc cấu trúc bộ nhớ khác. Ngoại trừ việc biết trước chi phí nút, giải pháp này hoạt động rất giống với BFS hoặc DFS.

**Tìm kiếm tham lam:** Giống như tìm kiếm tốt nhất trước, đường dẫn cần theo dõi được thông báo bởi chi phí nút. Tuy nhiên, tìm kiếm tham lam chỉ tìm kiếm một nút trước, giúp tiết kiệm thời gian xử lý trong hàm đánh giá, nhưng không phải lúc nào cũng đảm bảo giải pháp tối ưu.

**Tìm kiếm A\*:** Một dạng mở rộng của tìm kiếm tốt nhất trước tiên, thực tế sử dụng hai chi phí: chi phí di chuyển từ điểm bắt đầu đến một vị trí nhất định khác trên đồ thị và chi phí di chuyển từ nút nhất định đó trên đồ thị đến đích cuối cùng.

## 2. Chơi trò chơi đối kháng

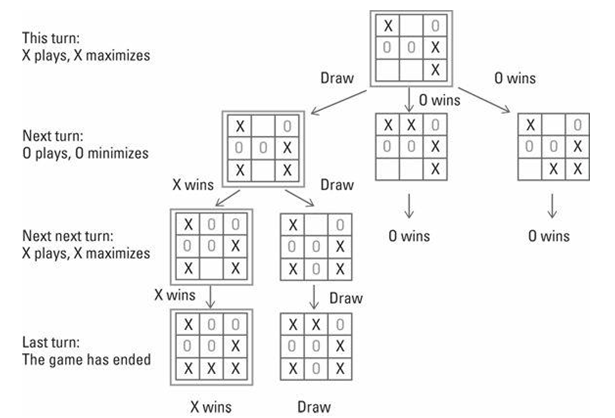
Điều thú vị về tìm kiếm không gian trạng thái là nó đại diện cho cả chức năng hiện tại và các cơ hội trong tương lai của AI. Đây là trường hợp của *các trò chơi đối kháng* (trò chơi mà một người thắng và những người khác thua) hoặc với bất kỳ tình huống tương tự nào trong đó người chơi theo đuổi một mục tiêu xung đột với mục tiêu của những người khác. Một trò chơi đơn giản như ô ăn quan là một ví dụ hoàn hảo về trò chơi tìm kiếm không gian mà bạn có thể đã thấy AI chơi. Trong bộ phim *WarGames năm 1983,* siêu máy tính WOPR (War Operation Plan Response) chơi với chính nó với tốc độ cực nhanh, nhưng nó không thể thắng vì trò chơi thực sự đơn giản và nếu bạn sử dụng tìm kiếm không gian trạng thái, bạn sẽ không bao giờ thua.

Bạn có chín ô để điền X và O cho mỗi người chơi. Người đầu tiên đặt ba dấu liên tiếp (ngang, dọc hoặc chéo) sẽ thắng. Khi xây dựng cây không gian trạng thái cho trò chơi, mỗi cấp độ của cây đại diện cho một lượt chơi. Các nút kết thúc đại diện cho trạng thái bảng cuối cùng và xác định chiến thắng, hòa hoặc thất bại cho AI. Mỗi nút đầu cuối có điểm cao hơn khi thắng, thấp hơn khi hòa và thậm chí thấp hơn hoặc âm khi thua. AI truyền điểm đến các nút và nhánh trên cùng bằng cách sử dụng phép cộng cho đến khi đạt đến nút bắt đầu. Nút bắt đầu đại diện cho tình huống thực tế. Sử dụng một chiến lược đơn giản cho phép bạn duyệt cây: Khi đến lượt AI và bạn phải truyền giá trị của nhiều nút, bạn cộng giá trị tối đa (có lẽ vì AI phải có kết quả tối đa từ trò chơi); khi đến lượt đối thủ, bạn cộng giá trị tối thiểu thay thế. Cuối cùng, bạn sẽ nhận được một cây có các nhánh được xác định theo điểm. Khi đến lượt AI, nó sẽ chọn nước đi dựa trên nhánh có giá trị cao nhất vì nó ngụ ý mở rộng các nút có khả năng chiến thắng cao nhất. [Hình 3-3](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-03-considering-the-use-of-algorithms/#page91) cho thấy một ví dụ trực quan về chiến lược này.

Cách tiếp cận này được gọi là xấp xỉ min-max. Ronald Rivest, từ phòng thí nghiệm khoa học máy tính tại MIT, đã giới thiệu nó vào năm 1987 (bạn có thể đọc bài báo của ông tại

<https://people.csail.mit.edu/rivest/pubs/Riv87c.pdf> ). Kể từ đó, thuật toán này và các biến thể của nó đã hỗ trợ nhiều trò chơi cạnh tranh, cùng với những tiến bộ gần đây trong trò chơi, chẳng hạn như AlphaGo của GoogleDeepMind, sử dụng phương pháp tiếp cận mô phỏng phép tính xấp xỉ min-max (cũng xuất hiện trong bộ phim *WarGames* năm 1983).

\* Đôi khi bạn nghe nói về việc cắt tỉa alpha-beta liên quan đến phép xấp xỉ min-max. *Cắt tỉa alpha-beta* là một cách thông minh để truyền bá các giá trị lên hệ thống phân cấp cây trong các không gian trạng thái phức tạp hạn chế các phép tính. Không phải tất cả các trò chơi đều có cây không gian trạng thái nhỏ gọn; khi các nhánh của bạn có số lượng hàng triệu, bạn cần cắt tỉa chúng và rút ngắn các phép tính của mình.



[**HÌNH 3-3:**](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-03-considering-the-use-of-algorithms/#page90)Nhìn thoáng qua phép tính gần đúng min-max trong trò chơi ô ăn quan.

***Sử dụng tìm kiếm cục bộ và phương pháp tìm kiếm***

Có rất nhiều thứ diễn ra đằng sau phương pháp tìm kiếm không gian trạng thái. Cuối cùng, không có máy móc nào, dù mạnh đến đâu, có thể liệt kê tất cả các khả năng phát sinh từ một tình huống phức tạp. Phần này tiếp tục với các trò chơi vì chúng có thể dự đoán được và có các quy tắc cố định, trong khi nhiều tình huống trong thế giới thực không thể đoán trước và thiếu các quy tắc rõ ràng, khiến trò chơi trở thành một bối cảnh lạc quan và thuận lợi.

Cờ đam, một trò chơi tương đối đơn giản so với cờ vua hoặc cờ vây, có 500 tỷ tỷ (500.000.000.000.000.000.000) vị trí có thể có trên bàn cờ, một con số mà theo tính toán của các nhà toán học tại Đại học Hawaii, tương đương với tất cả các hạt cát trên Trái đất. Đúng là càng về sau, ván cờ đam càng có ít nước đi hơn. Tuy nhiên, con số có thể đánh giá ở mỗi nước đi lại quá cao. Phải mất 18 năm sử dụng máy tính mạnh (xem “Checkers Solved” tại ScienceNetLinks.com) để tính toán tất cả 500 tỷ tỷ nước đi có thể có. Hãy tưởng tượng xem máy tính của người tiêu dùng có thể mất bao lâu để tính toán được ngay cả một tập hợp con nhỏ hơn các nước đi. Để có thể quản lý được, nó phải là một tập hợp con rất nhỏ trong số tất cả các nước đi tiềm năng.

Tối ưu hóa bằng cách sử dụng tìm kiếm cục bộ và phương pháp tìm kiếm giúp ích bằng cách sử dụng các ràng buộc để giới hạn số lượng đánh giá có thể có ban đầu (như trong cắt tỉa alpha, trong đó một số phép tính bị bỏ qua vì chúng không thêm bất cứ thứ gì vào thành công tìm kiếm). *Tìm kiếm cục bộ* là một phương pháp giải quyết vấn đề chung bao gồm nhiều thuật toán giúp bạn thoát khỏi sự phức tạp theo cấp số nhân của nhiều bài toán NP. Tìm kiếm cục bộ bắt đầu từ tình huống hiện tại của bạn hoặc một giải pháp bài toán không hoàn hảo và di chuyển ra khỏi nó, từng bước một. Tìm kiếm cục bộ xác định tính khả thi của các giải pháp gần đó, có khả năng dẫn đến một giải pháp hoàn hảo, dựa trên sự lựa chọn ngẫu nhiên hoặc phương pháp tìm kiếm thông minh (có nghĩa là không có phương pháp chính xác nào được sử dụng).

Thuật toán tìm kiếm cục bộ cải thiện theo từng bước từ trạng thái bắt đầu, di chuyển từng bước một qua các giải pháp lân cận trong không gian trạng thái cho đến khi chúng không thể cải thiện giải pháp thêm nữa. Vì thuật toán tìm kiếm cục bộ rất đơn giản và trực quan, nên việc thiết kế phương pháp tìm kiếm cục bộ cho một bài toán thuật toán không khó; việc làm cho nó hiệu quả thường khó hơn. Chìa khóa là xác định quy trình chính xác:

1. Bắt đầu với một tình huống hiện có (có thể là tình huống hiện tại hoặc một giải pháp ngẫu nhiên hoặc đã biết).

* Tìm kiếm một tập hợp các giải pháp mới có thể có trong vùng lân cận của giải pháp hiện tại, tạo nên danh sách ứng viên.
* Xác định giải pháp nào sẽ được sử dụng thay cho giải pháp hiện tại dựa trên kết quả của thuật toán tìm kiếm chấp nhận danh sách ứng viên làm đầu vào.
* Tiếp tục thực hiện Bước 2 và Bước 3 cho đến khi bạn không thấy giải pháp nào cải thiện nữa, điều đó có nghĩa là bạn đã có giải pháp tốt nhất.

Mặc dù dễ thiết kế, các giải pháp tìm kiếm cục bộ có thể không tìm ra giải pháp trong thời gian hợp lý (bạn có thể dừng quy trình và sử dụng giải pháp hiện tại) hoặc tạo ra giải pháp có chất lượng tối thiểu. Bạn không đảm bảo rằng tìm kiếm cục bộ sẽ đưa ra giải pháp có vấn đề, nhưng cơ hội của bạn sẽ cải thiện từ điểm bắt đầu khi bạn cung cấp đủ thời gian để tìm kiếm chạy các phép tính của nó. Nó chỉ dừng lại sau khi không thể tìm ra bất kỳ cách nào khác để cải thiện giải pháp. Bí quyết là xác định đúng vùng lân cận để khám phá. Nếu bạn khám phá mọi thứ, bạn sẽ quay lại tìm kiếm toàn diện, ngụ ý một vụ nổ các khả năng để khám phá và thử nghiệm.

Dựa vào giới hạn của phép thử nghiệm khi bạn nhìn dựa trên quy tắc chung. Đôi khi phép thử nghiệm là tính ngẫu nhiên, và một giải pháp như vậy, mặc dù là một cách tiếp cận không thông minh, có thể hoạt động tốt. Ví dụ, ít người biết rằng Roomba, máy hút bụi rô-bốt tự động do ba sinh viên tốt nghiệp MIT tạo ra (xem ví dụ, iRobot Roomba, tại Amazon.com), ban đầu không lập kế hoạch cho đường đi vệ sinh của nó mà chỉ đơn giản là đi lang thang ngẫu nhiên. Tuy nhiên, nó được chủ sở hữu coi là một thiết bị thông minh và đã làm rất tốt công việc vệ sinh. (Thực ra, trí thông minh nằm ở ý tưởng sử dụng tính ngẫu nhiên để giải quyết một vấn đề vốn quá phức tạp.)

Lựa chọn ngẫu nhiên không phải là phương pháp tìm kiếm duy nhất khả dụng. Tìm kiếm cục bộ có thể dựa vào các giải pháp khám phá hợp lý hơn bằng cách sử dụng phương pháp tìm kiếm được thiết kế tốt để có chỉ dẫn:

**Leo đồi:** Dựa trên quan sát rằng khi một quả bóng lăn xuống thung lũng, nó sẽ đi xuống dốc nhất. Khi leo đồi, một quả bóng có xu hướng đi theo hướng thẳng nhất lên trên để lên đến đỉnh, đó là hướng có độ nghiêng lớn nhất. Do đó, vấn đề AI được coi là

một sự xuống dốc đến một thung lũng hoặc một sự lên dốc đến đỉnh núi, và phương pháp tìm kiếm là bất kỳ quy tắc nào gợi ý về cách tiếp cận xuống dốc hoặc lên dốc tốt nhất trong số các trạng thái có thể có của không gian trạng thái. Đây là một thuật toán hiệu quả, mặc dù đôi khi nó bị kẹt trong các tình huống được gọi là cao nguyên (thung lũng trung gian) và đỉnh (điểm cực đại cục bộ).

**Twiddle (thuật toán hạ tọa độ):** Tương tự như thuật toán leo đồi, nhưng nó khám phá mọi hướng có thể. Nó tập trung tìm kiếm theo hướng lân cận hoạt động tốt nhất. Khi làm như vậy, nó hiệu chỉnh bước của mình, chậm lại khi thấy việc khám phá ra các giải pháp tốt hơn là khó khăn, cho đến khi dừng lại.

**Ủ mô phỏng:** Lấy tên từ một kỹ thuật luyện kim nung nóng kim loại rồi làm nguội từ từ để làm mềm kim loại để gia công nguội và loại bỏ các khuyết tật về tinh thể (xem “Xử lý nhiệt thép và kim loại” tại Bright Hub Engineering.com để biết chi tiết). Tìm kiếm cục bộ sao chép kỹ thuật này bằng cách xem tìm kiếm giải pháp như một cấu trúc nguyên tử thay đổi để cải thiện khả năng gia công của nó. Nhiệt độ là yếu tố thay đổi cuộc chơi trong quá trình tối ưu hóa. Cũng giống như nhiệt độ cao làm cho cấu trúc của vật liệu giãn ra (chất rắn tan chảy và chất lỏng bay hơi ở nhiệt độ cao), nhiệt độ cao trong thuật toán tìm kiếm cục bộ gây ra sự giãn ra của hàm mục tiêu, cho phép nó ưu tiên các giải pháp tệ hơn so với các giải pháp tốt hơn. Ủ mô phỏng sửa đổi quy trình leo đồi, giữ nguyên hàm mục tiêu để đánh giá giải pháp lân cận, nhưng cho phép nó xác định lựa chọn giải pháp tìm kiếm theo một cách khác.

**Tìm kiếm cấm kỵ:** Sử dụng trí nhớ để nhớ những phần nào của khu phố cần khám phá. Khi có vẻ như đã tìm ra giải pháp, nó có xu hướng cố gắng quay lại những con đường khả thi khác mà nó chưa thử để xác định giải pháp tốt nhất.

Sử dụng các biện pháp về hướng (lên, xuống) hoặc nhiệt độ (tính ngẫu nhiên được kiểm soát), hoặc chỉ cần hạn chế hoặc quay lại một phần tìm kiếm đều là những cách để tránh thử mọi thứ và tập trung vào một giải pháp tốt. Ví dụ, hãy xem xét một con robot đang đi bộ. Việc dẫn dắt một con robot trong một môi trường không xác định có nghĩa là tránh chướng ngại vật để đến được một mục tiêu cụ thể. Đây vừa là một nhiệm vụ cơ bản vừa đầy thách thức trong trí tuệ nhân tạo. Robot có thể dựa vào máy đo khoảng cách laser (lidar) hoặc sonar (bao gồm các thiết bị sử dụng âm thanh để nhìn thấy môi trường của chúng) để điều hướng xung quanh. Tuy nhiên, bất kể mức độ tinh vi của phần cứng, robot vẫn cần các thuật toán phù hợp để

 Tìm đường đi ngắn nhất đến đích (hoặc ít nhất là đường đi tương đối ngắn)

 Tránh chướng ngại vật trên đường đi

 Thực hiện các hành vi tùy chỉnh như giảm thiểu việc rẽ hoặc phanh

Thuật toán tìm đường giúp robot bắt đầu ở một vị trí và đến đích bằng cách sử dụng đường đi ngắn nhất giữa hai vị trí, dự đoán và tránh chướng ngại vật trên đường đi. (Phản ứng sau khi đâm vào tường là không đủ.) Tìm đường cũng hữu ích khi di chuyển bất kỳ thiết bị nào khác đến mục tiêu trong không gian, ngay cả mục tiêu ảo, chẳng hạn như trong trò chơi điện tử hoặc trang web. Khi sử dụng tìm đường với robot, robot nhận thức chuyển động như một luồng không gian trạng thái đến ranh giới của các cảm biến của nó. Nếu mục tiêu không nằm trong phạm vi, robot sẽ không biết phải đi đâu. Các phương pháp tìm kiếm có thể chỉ cho robot đúng hướng (ví dụ, robot có thể biết rằng mục tiêu ở hướng bắc) và giúp robot tránh chướng ngại vật kịp thời mà không cần phải xác định mọi cách có thể để thực hiện việc đó.

## 3. Khám phá máy học

Tất cả các ví dụ về thuật toán cho đến nay đều liên quan đến AI vì chúng là các giải pháp thông minh giải quyết các vấn đề lặp đi lặp lại và được phân định rõ ràng nhưng phức tạp, đòi hỏi trí thông minh. Chúng đòi hỏi một kiến ​​trúc sư nghiên cứu vấn đề và chọn đúng thuật toán để giải quyết. Các thay đổi, đột biến hoặc hiển thị đặc điểm bất thường của vấn đề có thể trở thành vấn đề thực sự đối với việc thực thi thành công thuật toán. Điều này là do việc học vấn đề và giải pháp của nó chỉ xảy ra một lần khi bạn đào tạo thuật toán. Ví dụ, bạn có thể lập trình an toàn cho AI để giải Sudoku (một trò chơi trong đó bạn đặt các con số vào bảng theo các quy tắc nhất định, như đã giải thích tại [Learn-Soduku.com](https://www.learn-sudoku.com/what-is-sudoku.html) ). Bạn thậm chí có thể cung cấp tính linh hoạt cho phép thuật toán chấp nhận nhiều quy tắc hơn hoặc bảng lớn hơn sau này. Peter Norvig, giám đốc nghiên cứu tại Google, đã viết một bài luận thú vị về chủ đề này ( <http://norvig.com/sudoku.html> ) chứng minh cách sử dụng DFS một cách khôn ngoan, giới hạn số lần tính toán (nếu không, các lần tính toán có thể mất gần như mãi mãi), sử dụng các ràng buộc và khám phá các nhánh nhỏ hơn trước tiên có thể giúp giải quyết Sudoku trở nên khả thi.

Thật không may, không phải mọi vấn đề đều có thể dựa vào giải pháp giống Sudoku. Các vấn đề trong đời thực không bao giờ được đặt trong thế giới đơn giản với thông tin hoàn hảo và hành động được xác định rõ ràng. Hãy xem xét vấn đề tìm kẻ gian lận gian lận trong các yêu cầu bồi thường bảo hiểm hoặc vấn đề chẩn đoán bệnh lý. Bạn phải đối mặt với những điều sau:

**Một tập hợp lớn các quy tắc và khả năng:** Số lượng các vụ gian lận có thể xảy ra là cực kỳ cao; nhiều bệnh có các triệu chứng tương tự nhau.

**Thiếu thông tin:** Kẻ gian có thể che giấu thông tin; bác sĩ thường dựa vào thông tin không đầy đủ (có thể thiếu kết quả xét nghiệm).

**Quy luật của vấn đề không phải là bất biến:** Kẻ lừa đảo tìm ra những cách mới để lừa đảo hoặc gian lận; nhiều căn bệnh mới phát sinh hoặc được phát hiện.

Để giải quyết những vấn đề như vậy, bạn không thể sử dụng một cách tiếp cận được định sẵn mà cần một cách tiếp cận linh hoạt và phải tích lũy kiến ​​thức hữu ích để đối mặt với bất kỳ thách thức mới nào. Nói cách khác, bạn tiếp tục học hỏi, như con người

phải làm trong suốt cuộc đời để ứng phó với môi trường thay đổi và đầy thử thách.

## 4. Tận dụng hệ thống chuyên gia

*Hệ thống chuyên gia,* một hệ thống sử dụng các quy tắc để đưa ra quyết định, là nỗ lực đầu tiên thoát khỏi phạm vi của các thuật toán được mã hóa cứng và tạo ra những cách linh hoạt và thông minh hơn để giải quyết các vấn đề trong đời thực. Ý tưởng cốt lõi của hệ thống chuyên gia rất đơn giản và phù hợp vào thời điểm lưu trữ và xử lý nhiều dữ liệu trong bộ nhớ máy tính vẫn còn tốn kém. Nghe có vẻ lạ ngày nay, nhưng vào những năm 1970, các nhà khoa học AI như Ross Quillian đã phải chứng minh cách xây dựng các mô hình ngôn ngữ hoạt động dựa trên vốn từ vựng chỉ có 20 từ, vì bộ nhớ máy tính thời đó chỉ có thể chứa được chừng đó. Có rất ít lựa chọn khả thi nếu máy tính không thể chứa tất cả dữ liệu và giải pháp là xử lý thông tin vấn đề quan trọng và lấy thông tin đó từ những người hiểu rõ nhất.

\* Hệ thống chuyên gia là chuyên gia không phải vì chúng dựa trên kiến ​​thức của mình vào quá trình học tập của chính chúng, mà đúng hơn là vì chúng thu thập kiến ​​thức từ các chuyên gia con người, những người cung cấp một hệ thống thông tin chính được tiêu hóa trước lấy từ việc nghiên cứu sách, học hỏi từ các chuyên gia khác hoặc tự mình khám phá ra. Về cơ bản, đó là một cách thông minh để đưa kiến ​​thức ra bên ngoài vào máy móc.

***MYCIN: Một hệ thống chuyên gia mới bắt đầu***

Một ví dụ về một trong những hệ thống đầu tiên thuộc loại này là MYCIN, một hệ thống chẩn đoán các bệnh đông máu hoặc nhiễm trùng do vi khuẩn gây ra, chẳng hạn như nhiễm khuẩn huyết (khi vi khuẩn xâm nhập vào máu) và viêm màng não (viêm màng bảo vệ não và tủy sống). MYCIN khuyến nghị liều lượng kháng sinh phù hợp bằng cách sử dụng hơn 500 quy tắc và khi cần, hệ thống này dựa vào bác sĩ sử dụng hệ thống. Khi không có đủ thông tin — ví dụ, xét nghiệm trong phòng thí nghiệm bị thiếu — MYCIN sau đó bắt đầu đối thoại tư vấn bằng cách đặt các câu hỏi có liên quan để đưa ra chẩn đoán và liệu pháp đáng tin cậy.

Được viết bằng Lisp như một luận án tiến sĩ của Edward Shortliffe tại Đại học Stanford, MYCIN mất hơn năm năm để hoàn thành và hoạt động tốt hơn bất kỳ bác sĩ mới vào nghề nào, đạt được độ chính xác chẩn đoán cao như một bác sĩ giàu kinh nghiệm. Nó đến từ cùng một phòng thí nghiệm đã phát minh ra DENDRAL, hệ thống chuyên gia đầu tiên từng được tạo ra, vài năm trước đó. DENDRAL, chuyên về hóa học hữu cơ, là một ứng dụng đầy thách thức trong đó các thuật toán dùng vũ lực tỏ ra không khả thi khi đối mặt với các phương pháp tìm kiếm dựa trên kinh nghiệm thực tế của con người.

Đối với thành công của MYCIN, một số vấn đề đã nảy sinh. Đầu tiên, các điều khoản trách nhiệm không rõ ràng. (Nếu hệ thống đưa ra chẩn đoán sai, ai sẽ chịu trách nhiệm?) Thứ hai, MYCIN có vấn đề về khả năng sử dụng vì bác sĩ phải kết nối với MYCIN bằng cách sử dụng thiết bị đầu cuối từ xa đến máy chủ lớn ở Stanford, một việc khá khó khăn và chậm chạp vào thời điểm Internet vẫn còn ở giai đoạn sơ khai. MYCIN vẫn chứng minh được tính hiệu quả và hữu ích của mình trong việc hỗ trợ các quyết định của con người, và nó mở đường cho nhiều hệ thống chuyên gia khác phát triển mạnh vào những năm 1970 và 1980.

***Các thành phần của hệ thống chuyên gia***

Nhìn chung, các hệ thống chuyên gia thời đó được tạo thành từ hai thành phần riêng biệt: cơ sở tri thức và công cụ suy luận. Cơ *sở tri thức* lưu giữ tri thức dưới dạng tập hợp các quy tắc dưới dạng các câu lệnh if-then (với *if* bao gồm một hoặc nhiều điều kiện và *sau đó* bao gồm các câu lệnh kết luận). Các câu lệnh này xuất hiện dưới dạng biểu tượng, phân biệt giữa các trường hợp (sự kiện hoặc sự kiện đơn lẻ), lớp và lớp con, tất cả đều có thể được thao tác bằng logic Boolean hoặc logic bậc nhất phức tạp, bao gồm nhiều phép toán khả thi hơn.

*\* Logic bậc nhất* là một tập hợp các phép toán vượt ra ngoài việc chỉ bị ràng buộc để kết hợp các khẳng định ĐÚNG và SAI. Ví dụ, nó giới thiệu các khái niệm như FOR ALL hoặc THERE EXIST, cho phép bạn xử lý các câu lệnh có thể đúng nhưng không thể chứng minh được bằng bằng chứng bạn có trong tay tại thời điểm đó. Bạn có thể đọc và khám phá thêm về dạng logic này bắt đầu từ bài viết này tại TechTarget.com, “Logic bậc nhất” (yêu cầu đăng ký) hoặc một bài viết khác

một bài có cùng tiêu đề tại [Wikipedia.org](https://en.wikipedia.org/wiki/First-order_logic) (có thể kém chính xác và đầy đủ hơn).

Công *cụ suy luận* là một tập hợp các hướng dẫn cho hệ thống biết cách thao tác các điều kiện dựa trên tập hợp logic Boolean của các toán tử như AND, OR, NOT. Sử dụng tập hợp logic này, các điều kiện biểu tượng TRUE hoặc FALSE có thể kết hợp thành lý luận phức tạp. (Khi TRUE, một quy tắc được kích hoạt hoặc về mặt kỹ thuật là “bị kích hoạt”; khi FALSE, quy tắc không áp dụng.)

Bởi vì hệ thống được tạo ra ở cốt lõi của một loạt các if (điều kiện) và then (kết luận), và được lồng vào nhau và cấu trúc thành nhiều lớp, việc thu thập thông tin ban đầu giúp loại trừ một số kết luận đồng thời cũng giúp hệ thống tương tác với người dùng về thông tin có thể dẫn đến câu trả lời. Khi xử lý công cụ suy luận, các hoạt động chung của hệ thống chuyên gia như sau:

**Chuỗi chuyển tiếp:** Bằng chứng có sẵn kích hoạt một loạt các quy tắc và loại trừ các quy tắc khác ở mỗi giai đoạn. Hệ thống ban đầu tập trung vào các quy tắc có thể kích hoạt kết luận cuối cùng bằng cách kích hoạt. Cách tiếp cận này rõ ràng là dựa trên dữ liệu.

**Chuỗi ngược:** Hệ thống đánh giá mọi kết luận có thể và cố gắng chứng minh từng kết luận dựa trên bằng chứng có sẵn. Cách tiếp cận theo mục tiêu này giúp xác định câu hỏi nào cần đặt ra và loại trừ toàn bộ các mục tiêu. MYCIN, được mô tả trước đó, sử dụng chuỗi ngược; tiến triển từ giả thuyết ngược đến bằng chứng là một chiến lược phổ biến trong chẩn đoán y khoa.

**Giải quyết xung đột:** Nếu một hệ thống đưa ra nhiều hơn một kết luận cùng một lúc, hệ thống sẽ ưu tiên kết luận có một số đặc điểm nhất định (về tác động, rủi ro hoặc các yếu tố khác). Đôi khi hệ thống tham khảo ý kiến ​​người dùng và giải pháp được thực hiện dựa trên đánh giá của người dùng. Ví dụ, MYCIN đã sử dụng một hệ số chắc chắn ước tính xác suất chẩn đoán chính xác.

Một lợi thế lớn của các hệ thống như vậy là thể hiện kiến ​​thức dưới dạng con người có thể đọc được, giúp quá trình ra quyết định trở nên minh bạch. Nếu hệ thống đi đến kết luận, nó sẽ trả về các quy tắc được sử dụng để đi đến kết luận đó. Người dùng có thể xem xét lại công việc của hệ thống một cách có hệ thống và đồng ý hoặc xem xét các dấu hiệu lỗi đầu vào. Hơn nữa, các hệ thống chuyên gia dễ lập trình bằng các ngôn ngữ như Lisp, Prolog hoặc ALGOL. Người dùng đã cải thiện các hệ thống chuyên gia theo thời gian bằng cách thêm các quy tắc mới hoặc cập nhật các quy tắc hiện có. Chúng thậm chí có thể được thực hiện để hoạt động trong các điều kiện không chắc chắn bằng cách áp dụng *logic mờ* , một loại logic đa giá trị trong đó giá trị có thể chứa bất kỳ giá trị nào giữa 0 hoặc hoàn toàn sai và 1 hoặc hoàn toàn đúng (xem "Trí tuệ nhân tạo — Hệ thống logic mờ" tại TutorialsPoint.com để biết thêm chi tiết). Logic mờ tránh các bước đột ngột kích hoạt quy tắc dựa trên ngưỡng. Ví dụ: nếu một quy tắc được đặt để kích hoạt khi phòng nóng, quy tắc đó không được kích hoạt ở nhiệt độ chính xác mà thay vào đó là khi nhiệt độ ở quanh ngưỡng đó.

Hệ thống chuyên gia đã chứng kiến ​​sự lụi tàn vào cuối những năm 1980 và sự phát triển của chúng đã dừng lại, chủ yếu vì những lý do sau:

 Logic và biểu tượng của những hệ thống như vậy tỏ ra hạn chế trong việc thể hiện các quy tắc đằng sau một quyết định, dẫn đến việc tạo ra các hệ thống tùy chỉnh, tức là lại quay trở lại với các quy tắc mã hóa cứng bằng các thuật toán cổ điển.

 Đối với nhiều vấn đề khó khăn, hệ thống chuyên gia trở nên quá phức tạp và rắc rối đến mức mất đi sức hấp dẫn về mặt khả thi và chi phí kinh tế.

 Do dữ liệu ngày càng trở nên phân tán và dễ tiếp cận hơn, nên việc phải vất vả phỏng vấn, thu thập và chắt lọc cẩn thận kiến ​​thức chuyên môn hiếm có không còn ý nghĩa nữa khi có thể sàng lọc được kiến ​​thức tương tự (hoặc thậm chí tốt hơn) từ dữ liệu.

Hệ thống chuyên gia vẫn tồn tại. Bạn có thể thấy chúng được sử dụng trong chấm điểm tín dụng, phát hiện gian lận và các lĩnh vực khác với yêu cầu không chỉ cung cấp câu trả lời mà còn nêu rõ ràng và minh bạch các quy tắc đằng sau quyết định theo cách mà người dùng hệ thống cho là chấp nhận được (như một chuyên gia về chủ đề sẽ làm). Ngoài ra, chúng được sử dụng trong các tình huống mà các hình thức AI khác quá chậm, chẳng hạn như một số ứng dụng xe tự lái (xem “Hệ thống chuyên gia và xe tự lái AI: Các kỹ thuật sáng tạo quan trọng” tại aitrends.com để biết chi tiết).

## 5. Giới thiệu về máy học

Các giải pháp có khả năng học trực tiếp từ dữ liệu mà không cần bất kỳ quá trình tiêu hóa trước nào để biến nó thành các ký hiệu đã xuất hiện trước các hệ thống chuyên gia vài thập kỷ. Một số có bản chất thống kê; một số khác mô phỏng tự nhiên theo những cách khác nhau; và một số khác nữa cố gắng tạo ra logic biểu tượng tự chủ dưới dạng các quy tắc từ thông tin thô. Tất cả các giải pháp này đều bắt nguồn từ các trường phái khác nhau và xuất hiện dưới những tên gọi khác nhau mà ngày nay bao gồm *học máy* . Học máy là một phần của thế giới thuật toán, mặc dù, trái ngược với nhiều thuật toán đã thảo luận cho đến nay, nó không có ý định là một loạt các bước được xác định trước có khả năng giải quyết một vấn đề. Theo quy tắc, học máy giải quyết các vấn đề mà con người không biết cách chi tiết thành các bước, nhưng con người tự nhiên giải quyết được. Một ví dụ về vấn đề như vậy là nhận dạng khuôn mặt trong hình ảnh hoặc một số từ nhất định trong một cuộc thảo luận được nói. Học máy được đề cập trong hầu hết mọi chương của cuốn sách này, nhưng [Chương 9](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-03-considering-the-use-of-algorithms/#page215) đến [11](https://book2best.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-03-considering-the-use-of-algorithms/#page263) dành riêng để tiết lộ cách thức hoạt động của các thuật toán học máy chính, đặc biệt là học sâu, đây là công nghệ thúc đẩy làn sóng ứng dụng AI mới, xuất hiện trên các tiêu đề tin tức hầu như mỗi ngày.

***Chạm đến tầm cao mới***

Vai trò của máy học trong làn sóng thuật toán AI mới là một phần thay thế, một phần bổ sung, các thuật toán hiện có. Máy học hoạt động với các hoạt động đòi hỏi trí thông minh theo quan điểm của con người nhưng không dễ để chính thức hóa thành một chuỗi các bước chính xác. Một ví dụ rõ ràng về vai trò này là sự thành thạo được thể hiện bởi một chuyên gia cờ vây, người mà chỉ cần nhìn thoáng qua đã hiểu được các mối đe dọa và cơ hội của một cấu hình bàn cờ và nắm bắt trực giác các nước đi đúng. (Đọc lịch sử của cờ vây tại <https://www.usgo.org/brief-history-go> ).

Cờ vây là một trò chơi cực kỳ phức tạp đối với AI. Cờ vua có trung bình 35 nước đi có thể đánh giá trên bàn cờ và một ván cờ thường kéo dài hơn 80 nước đi, trong khi một ván cờ vây có khoảng 140 nước đi để đánh giá và một ván cờ thường kéo dài hơn 240 nước đi. Hiện tại, không có sức mạnh tính toán nào tồn tại trên thế giới để tạo ra một không gian trạng thái hoàn chỉnh cho một ván cờ vây. Nhóm DeepMind của Google tại London đã phát triển AlphaGo, một chương trình đã đánh bại một số kỳ thủ cờ vây hàng đầu (xem

<https://deepmind.com/research/alphago/> và [https://www.kdnuggets.com/2020/05/deepmind-gaming-ai-](https://www.kdnuggets.com/2020/05/deepmind-gaming-ai-dominance.html)

[dominate.html). Chương trình không dựa vào](https://www.kdnuggets.com/2020/05/deepmind-gaming-ai-dominance.html) phương pháp tiếp cận thuật toán dựa trên việc tìm kiếm không gian trạng thái rộng lớn, mà thay vào đó sử dụng phương pháp sau:

 Một phương pháp tìm kiếm thông minh dựa trên các thử nghiệm ngẫu nhiên về một nước đi có thể. AI áp dụng DFS nhiều lần để xác định xem kết quả đầu tiên tìm thấy là kết quả dương hay âm (không gian trạng thái không đầy đủ và một phần).

 Thuật toán học sâu xử lý hình ảnh của bàn cờ (nhìn thoáng qua) và đưa ra cả nước đi tốt nhất có thể trong tình huống đó (thuật toán này được gọi là *mạng chính sách* ) và ước tính khả năng AI giành chiến thắng trong trò chơi khi sử dụng nước đi đó (thuật toán này được gọi là *mạng giá trị* ).

 Khả năng học bằng cách xem các ván cờ trước đây của các chuyên gia cờ vây và bằng cách chơi với chính nó. Một phiên bản của chương trình, được gọi là AlphaGo Zero, có thể tự học mà không cần bất kỳ ví dụ nào của con người (xem <https://deepmind.com/blog/alphago-zero-learning-scratch> ).

Khả năng học tập này được gọi là *học tăng cường.*

# Chương 4: Xem xét việc sử dụng phần cứng

**TRONG CHƯƠNG NÀY**

* Sử dụng phần cứng tiêu chuẩn.
* Sử dụng phần cứng chuyên dụng.
* Cải thiện phần cứng.
* Tương tác với môi trường.

Trong Chương 1, bạn sẽ khám phá ra rằng một trong những lý do khiến những nỗ lực AI ban đầu thất bại là do thiếu phần cứng phù hợp. Phần cứng không thể thực hiện các tác vụ đủ nhanh ngay cả đối với những nhu cầu tầm thường, chứ đừng nói đến việc phức tạp như mô phỏng suy nghĩ của con người. Vấn đề này được mô tả khá chi tiết trong bộ phim The Imitation Game (xem phần mô tả tại Amazon.com), trong đó Alan Turing cuối cùng đã bẻ khóa được mã Enigma bằng cách khéo léo tìm kiếm một cụm từ cụ thể, "Heil Hitler", trong mỗi tin nhắn. Nếu không có lỗi cụ thể đó trong cách người vận hành sử dụng Enigma, thiết bị máy tính mà Turing sử dụng sẽ không bao giờ hoạt động đủ nhanh để giải quyết vấn đề (và bộ phim đã có không ít lời phàn nàn về vấn đề này). Nếu có bất kỳ điều gì, thì tài khoản lịch sử — phần nhỏ được giải mật hoàn toàn — cho thấy các vấn đề của Turing sâu sắc hơn những gì bộ phim thể hiện (xem "Cracking the Uncrackable" tại ScienceABC.com để biết chi tiết). May mắn thay, phần cứng tiêu chuẩn, có sẵn có thể khắc phục vấn đề về tốc độ đối với nhiều vấn đề hiện nay, đó là nơi chương này bắt đầu.

Ngay cả với phần cứng tùy chỉnh được thiết kế đặc biệt để tăng tốc độ tính toán, một cỗ máy được thiết kế để mô phỏng suy nghĩ của con người chỉ có thể chạy nhanh như đầu vào và đầu ra của nó cho phép. Do đó, mọi người đang nỗ lực tạo ra một môi trường tốt hơn để phần cứng có thể hoạt động. Nhu cầu này có thể được giải quyết theo một số cách, nhưng chương này sẽ xem xét hai cách: tăng cường khả năng của phần cứng cơ bản và sử dụng các cảm biến chuyên dụng. Những thay đổi này đối với môi trường phần cứng hoạt động tốt, nhưng như tài liệu sau đây giải thích, vẫn chưa đủ để xây dựng bộ não con người. Cuối cùng, phần cứng vô dụng, ngay cả khi được cải tiến, nếu con người dựa vào nó không thể tương tác hiệu quả với nó. Các phần cuối của chương này mô tả các kỹ thuật để làm cho những tương tác đó hiệu quả hơn. Điều đặc biệt quan trọng hiện nay là việc sử dụng Bộ xử lý học sâu (DLP), được thiết kế riêng để hoạt động với các thuật toán học sâu. Tuy nhiên, cũng có những cách tiếp cận tầm thường hơn chỉ đơn giản là kết quả của sự kết hợp giữa đầu ra được cải thiện và lập trình thông minh. Cũng giống như Alan Turing đã sử dụng một mẹo để khiến máy tính của mình dường như làm được nhiều hơn những gì nó có thể làm, những kỹ thuật này khiến máy tính hiện đại trông giống như những người làm phép lạ. Trên thực tế, máy tính không hiểu gì cả; mọi công lao đều thuộc về những người lập trình máy tính.

## Dựa vào phần cứng tiêu chuẩn

Hầu hết các dự án AI mà bạn tạo ra ít nhất sẽ bắt đầu bằng phần cứng tiêu chuẩn vì các thành phần hiện đại có sẵn thực sự cung cấp sức mạnh xử lý đáng kể, đặc biệt là khi so sánh với các thành phần từ những năm 1980 khi AI lần đầu tiên bắt đầu tạo ra kết quả có thể sử dụng được. Do đó, ngay cả khi cuối cùng bạn không thể thực hiện công việc ở cấp độ sản xuất bằng cách sử dụng phần cứng tiêu chuẩn, bạn vẫn có thể tiến đủ xa với mã thử nghiệm và tiền sản xuất của mình để tạo ra một mô hình hoạt động cuối cùng sẽ xử lý toàn bộ tập dữ liệu.

**Hiểu về phần cứng tiêu chuẩn**

Kiến trúc (cấu trúc) của PC tiêu chuẩn không thay đổi kể từ lần đầu tiên John von Neumann đề xuất vào năm 1946 (xem bài viết “John von Neumann: Cha đẻ của máy tính hiện đại” tại <https://www.maa.org/external_archive/devlin/devlin_12_03.html> để biết chi tiết). Xem lại lịch sử tại <https://lennartb.home.xs4all.nl/coreboot/col2.html> cho bạn thấy rằng bộ xử lý kết nối với bộ nhớ và các thiết bị ngoại vi **thông qua một bus** trong các sản phẩm PC từ năm 1981 (và lâu trước đó). Tất cả các hệ thống này đều sử dụng kiến ​​trúc von Neumann vì kiến ​​trúc này mang lại những lợi ích đáng kể về tính mô-đun. Đọc lịch sử cho bạn biết rằng các thiết bị này cho phép nâng cấp mọi thành phần dưới dạng các quyết định riêng lẻ, cho phép tăng khả năng. Ví dụ: trong giới hạn, bạn có thể tăng dung lượng bộ nhớ hoặc dung lượng lưu trữ khả dụng cho bất kỳ PC nào. Bạn cũng có thể sử dụng các thiết bị ngoại vi tiên tiến. Tuy nhiên, tất cả các thành phần này đều kết nối thông qua một bus.

**Mô tả các khiếm khuyết phần cứng tiêu chuẩn**

Khả năng tạo ra một hệ thống mô-đun có những lợi ích đáng kể, đặc biệt là trong kinh doanh. Khả năng loại bỏ và thay thế các thành phần riêng lẻ giúp giảm chi phí trong khi cho phép cải thiện gia tăng cả về tốc độ và hiệu quả. Tuy nhiên, giống như hầu hết mọi thứ, không có bữa trưa miễn phí. Tính mô-đun do kiến ​​trúc von Neumann cung cấp đi kèm với một số thiếu sót nghiêm trọng:

1. Nút thắt cổ chai von Neumann: Trong tất cả các khiếm khuyết, nút thắt cổ chai von Neumann là nghiêm trọng nhất khi xem xét các yêu cầu của các ngành như AI, học máy và thậm chí là khoa học dữ liệu. Bạn có thể tìm thấy khiếm khuyết cụ thể này được thảo luận chi tiết hơn trong phần “Xem xét nút thắt cổ chai von Neumann” ở phần sau của chương này.

2. Điểm lỗi đơn: Bất kỳ mất kết nối nào với bus đều có nghĩa là máy tính sẽ hỏng ngay lập tức, thay vì hỏng một cách nhanh chóng. Ngay cả trong các hệ thống có nhiều bộ xử lý, việc mất một bộ xử lý duy nhất, vốn chỉ đơn giản là gây ra mất khả năng, thay vào đó lại gây ra lỗi toàn bộ hệ thống. Vấn đề tương tự cũng xảy ra với việc mất các thành phần hệ thống khác: Thay vì giảm chức năng, toàn bộ hệ thống sẽ hỏng. Vì AI thường yêu cầu hoạt động liên tục của hệ thống, nên khả năng xảy ra hậu quả nghiêm trọng sẽ tăng theo cách mà ứng dụng dựa vào phần cứng.

3. Đơn nhất: Bus von Neumann có thể lấy lệnh hoặc lấy dữ liệu cần thiết để thực hiện lệnh, nhưng không thể làm cả hai. Do đó, khi việc lấy dữ liệu đòi hỏi nhiều chu kỳ bus, bộ xử lý sẽ ở trạng thái nhàn rỗi, làm giảm khả năng thực hiện các tác vụ AI đòi hỏi nhiều lệnh hơn nữa.

|  |
| --- |
| **KIỂM TRA SỰ KHÁC BIỆT CỦA KIẾN TRÚC HARVARD** Bạn có thể gặp kiến ​​trúc Harvard trong quá trình di chuyển phần cứng của mình vì một số hệ thống sử dụng dạng đã sửa đổi của kiến ​​trúc này để tăng tốc độ xử lý. Cả kiến ​​trúc von Neumann và kiến ​​trúc Harvard đều dựa trên cấu trúc bus. Tuy nhiên, khi làm việc với hệ thống kiến ​​trúc von Neumann, phần cứng dựa trên một bus duy nhất và một vùng bộ nhớ duy nhất cho cả lệnh và dữ liệu, trong khi kiến ​​trúc Harvard dựa trên các bus riêng lẻ cho lệnh và dữ liệu và có thể sử dụng các vùng bộ nhớ vật lý riêng biệt (xem phần so sánh trong “Sự khác biệt giữa kiến ​​trúc Von Neumann và Harvard” tại GeeksforGeeks.org). Việc sử dụng các bus riêng lẻ cho phép hệ thống kiến ​​trúc Harvard truy xuất lệnh tiếp theo trong khi chờ dữ liệu đến từ bộ nhớ cho lệnh hiện tại, do đó làm cho kiến ​​trúc Harvard vừa nhanh hơn vừa hiệu quả hơn. Tuy nhiên, độ tin cậy bị ảnh hưởng vì giờ đây bạn có hai điểm lỗi cho mỗi thao tác: bus lệnh và bus dữ liệu. Các bộ vi điều khiển, chẳng hạn như các bộ vi điều khiển cung cấp năng lượng cho lò vi sóng của bạn, thường sử dụng kiến ​​trúc Harvard. Ngoài ra, bạn có thể tìm thấy nó ở một số nơi bất thường vì một lý do cụ thể. iPhone và Xbox 360 đều sử dụng các phiên bản đã sửa đổi của kiến ​​trúc Harvard dựa trên một vùng bộ nhớ duy nhất (thay vì hai), nhưng vẫn dựa trên các bus riêng biệt. Lý do sử dụng kiến ​​trúc trong trường hợp này là Quản lý bản quyền kỹ thuật số (DRM). Bạn có thể biến vùng mã của bộ nhớ thành chỉ đọc để không ai có thể sửa đổi hoặc tạo ứng dụng mới mà không được phép. Theo quan điểm của AI, điều này có thể gây ra vấn đề vì một trong những khả năng của AI là viết các thuật toán mới (mã thực thi) khi cần để xử lý các tình huống không lường trước được. Vì PC hiếm khi triển khai kiến ​​trúc Harvard ở dạng thuần túy hoặc là cấu trúc bus chính của nó, nên kiến ​​trúc Harvard không nhận được nhiều sự chú ý trong cuốn sách này. |

4. Nhiệm vụ: Khi não thực hiện một nhiệm vụ, một số khớp thần kinh sẽ kích hoạt cùng một lúc, cho phép thực hiện đồng thời nhiều thao tác. Thiết kế von Neumann ban đầu chỉ cho phép thực hiện một thao tác tại một thời điểm và chỉ sau khi hệ thống đã lấy được cả hướng dẫn và dữ liệu cần thiết. Máy tính ngày nay thường có nhiều lõi, cho phép thực hiện đồng thời các thao tác trong mỗi lõi. Tuy nhiên, mã ứng dụng phải giải quyết cụ thể yêu cầu này, do đó chức năng đôi khi vẫn không được sử dụng.

**Dựa vào các kỹ thuật tính toán mới**

Đọc tài liệu về cách thực hiện nhiệm vụ bằng AI có thể khiến bạn cảm thấy như đang xem một nhà tiếp thị trên TV tuyên bố, "Nó mới! Nó đã được cải thiện! Nó thực sự chói sáng!" Vì vậy, bạn không nên ngạc nhiên nhiều khi mọi người luôn nghĩ ra những cách để làm cho trải nghiệm phát triển AI nhanh hơn, chính xác hơn và tốt hơn theo những cách khác. Vấn đề là nhiều kỹ thuật mới này chưa được thử nghiệm, vì vậy chúng có thể trông tuyệt vời, nhưng bạn phải suy nghĩ về chúng trong một thời gian. Một cách giải quyết các vấn đề khác nhau xung quanh tốc độ tính toán của AI là tạo ra các kỹ thuật mới để thực hiện các nhiệm vụ. Mặc dù nhiều nhà khoa học dữ liệu dựa vào Bộ xử lý đồ họa (GPU) để tăng tốc độ thực thi mã phức tạp, bài viết "The startup making deep learning possiblewithout specialist hardware" tại MIT Technology Review.com mô tả một cách tiếp cận khác, dựa trên một sản phẩm có tên là Neural Magic ( [https://neuralmagic.com](https://neuralmagic.com/) ), về cơ bản là nén dữ liệu để làm cho CPU hiệu quả hơn. Neural Magic cũng giúp giảm chi phí so với việc sử dụng phần cứng chuyên dụng. (Phần cứng càng chuyên dụng thì chi phí càng cao.)

Ưu điểm của việc sử dụng phương pháp Neural Magic là một tổ chức có thể chỉ cần mua một vài máy móc đắt tiền để thực hiện nghiên cứu và tạo ứng dụng. Sau đó, tổ chức có thể chạy ứng dụng kết quả trên nhiều hệ thống giá rẻ tùy theo nhu cầu để đáp ứng yêu cầu của người dùng. Lợi ích lớn nhất là các hệ thống này không cần dựa vào công nghệ máy tính để bàn mà còn có thể sử dụng cả thiết bị di động, do đó ứng dụng có thể chạy ở bất kỳ đâu. Do đó, đây là một phương pháp có giá trị trong phạm vi công nghệ mà nó hiện đang sử dụng. Một phương pháp mới khác dựa vào việc sử dụng bảng băm thay vì ma trận để mô hình hóa các vấn đề. Theo bài viết “Thuật toán CPU đào tạo mạng nơ-ron sâu nhanh hơn tới 15 lần so với các trình đào tạo GPU hàng đầu” trên TechXplore.com, Công cụ học sâu dưới tuyến tính (SLIDE) có thể đào tạo các mô hình bằng bộ xử lý hàng hóa thay vì sử dụng GPU. Bên cạnh việc sử dụng bảng băm thay cho phép nhân ma trận, việc sử dụng SLIDE cũng loại bỏ một số yếu tố lãng phí hơn khi đào tạo mô hình (xem bài viết “Đột ​​phá về học sâu: thuật toán học sâu dưới tuyến tính không cần GPU?” trên KD-nuggets.com). Vấn đề với cách tiếp cận mới này (cũng như nhiều cách tiếp cận mới khác) là nó đòi hỏi phải thay đổi hoàn toàn cách thực hiện nhiệm vụ. Rõ ràng là các tổ chức sẽ không vui khi phải chi hàng triệu đô la cho quá trình phát triển hiện tại để thử một cái gì đó mới. Sách trắng “SLIDE: Bảo vệ thuật toán thông minh hơn là tăng tốc phần cứng cho hệ thống học sâu quy mô lớn” tại arXiv.org cung cấp một cuộc thảo luận chính thức hơn về phương pháp luận mới này.

|  |
| --- |
| **XEM XÉT MÁY BOMBE CỦA ALAN TURING** Máy Bombe của Alan Turing không phải là bất kỳ dạng AI nào. Trên thực tế, nó thậm chí không phải là một máy tính thực sự. Nó đã phá vỡ các thông điệp mật mã Enigma và thế là hết. Tuy nhiên, nó đã cung cấp thức ăn cho suy nghĩ của Turing, cuối cùng dẫn đến một bài báo có tiêu đề "Máy tính và trí thông minh". Turing đã xuất bản bài báo đó, mô tả trò chơi bắt chước, vào những năm 1950 (bộ phim The Imitation Game là mô tả về các sự kiện xung quanh việc tạo ra trò chơi này). Tuy nhiên, bản thân Bombe thực sự dựa trên một cỗ máy của Ba Lan có tên là Bomba. Mặc dù một số nguồn ngụ ý rằng Alan Turing đã làm việc một mình, nhưng Bombe được sản xuất với sự giúp đỡ của nhiều người, đặc biệt là Gordon Welchman. Turing cũng không xuất hiện từ chân không, được chế tạo sẵn để phá vỡ mã hóa của Đức. Thời gian của ông tại Princeton đã dành cho những người vĩ đại như Albert Einstein và John von Neumann (người sau này sẽ phát minh ra khái niệm phần mềm máy tính). Các bài báo Turing viết đã truyền cảm hứng cho những nhà khoa học khác này để thử nghiệm và xem điều gì là khả thi. Phần cứng chuyên dụng đủ loại sẽ tiếp tục xuất hiện miễn là các nhà khoa học vẫn viết các bài báo, trao đổi ý tưởng với nhau, tạo ra những ý tưởng mới của riêng họ và thử nghiệm. Khi bạn xem phim hoặc các phương tiện truyền thông khác, giả sử rằng chúng hoàn toàn chính xác về mặt lịch sử, đừng rời đi với cảm giác rằng những người này vừa thức dậy vào một buổi sáng, tuyên bố, "Hôm nay, tôi sẽ trở nên xuất sắc!" và tiếp tục làm một điều gì đó tuyệt vời. Mọi thứ đều xây dựng trên một thứ khác, vì vậy lịch sử rất quan trọng vì nó giúp chỉ ra con đường đã đi qua và làm sáng tỏ những con đường đầy hứa hẹn khác — những con đường chưa được đi qua. |

## Sử dụng GPU – phần cứng chuyên dụng

Sau khi tạo một thiết lập nguyên mẫu để thực hiện các tác vụ cần thiết để mô phỏng suy nghĩ của con người về một chủ đề nhất định, bạn có thể cần phần cứng bổ sung để cung cấp đủ sức mạnh xử lý để làm việc với toàn bộ tập dữ liệu cần thiết của một hệ thống sản xuất. Có nhiều cách để cung cấp sức mạnh xử lý như vậy, nhưng một cách phổ biến là sử dụng Bộ xử lý đồ họa (GPU) ngoài bộ xử lý trung tâm của máy. Các phần sau đây mô tả miền vấn đề mà GPU giải quyết, chính xác thuật ngữ GPU có nghĩa là gì và tại sao GPU giúp xử lý nhanh hơn.

**Xem xét nút thắt cổ chai von Neumann**

Nút thắt cổ chai von Neumann là kết quả tự nhiên của việc sử dụng bus để truyền dữ liệu giữa bộ xử lý, bộ nhớ, bộ lưu trữ dài hạn và các thiết bị ngoại vi. Bất kể bus thực hiện nhiệm vụ của mình nhanh đến mức nào, việc quá tải bus — tức là tạo ra nút thắt cổ chai làm giảm tốc độ — luôn có thể xảy ra. Theo thời gian, tốc độ bộ xử lý tiếp tục tăng trong khi bộ nhớ và các cải tiến thiết bị khác tập trung vào mật độ — khả năng lưu trữ nhiều hơn trong không gian nhỏ hơn. Do đó, nút thắt cổ chai trở thành vấn đề lớn hơn với mọi cải tiến, khiến bộ xử lý dành nhiều thời gian ở chế độ chờ. Trong phạm vi hợp lý, bạn có thể khắc phục một số vấn đề xung quanh nút thắt cổ chai von Neumann và tạo ra sự gia tăng nhỏ nhưng đáng chú ý về tốc độ ứng dụng. Sau đây là các giải pháp phổ biến nhất:

- Bộ nhớ đệm: Khi các vấn đề với việc lấy dữ liệu từ bộ nhớ đủ nhanh với kiến ​​trúc von Neumann trở nên rõ ràng, các nhà cung cấp phần cứng đã nhanh chóng phản ứng bằng cách thêm bộ nhớ cục bộ không yêu cầu truy cập bus. Bộ nhớ này xuất hiện bên ngoài bộ xử lý nhưng là một phần của gói bộ xử lý. Tuy nhiên, bộ nhớ đệm tốc độ cao rất đắt tiền, do đó kích thước bộ nhớ đệm có xu hướng nhỏ.

- Bộ nhớ đệm bộ xử lý: Thật không may, bộ nhớ đệm ngoài vẫn không cung cấp đủ tốc độ. Ngay cả khi sử dụng RAM nhanh nhất hiện có và cắt hoàn toàn quyền truy cập bus cũng không đáp ứng được nhu cầu về khả năng xử lý của bộ xử lý. Do đó, các nhà cung cấp bắt đầu thêm bộ nhớ trong — một bộ nhớ đệm nhỏ hơn bộ nhớ đệm ngoài, nhưng có khả năng truy cập thậm chí còn nhanh hơn vì nó là một phần của bộ xử lý.

- Prefetching: Vấn đề với bộ nhớ đệm là chúng chỉ hữu ích khi chúng chứa dữ liệu chính xác. Thật không may, số lần truy cập bộ nhớ đệm thấp trong các ứng dụng sử dụng nhiều dữ liệu và thực hiện nhiều tác vụ khác nhau. Bước tiếp theo để bộ xử lý hoạt động nhanh hơn là đoán dữ liệu nào ứng dụng sẽ yêu cầu tiếp theo và tải dữ liệu đó vào bộ nhớ đệm trước khi ứng dụng yêu cầu.

-Sử dụng RAM chuyên dụng: Bạn có thể bị chôn vùi trong mớ chữ cái RAM vì có nhiều loại RAM hơn hầu hết mọi người tưởng tượng. Mỗi loại RAM đều có mục đích giải quyết ít nhất một phần của vấn đề nút thắt cổ chai von Neumann và chúng hoạt động — trong giới hạn. Trong hầu hết các trường hợp, các cải tiến xoay quanh ý tưởng lấy dữ liệu từ bộ nhớ và đưa vào bus nhanh hơn. Hai yếu tố chính (và nhiều yếu tố phụ) ảnh hưởng đến tốc độ: tốc độ bộ nhớ (bộ nhớ di chuyển dữ liệu nhanh như thế nào) và độ trễ (mất bao lâu để định vị một phần dữ liệu cụ thể). Bạn có thể đọc thêm về bộ nhớ và các yếu tố ảnh hưởng đến bộ nhớ trong “Các loại RAM khác nhau và cách sử dụng” tại Computer Memory Upgrade.net.

**Định nghĩa GPU**

Mục đích ban đầu của GPU là xử lý dữ liệu hình ảnh nhanh chóng và sau đó hiển thị hình ảnh kết quả trên màn hình. Trong giai đoạn đầu của quá trình tiến hóa PC, CPU thực hiện tất cả các quá trình xử lý, điều đó có nghĩa là đồ họa có thể xuất hiện chậm trong khi CPU thực hiện các tác vụ khác. Trong thời gian này, PC thường được trang bị bộ điều hợp hiển thị, chứa ít hoặc không có sức mạnh xử lý. Bộ điều hợp hiển thị chỉ chuyển đổi dữ liệu máy tính thành dạng trực quan. Trên thực tế, việc chỉ sử dụng một bộ xử lý đã chứng tỏ là gần như không thể sau khi PC vượt qua màn hình chỉ có văn bản hoặc đồ họa 16 màu cực kỳ đơn giản. Tuy nhiên, GPU thực sự không tạo ra nhiều đột phá trong điện toán cho đến khi mọi người bắt đầu muốn có đầu ra 3D. Vào thời điểm này, sự kết hợp giữa CPU và bộ điều hợp hiển thị đơn giản là không thể thực hiện được công việc. Một bước đầu tiên theo hướng này đã được thực hiện bởi các hệ thống như Hauppauge 4860 (xem chi tiết tại Geekdot.com), bao gồm một CPU và một chip đồ họa đặc biệt (80860, trong trường hợp này) trên bo mạch chủ. 80860 có ưu điểm là thực hiện các phép tính cực kỳ nhanh (xem “Dòng CPU Intel 80860 (i860)” tại CPU-World.com để biết chi tiết). Thật không may, các hệ thống đa xử lý, không đồng bộ này không đáp ứng được kỳ vọng của mọi người đối với chúng (mặc dù chúng cực kỳ nhanh đối với các hệ thống thời đó) và chúng tỏ ra cực kỳ đắt đỏ. Thêm vào đó, còn có toàn bộ vấn đề về việc viết các ứng dụng bao gồm chip thứ hai (hoặc các chip tiếp theo). Hai chip cũng chia sẻ bộ nhớ (rất nhiều cho các hệ thống này). GPU di chuyển xử lý đồ họa từ bo mạch chủ sang bo mạch ngoại vi đồ họa. CPU có thể yêu cầu GPU thực hiện một tác vụ, sau đó GPU xác định phương pháp tốt nhất để thực hiện tác vụ đó mà không cần CPU. GPU có bộ nhớ riêng và đường dẫn dữ liệu cho bus của nó rất lớn. Ngoài ra, GPU có thể truy cập bộ nhớ chính để lấy dữ liệu cần thiết để thực hiện tác vụ và đăng kết quả mà không cần CPU. Do đó, thiết lập này giúp màn hình đồ họa hiện đại trở nên khả thi.

**Xem xét lý do tại sao GPU hoạt động tốt**

Giống như chip 80860 được mô tả trong phần trước, GPU ngày nay rất xuất sắc trong việc thực hiện các tác vụ chuyên biệt liên quan đến xử lý đồ họa, bao gồm làm việc với các vectơ. Tất cả các lõi thực hiện các tác vụ song song thực sự tăng tốc các phép tính AI. Ví dụ, chúng không thể thiếu trong việc tạo ra các mô hình AI chuyên sâu về tính toán, như Mạng đối nghịch tạo sinh (GAN) thực hiện các tác vụ như các tác vụ được mô tả trong bài viết "18 ứng dụng ấn tượng của Mạng đối nghịch tạo sinh (GAN)" tại Machine Learning Mastery.com. Năm 2011, Dự án Google Brain ( <https://research.google.com/teams/brain/> ) đã đào tạo một AI để nhận ra sự khác biệt giữa mèo và người bằng cách xem phim trên YouTube. Tuy nhiên, để thực hiện nhiệm vụ này, Google đã sử dụng 2.000 CPU tại một trong những trung tâm dữ liệu khổng lồ của Google. Rất ít người có đủ nguồn lực cần thiết để sao chép công trình của Google. Mặt khác, Bryan Catanzaro (nhóm nghiên cứu của NVIDIA) và Andrew Ng (Stanford) đã có thể sao chép công trình của Google bằng cách sử dụng bộ 12 GPU NVIDIA (xem bài đăng “Tăng tốc AI bằng GPU: Mô hình điện toán mới” trên blog NVIDIA.com để biết chi tiết). Sau khi mọi người hiểu rằng GPU có thể thay thế nhiều hệ thống máy tính được trang bị CPU, họ có thể bắt đầu tiến hành nhiều dự án AI khác nhau. Năm 2012, Alex Krizhevsky (Đại học Toronto) đã giành chiến thắng trong cuộc thi nhận dạng hình ảnh máy tính ImageNet bằng cách sử dụng GPU. Trên thực tế, một số nhà nghiên cứu hiện đã sử dụng GPU và đạt được thành công đáng kinh ngạc (xem “9 bài báo về học sâu mà bạn cần biết” tại https://adeshpande3.github.io/The-9-Deep-Learning-Papers-YouNeed-To-Know-About.html để biết chi tiết).

**Làm việc với Bộ xử lý học sâu (DLP)**

Các nhà nghiên cứu tham gia vào cuộc đấu tranh liên tục để khám phá ra những cách tốt hơn để đào tạo, xác minh và thử nghiệm các mô hình được sử dụng để tạo ra các ứng dụng AI. Một trong những cách đó là sử dụng các kỹ thuật điện toán mới, như đã mô tả trong phần "Dựa vào các kỹ thuật tính toán mới" ở đầu chương này. Một cách khác là tăng thêm sức mạnh xử lý cho vấn đề, chẳng hạn như sử dụng GPU. Tuy nhiên, GPU chỉ có lợi vì nó có thể thực hiện thao tác ma trận nhanh chóng và ở cấp độ song song hàng loạt. Nếu không, việc sử dụng GPU cũng có thể tạo ra vấn đề, như đã thảo luận trong phần "Sử dụng GPU" của chương này. Vì vậy, việc tìm kiếm thứ gì đó tốt hơn vẫn đang diễn ra và bạn có thể tìm thấy một bảng chữ cái thực sự về các loại bộ xử lý được mô tả trên các trang web như Primo.ai, với trang này: "Đơn vị xử lý - CPU, GPU, APU, TPU, VPU, FPGA, QPU." Trang tài nguyên này sẽ giúp bạn làm quen với tất cả các loại bộ xử lý hiện tại. Tuy nhiên, bạn nên bắt đầu với phần tổng quan được cung cấp trong các phần sau vì rất dễ bị sa lầy vào quá nhiều lựa chọn (và sau đó đầu bạn sẽ nổ tung). Định nghĩa Bộ xử lý học sâu DLPA (DLP) chỉ đơn giản là một bộ xử lý chuyên dụng cung cấp một số lợi thế trong việc đào tạo, xác minh, thử nghiệm và chạy các ứng dụng AI. Chúng cố gắng tạo ra một môi trường mà các ứng dụng AI chạy nhanh ngay cả trên các thiết bị nhỏ hơn hoặc có khả năng kém hơn. Hầu hết các DLP đều tuân theo một mô hình tương tự bằng cách cung cấp Các vùng bộ nhớ dữ liệu và mã riêng biệt Các bus dữ liệu và mã riêng biệt Các bộ lệnh chuyên dụng Bộ nhớ trên chip lớn Bộ đệm lớn để khuyến khích các mô hình tái sử dụng dữ liệu Vào năm 2014, Tianshi Chen (và những người khác) đã đề xuất DLP đầu tiên, được gọi là DianNoa (tiếng Trung có nghĩa là não điện), trong một sách trắng tại <http://novel.ict.ac.cn/ychen/pdf/DianNao.pdf> . Tất nhiên, lần thử đầu tiên không bao giờ là đủ tốt, vì vậy có cả một họ chip DianNoa: DaDianNao, ShiDianNao và PuDianNao (và có thể còn nhiều loại khác nữa).

**Sử dụng Bộ xử lý thần kinh di động (NPU)**

Một số thiết bị di động, đặc biệt là các thiết bị của Huawei và Samsung, có một Bộ xử lý nơ-ron (NPU) ngoài một CPU chung để thực hiện các tác vụ dự đoán AI bằng các mô hình như Mạng nơ-ron nhân tạo (ANN) và Rừng ngẫu nhiên (RF). Bạn không thể sử dụng NPU cho các nhu cầu tính toán chung vì nó quá chuyên biệt. Tuy nhiên, NPU thường hoạt động nhanh hơn GPU tới mười lần cho cùng một tác vụ. NPU được chuyên biệt theo những cách sau: Nó tăng tốc độ chạy các mô hình được xác định trước (trái ngược với đào tạo, xác minh và thử nghiệm) Nó được thiết kế để sử dụng với các thiết bị nhỏ Nó tiêu thụ ít điện năng khi so sánh với các loại bộ xử lý khác Nó sử dụng tài nguyên, chẳng hạn như bộ nhớ, một cách hiệu quả Vì ranh giới chính xác giữa các loại bộ xử lý rất khó xác định, bạn có thể thấy một số NPU giống hoặc thay thế được phân loại là NPU. Tuy nhiên, sau đây là danh sách các bộ xử lý mà hiện tại bạn có thể phân loại là NPU thực sự: Ali-NPU của AlibabaAscend, của HuaweiNeural Engine, của AppleNeural Processing Unit (NPU), của SamsungNNP, Myriad, EyeQ của IntelNVDLA (chủ yếu được sử dụng cho các thiết bị Internet vạn vật [IoT]), của NVIDIA.

**Truy cập vào Tenser Processing Unit (TPU) dựa trên đám mây**

Google đã thiết kế riêng Bộ xử lý Tensor (TPU) vào năm 2015 để chạy nhanh hơn các ứng dụng được xây dựng trên nền tảng TensorFlow. Nó đại diện cho một chuyên môn hóa chip thực sự ở chỗ bạn không thể sử dụng nó hiệu quả nếu không có TensorFlow. Tuy nhiên, nó khác ở một điểm khác là nó là Mạch tích hợp dành riêng cho ứng dụng (ASIC), chứ không phải là chip loại CPU hoàn chỉnh. Những điểm khác biệt rất quan trọng: Một ASIC chỉ có thể thực hiện một tác vụ và bạn không thể thay đổi tác vụ đó. Do chuyên môn hóa của mình, ASIC thường rẻ hơn nhiều so với CPU. Hầu hết các triển khai ASIC đều nhỏ hơn nhiều so với cùng một triển khai được tạo bằng CPU. So với triển khai CPU, ASIC tiết kiệm điện hơn. ASIC cực kỳ đáng tin cậy.

**Tạo ra một môi trường xử lý chuyên biệt**

Học sâu và AI đều là những quy trình không phải của von Neumann, theo nhiều chuyên gia, bao gồm cả Massimiliano Versace, CEO của Neurala Inc. ( <https://www.neurala.com/> ). Vì nhiệm vụ mà thuật toán thực hiện không khớp với phần cứng cơ bản nên tồn tại đủ loại tình trạng kém hiệu quả, cần phải hack và việc thu được kết quả khó hơn nhiều so với mức cần thiết. Do đó, việc thiết kế phần cứng khớp với phần mềm khá hấp dẫn. Cơ quan các dự án nghiên cứu tiên tiến quốc phòng (DARPA) đã thực hiện một dự án như vậy dưới hình thức Hệ thống điện tử có thể mở rộng bằng nhựa thích ứng thần kinh (SyNAPSE). Ý tưởng đằng sau cách tiếp cận này là sao chép cách tiếp cận của tự nhiên để giải quyết vấn đề bằng cách kết hợp bộ nhớ và sức mạnh xử lý, thay vì giữ hai thứ này riêng biệt. Họ thực sự đã xây dựng hệ thống (nó rất lớn), và bạn có thể đọc thêm về nó tại <https://www.darpa.mil/program/systems-ofneuromorphic-adaptive-plastic-scalable-electronics> và [https://www.darpa.mil/news-events/2014-08-07](https://www.darpa.mil/news-events/2014-08-07.The) . Dự án SyNAPSE đã tiến triển. IBM đã xây dựng một hệ thống nhỏ hơn bằng cách sử dụng công nghệ hiện đại vừa cực kỳ nhanh vừa tiết kiệm điện (xem <https://www.research.ibm.com/articles/brain-chip.shtml> ). Vấn đề duy nhất là không ai mua chúng. Cũng giống như nhiều người cho rằng Betamax là cách lưu trữ dữ liệu tốt hơn VHS, VHS đã chiến thắng về chi phí, tính dễ sử dụng và các tính năng hấp dẫn (xem “Betamax so với VHS: Sony đã thua cuộc chiến định dạng video gia đình ban đầu như thế nào” tại GIZMODO.com). Điều tương tự cũng đúng đối với sản phẩm SyNAPSE của IBM, TrueNorth. Thật khó để tìm được những người sẵn sàng trả mức giá cao hơn, những lập trình viên có thể phát triển phần mềm bằng kiến ​​trúc mới và những sản phẩm thực sự được hưởng lợi từ con chip. Do đó, sự kết hợp giữa CPU và GPU, ngay cả với những điểm yếu cố hữu của nó, vẫn tiếp tục chiến thắng.

## Tăng khả năng phần cứng

CPU vẫn hoạt động tốt đối với các hệ thống kinh doanh hoặc trong các ứng dụng mà nhu cầu về tính linh hoạt chung trong lập trình lớn hơn sức mạnh xử lý thuần túy. Tuy nhiên, GPU hiện là tiêu chuẩn cho nhiều loại khoa học dữ liệu, học máy, AI và nhu cầu học sâu. Tất nhiên, mọi người liên tục tìm kiếm điều lớn lao tiếp theo trong môi trường phát triển. Cả CPU và GPU đều là bộ xử lý cấp sản xuất. Trong tương lai, bạn có thể thấy một trong hai loại bộ xử lý được sử dụng thay cho các tiêu chuẩn này: Mạch tích hợp dành riêng cho ứng dụng (ASIC): Trái ngược với bộ xử lý chung, nhà cung cấp tạo ra ASIC cho một mục đích cụ thể. Giải pháp ASIC cung cấp hiệu suất cực nhanh chỉ sử dụng rất ít năng lượng, nhưng lại thiếu tính linh hoạt. Bạn có thể tìm thấy ví dụ về ASIC trước đó trong chương này dưới dạng TPU (xem phần "Truy cập Đơn vị xử lý Tenser dựa trên đám mây (TPU)" để biết chi tiết). Mảng cổng lập trình trường (FPGA): Giống như ASIC, nhà cung cấp thường chế tạo FPGA cho một mục đích cụ thể. Tuy nhiên, trái ngược với ASIC, bạn có thể lập trình FPGA để thay đổi chức năng cơ bản của nó. Một ví dụ về giải pháp FPGA là Brainwave của Microsoft, được sử dụng cho các dự án học sâu (xem “Microsoft Brainwave có mục tiêu tăng tốc học sâu bằng FPGA” tại TechCrunch.com).

Các nhà cung cấp cũng đang nghiên cứu các loại xử lý hoàn toàn mới, có thể hoặc không thực sự hoạt động như mong đợi. Ví dụ, Graphcore đang nghiên cứu một Bộ xử lý thông minh (IPU), như được mô tả tại https://www.prnewswire.com/news-releases/sequoia-backsgraphcore-as-the-future-of-artificial-intelligence-processors- 300554316.html. Công ty đã phát triển dòng bộ xử lý được hiển thị tại <https://www.graphcore.ai/products/ipu> . Tuy nhiên, bạn phải xem xét tin tức về các bộ xử lý mới này một cách thận trọng, vì sự cường điệu đã bao quanh ngành công nghiệp này trong quá khứ. Khi bạn thấy các ứng dụng thực tế từ các công ty lớn như Google và Microsoft, bạn có thể bắt đầu cảm thấy chắc chắn hơn một chút về tương lai của công nghệ liên quan.

**Thêm cảm biến chuyên dụng**

Một thành phần thiết yếu của AI là khả năng của AI trong việc mô phỏng trí thông minh của con người bằng cách sử dụng một bộ đầy đủ các giác quan. Đầu vào được cung cấp thông qua các giác quan giúp con người phát triển các loại trí thông minh khác nhau được mô tả trong Chương 1. Các giác quan của con người cung cấp đúng loại đầu vào để tạo ra một con người thông minh. Ngay cả khi giả sử rằng AI có thể triển khai đầy đủ cả bảy loại trí thông minh, thì nó vẫn cần đúng loại đầu vào để trí thông minh đó hoạt động. Con người thường có năm giác quan để tương tác với môi trường: thị giác, thính giác, xúc giác, vị giác và thính giác. Thật kỳ lạ, con người vẫn chưa hiểu hết khả năng của chính mình, vì vậy không có gì ngạc nhiên khi máy tính chậm hơn khi nói đến việc cảm nhận môi trường theo cùng một cách mà con người làm. Ví dụ, cho đến gần đây, vị giác chỉ bao gồm bốn thành phần: mặn, ngọt, đắng và chua. Tuy nhiên, hiện có thêm hai vị nữa xuất hiện trong danh sách: umami và chất béo (xem “Ngọt, Chua, Mặn, Đắng, Umami … Và Chất béo?” tại FiveThirtyEight.com để biết chi tiết). Tương tự như vậy, một số phụ nữ là người có bốn sắc tố ( <https://concettaantico.com/tetrachromacy/> ), những người có thể nhìn thấy 100.000.000 màu thay vì 1.000.000 màu thông thường. (Chỉ phụ nữ mới có thể là người có bốn sắc tố do các yêu cầu về nhiễm sắc thể.) Thậm chí vẫn chưa thể biết có bao nhiêu phụ nữ có khả năng này, nhưng một số nguồn đưa ra con số cao tới 20 phần trăm; hãy xem http://sciencevibe.com/2016/12/11/the-women-that-see-100- million-colors-live-in-a-different-world/ để biết chi tiết. Việc sử dụng dữ liệu tĩnh và động đã lọc cho phép AI tương tác với con người theo những cách cụ thể ngày nay. Ví dụ, hãy xem xét Alexa, thiết bị của Amazon dường như có thể nghe thấy bạn và sau đó nói lại điều gì đó. Mặc dù Alexa thực sự không hiểu bất cứ điều gì bạn nói, nhưng vẻ ngoài của giao tiếp khá gây nghiện và khuyến khích mọi người nhân cách hóa các thiết bị này. Để thực hiện nhiệm vụ của mình, Alexa cần có quyền truy cập vào một cảm biến đặc biệt: một micrô cho phép nó nghe. Trên thực tế, Alexa có một số micrô giúp nó nghe đủ tốt để tạo ra ảo giác về sự hiểu biết. Thật không may, mặc dù Alexa tiên tiến đến mức nào, nó vẫn không thể nhìn thấy, cảm nhận, chạm vào hoặc nếm bất cứ thứ gì, điều này khiến nó không giống con người ngay cả ở những khía cạnh nhỏ nhất.

## Thiết kế các phương pháp tương tác với môi trường

Một AI tự chứa và không bao giờ tương tác với môi trường là vô dụng. Tất nhiên, tương tác đó diễn ra dưới dạng đầu vào và đầu ra. Phương pháp truyền thống để cung cấp đầu vào và đầu ra là trực tiếp thông qua các luồng dữ liệu mà máy tính có thể hiểu được, chẳng hạn như tập dữ liệu, truy vấn văn bản, v.v. Tuy nhiên, những cách tiếp cận này hầu như không thân thiện với con người và chúng đòi hỏi các kỹ năng đặc biệt để sử dụng.

Tương tác có thể có nhiều hình thức. Trên thực tế, số lượng và hình thức tương tác đang liên tục tăng lên. Ví dụ, AI hiện có thể ngửi (xem “Trí tuệ nhân tạo mọc thêm mũi” tại ScienceMag.org). Tuy nhiên, máy tính thực sự không ngửi thấy bất cứ thứ gì. Cảm biến cung cấp phương tiện để biến phát hiện hóa chất thành dữ liệu mà AI có thể sử dụng theo cùng cách mà nó làm với tất cả các dữ liệu khác. Khả năng phát hiện hóa chất không phải là mới; khả năng biến quá trình phân tích các hóa chất đó không phải là mới; cũng như các thuật toán được sử dụng để tương tác với dữ liệu kết quả không phải là mới. Cái mới là các tập dữ liệu được sử dụng để diễn giải dữ liệu đầu vào thành mùi và các tập dữ liệu đó đến từ các nghiên cứu trên người. “Mũi” của AI có đủ mọi cách sử dụng có thể. Ví dụ, hãy nghĩ về khả năng sử dụng mũi của AI khi làm việc trong một số môi trường nguy hiểm, chẳng hạn như ngửi thấy rò rỉ khí trước khi có thể nhìn thấy bằng cách sử dụng các cảm biến khác. Tương tác vật lý cũng đang gia tăng. Robot làm việc trong dây chuyền lắp ráp là chuyện cũ, nhưng hãy xem xét tác động của robot có thể lái xe. Đây là những cách sử dụng tương tác vật lý lớn hơn. Cũng cần cân nhắc rằng AI có thể phản ứng theo những cách nhỏ hơn. Ví dụ, Hugh Herr sử dụng AI để cung cấp tương tác với bàn chân thông minh, như được mô tả trong "Đây có phải là tương lai của chân rô-bốt không?" tại Smithsonian Magazine.com và "Phẫu thuật mới có thể giúp kiểm soát chân tay giả tốt hơn" tại MIT News.edu. Bàn chân động này cung cấp sự thay thế vượt trội cho những người đã mất bàn chân thật. Thay vì loại phản hồi tĩnh mà con người nhận được từ chân giả thông thường, bàn chân động này thực sự cung cấp loại phản hồi chủ động mà con người thường nhận được từ bàn chân thật. Ví dụ, lượng đẩy ngược từ bàn chân khác nhau khi đi lên dốc so với đi xuống dốc. Tương tự như vậy, việc điều hướng lề đường đòi hỏi lượng đẩy ngược khác so với việc điều hướng một bước. Vấn đề là khi AI có khả năng thực hiện các phép tính phức tạp hơn trong các gói nhỏ hơn với các tập dữ liệu ngày càng lớn hơn, khả năng thực hiện các nhiệm vụ thú vị của AI sẽ tăng lên. Tuy nhiên, các nhiệm vụ mà AI thực hiện hiện có thể không thuộc danh mục của con người. Bạn có thể không bao giờ thực sự tương tác với một AI có thể hiểu được lời nói của bạn, nhưng bạn có thể dựa vào một AI giúp bạn duy trì sự sống hoặc ít nhất là làm cho cuộc sống dễ sống hơn.

# Chương 5: Xem cách sử dụng AI trong các ứng dụng máy tính

**TRONG CHƯƠNG NÀY**

* **Định nghĩa và sử dụng AI trong các ứng dụng**
* **Sử dụng AI để chỉnh sửa và đề xuất**
* **Hiểu các lỗi tiềm ẩn của AI**

Bạn có thể đã sử dụng AI ở một số dạng trong nhiều ứng dụng máy tính mà bạn dựa vào cho công việc của mình. Ví dụ, khi nói chuyện với điện thoại thông minh, bạn cần sử dụng AI nhận dạng giọng nói. Tương tự như vậy, AI lọc ra tất cả thư rác có thể đến Hộp thư đến của bạn. Phần đầu tiên của chương này thảo luận về các loại ứng dụng AI, nhiều loại trong số đó sẽ khiến bạn ngạc nhiên, và các lĩnh vực thường dựa vào AI để thực hiện một số lượng lớn các tác vụ. Bạn cũng khám phá ra nguồn hạn chế khi tạo các ứng dụng dựa trên AI, điều này giúp bạn hiểu tại sao robot có tri giác có thể không bao giờ xuất hiện — hoặc ít nhất là không xuất hiện với công nghệ hiện có.

Tuy nhiên, bất kể AI có bao giờ đạt được tri giác hay không, thực tế vẫn là AI thực hiện một số lượng lớn các nhiệm vụ hữu ích. Hai cách thiết yếu mà AI hiện đang đóng góp vào nhu cầu của con người là thông qua các bản sửa lỗi và gợi ý. Bạn không muốn có quan điểm của con người về hai thuật ngữ này. Bản sửa lỗi không nhất thiết là phản hồi cho một lỗi. Tương tự như vậy, một gợi ý không nhất thiết là phản hồi cho một truy vấn. Ví dụ, hãy xem xét một chiếc ô tô hỗ trợ lái xe (một chiếc ô tô mà AI hỗ trợ thay vì thay thế người lái). Khi chiếc xe di chuyển, AI có thể thực hiện các bản sửa lỗi nhỏ cho phép điều kiện lái xe và đường sá, người đi bộ và vô số các vấn đề khác trước khi xảy ra lỗi thực sự. AI có cách tiếp cận chủ động đối với một vấn đề có thể xảy ra hoặc không. Tương tự như vậy, AI có thể gợi ý một con đường nhất định cho người lái xe có thể mang lại khả năng thành công cao nhất, chỉ để thay đổi gợi ý sau đó dựa trên các điều kiện mới. Phần thứ hai của chương này xem xét các bản sửa lỗi và gợi ý riêng biệt.

Phần chính thứ ba của chương thảo luận về các lỗi AI tiềm ẩn. Một lỗi xảy ra bất cứ khi nào kết quả khác với mong đợi. Kết quả có thể thành công, nhưng nó có thể vẫn không mong đợi. Tất nhiên, lỗi hoàn toàn cũng xảy ra: AI có thể không cung cấp kết quả thành công. Có lẽ kết quả thậm chí còn đi ngược lại mục tiêu ban đầu (có thể gây ra thiệt hại). Nếu bạn có ý tưởng rằng các ứng dụng AI cung cấp kết quả màu xám, thay vì đen hoặc trắng, thì bạn đang trên con đường hiểu cách AI sửa đổi các ứng dụng máy tính thông thường, trên thực tế, cung cấp kết quả hoàn toàn chính xác hoặc hoàn toàn không chính xác.

**Giới thiệu các loại ứng dụng phổ biến**

Cũng giống như điều duy nhất hạn chế các loại ứng dụng máy tính thủ tục là trí tưởng tượng của lập trình viên, các ứng dụng AI có thể xuất hiện ở bất kỳ địa điểm nào cho hầu hết mọi mục đích, phần lớn trong số đó chưa ai nghĩ đến. Trên thực tế, tính linh hoạt mà AI cung cấp có nghĩa là một số ứng dụng AI có thể xuất hiện ở những nơi khác ngoài những nơi mà lập trình viên ban đầu định nghĩa chúng. Trên thực tế, một ngày nào đó, phần mềm AI cũng có thể viết thế hệ tiếp theo của riêng mình (xem "Phần mềm AI học cách tạo phần mềm AI" tại MIT Technology Review.com để biết chi tiết). Công cụ GPT-3 ( <https://openai.com/blog/openai-api/> và <https://arxiv.org/abs/2005.14165> ) hiện nay viết một số lượng lớn các loại tài liệu, bao gồm cả mã (xem "AI này có thể mang đến cho chúng ta những chiếc máy tính có thể tự viết phần mềm của chúng" tại SingularityHub.com). Tuy nhiên, để có được ý tưởng tốt hơn về những gì làm cho AI hữu ích trong các ứng dụng, hãy xem các ứng dụng được áp dụng phổ biến nhất cho AI hiện nay (và những cạm bẫy tiềm ẩn liên quan đến những ứng dụng đó), như được mô tả trong các phần sau.

## Sử dụng AI trong các ứng dụng điển hình

Bạn có thể tìm thấy AI ở những nơi khó có thể tưởng tượng được việc sử dụng AI. Ví dụ, bộ điều nhiệt thông minh để kiểm soát nhiệt độ trong nhà của bạn có thể chứa AI nếu bộ điều nhiệt đủ phức tạp (xem “Bộ điều nhiệt thông minh tốt nhất năm 2020” tại NBC News.com để biết chi tiết). Việc sử dụng AI, ngay cả trong những ứng dụng đặc biệt này, thực sự có ý nghĩa khi AI được sử dụng cho những việc mà AI làm tốt nhất, chẳng hạn như theo dõi nhiệt độ ưa thích theo thời gian để tự động tạo lịch trình nhiệt độ. Sau đây là một số cách sử dụng AI phổ biến hơn mà bạn sẽ tìm thấy ở nhiều nơi:

* Sáng tạo nhân tạo
* Tầm nhìn máy tính, thực tế ảo và xử lý hình ảnh
* Chẩn đoán (trí tuệ nhân tạo)
* Nhận dạng khuôn mặt
* Trò chơi trí tuệ nhân tạo, bot trò chơi máy tính, lý thuyết trò chơi và lập kế hoạch chiến lược
* Nhận dạng chữ viết tay
* Xử lý ngôn ngữ tự nhiên, dịch thuật và chatterbot
* Điều khiển phi tuyến tính và robot
* Nhận dạng ký tự quang học
* Nhận dạng giọng nói

Tuy nhiên, khai thác AI có tiềm năng lớn nhất trong việc tạo ra các vấn đề nghiêm trọng là *deep fake* (AI mạo danh ai đó để nói hoặc làm những điều mà người thật sẽ không bao giờ làm). Bài viết “Năm Deepfakes trở nên phổ biến” trên MIT Technology Review.com mô tả công nghệ này khá chi tiết. Tuy nhiên, đọc về deep fake và chứng kiến ​​một deep fake hoạt động là hai chuyện khác nhau. Hãy xem deep fake của cựu Tổng thống Obama trên YouTube.com và bạn sẽ bắt đầu hiểu được mục đích thực sự xấu xa mà một số người có thể áp dụng AI. Tất nhiên, cách sử dụng AI mới này đang tạo ra các vấn đề nghiêm trọng cho hệ thống tòa án, như được mô tả trong “Tòa án và luật sư đấu tranh với sự phổ biến ngày càng tăng của deepfake” trên ABA Journal.com. Nếu bạn muốn xem deep fake đó thực sự hoạt động như thế nào, hãy xem video “Cách thức hoạt động của DEEPFAKE của Obama / Jordan Peele | Tin tức giả của Ian Hislop – BBC” trên YouTube.com.

***Nhận ra phạm vi rộng lớn của AI trong các lĩnh vực***

Các ứng dụng xác định các loại sử dụng cụ thể cho AI. Bạn cũng có thể tìm thấy AI được sử dụng chung hơn trong các lĩnh vực chuyên môn cụ thể. Danh sách sau đây chứa các lĩnh vực mà AI thường xuất hiện nhất:

* Cuộc sống nhân tạo
* Lý luận tự động
* Tự động hóa
* Máy tính lấy cảm hứng từ sinh học
* Khai thác khái niệm
* Khai thác dữ liệu
* Lọc thư rác email
* Hệ thống thông minh lai
* Tác nhân thông minh và điều khiển thông minh
* Biểu diễn kiến ​​thức
* kiện tụng
* Robot: robot dựa trên hành vi, nhận thức, điều khiển học, robot phát triển (biểu sinh) và robot tiến hóa
* Web ngữ nghĩa

***Xem xét lập luận về ‘Phòng Trung Quốc’***

Năm 1980, John Searle đã viết một bài báo có tựa đề “Minds, Brains, and Programs” được xuất bản trên tạp *chí Behavioral and Brain Sciences* . Trọng tâm của bài báo này là bác bỏ bài kiểm tra Turing, trong đó máy tính có thể đánh lừa con người nghĩ rằng máy tính là con người (chứ không phải máy tính) bằng cách sử dụng một loạt các câu hỏi (xem bài báo tại

<https://www.abelard.org/turpap/turpap.php> để biết chi tiết). Giả định cơ bản là chủ nghĩa chức năng, hay khả năng mô phỏng các đặc điểm cụ thể của tâm trí con người, không giống với suy nghĩ thực sự.

Lập luận ‘Phòng Trung Quốc’, như thí nghiệm tư duy này được gọi, dựa trên hai bài kiểm tra. Trong bài kiểm tra đầu tiên, ai đó tạo ra một AI có thể chấp nhận các ký tự tiếng Trung, sử dụng một bộ quy tắc để tạo phản hồi từ các ký tự đó, sau đó đưa ra phản hồi bằng các ký tự tiếng Trung. Câu hỏi là về một câu chuyện — AI phải diễn giải các câu hỏi được đặt ra sao cho câu trả lời phản ánh nội dung câu chuyện thực tế chứ không chỉ là một số phản hồi ngẫu nhiên. AI này giỏi đến mức không ai bên ngoài phòng có thể biết rằng AI đang thực hiện các nhiệm vụ được yêu cầu. Những người nói tiếng Trung hoàn toàn bị lừa khi nghĩ rằng AI thực sự có thể đọc và hiểu tiếng Trung.

Trong bài kiểm tra thứ hai, một người không nói tiếng Trung được đưa cho ba mục bắt chước những gì máy tính làm. Mục đầu tiên là một tập lệnh chứa một số lượng lớn các ký tự tiếng Trung; mục thứ hai là một câu chuyện bằng tiếng Trung; và mục thứ ba là một tập hợp các quy tắc để liên hệ mục đầu tiên với mục thứ hai. Một người nào đó gửi vào một tập hợp các câu hỏi, được viết bằng tiếng Trung, mà người đó hiểu được bằng cách sử dụng tập hợp các quy tắc để tìm vị trí trong câu chuyện chứa câu trả lời dựa trên cách diễn giải các ký tự tiếng Trung. Câu trả lời là tập hợp các ký tự tiếng Trung có liên quan đến câu hỏi dựa trên các quy tắc. Người đó trở nên giỏi trong nhiệm vụ này đến mức không ai có thể nhận ra sự thiếu hiểu biết về tiếng Trung.

Mục đích của hai bài kiểm tra này là để chứng minh rằng khả năng sử dụng các quy tắc chính thức để tạo ra kết quả (cú pháp) không giống với việc thực sự hiểu được ai đó đang làm gì (ngữ nghĩa). Searle đưa ra giả thuyết rằng cú pháp không đủ cho ngữ nghĩa, tuy nhiên đây lại là điều mà một số người triển khai AI đang cố gắng nói khi nói đến việc tạo ra nhiều công cụ dựa trên quy tắc khác nhau, chẳng hạn như Cơ chế áp dụng tập lệnh (SAM); xem “Sam– A Story Understander. Research Report No. 43” tại Eric.ed.gov để biết chi tiết.

Vấn đề cơ bản liên quan đến việc có một AI mạnh, một AI thực sự hiểu những gì nó đang cố gắng làm, và một AI yếu, một AI chỉ đơn giản là tuân theo các quy tắc. Tất cả AI ngày nay đều là AI yếu; nó thực sự không hiểu bất cứ điều gì. Những gì bạn thấy là chương trình thông minh mô phỏng suy nghĩ bằng cách sử dụng các quy tắc (chẳng hạn như các quy tắc ngầm định trong thuật toán). Tất nhiên, có nhiều tranh cãi nảy sinh về ý tưởng rằng cho dù máy móc có phức tạp đến đâu, chúng cũng sẽ không thực sự phát triển não bộ, nghĩa là chúng sẽ không bao giờ hiểu. Khẳng định của Searle là AI sẽ vẫn yếu. Bạn có thể xem thảo luận về chủ đề này tại <http://www.iep.utm.edu/chineser/> . Các lập luận và phản biện rất thú vị khi đọc vì chúng cung cấp những hiểu biết sâu sắc đáng kể về những gì thực sự phát huy tác dụng khi tạo ra AI.

## Xem cách AI làm cho các ứng dụng thân thiện hơn

Bạn có thể xem câu hỏi về tính thân thiện của ứng dụng được AI giải quyết theo nhiều cách khác nhau. Ở cấp độ cơ bản nhất, AI có thể dự đoán đầu vào của người dùng. Ví dụ, khi người dùng chỉ nhập một vài chữ cái của một từ cụ thể, AI sẽ đoán các ký tự còn lại. Bằng cách cung cấp dịch vụ này, AI đạt được một số mục tiêu:

 Người dùng sẽ làm việc hiệu quả hơn khi nhập ít ký tự hơn.

 Ứng dụng nhận được ít mục nhập sai hơn do lỗi đánh máy.

 Người dùng và ứng dụng đều tham gia vào mức độ giao tiếp cao hơn bằng cách nhắc nhở người dùng bằng các thuật ngữ chính xác hoặc nâng cao mà nếu không thì người dùng có thể không nhớ, tránh các thuật ngữ thay thế mà máy tính có thể không nhận ra.

AI cũng có thể học hỏi từ dữ liệu đầu vào trước đó của người dùng để sắp xếp lại các gợi ý theo cách phù hợp với phương pháp thực hiện nhiệm vụ của người dùng. Mức độ tương tác tiếp theo này nằm trong phạm vi gợi ý được mô tả trong phần " [Đưa ra gợi ý](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-05-seeing-ai-uses-in-computer-applications/#page143) " ở phần sau của chương này. Gợi ý cũng có thể bao gồm việc cung cấp cho người dùng những ý tưởng mà nếu không thì người dùng có thể chưa cân nhắc đến. Ngay cả trong lĩnh vực gợi ý, con người có thể bắt đầu nghĩ rằng AI đang suy nghĩ, nhưng thực tế không phải vậy. AI đang thực hiện một dạng so khớp mẫu nâng cao cũng như phân tích để xác định khả [năng cần đến một dữ liệu đầu vào cụ thể. Phần "Xem xét lập luận Phòng Trung Quốc" ở phần trước của chương này thảo luận về](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-05-seeing-ai-uses-in-computer-applications/#page137) sự khác biệt giữa AI yếu, loại AI được tìm thấy trong mọi ứng dụng hiện nay, và AI mạnh, thứ mà các ứng dụng cuối cùng có thể đạt được.

Sử dụng AI cũng có nghĩa là con người hiện có thể thực hiện các loại đầu vào thông minh khác. Ví dụ về giọng nói gần như đã bị lạm dụng, nhưng nó vẫn là một trong những phương pháp đầu vào thông minh phổ biến hơn. Tuy nhiên, ngay cả khi AI không có đầy đủ các giác quan, như đã mô tả trong [Chương 4](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-05-seeing-ai-uses-in-computer-applications/#page103) , nó vẫn có thể cung cấp nhiều loại đầu vào thông minh phi ngôn ngữ. Một lựa chọn rõ ràng là trực quan, chẳng hạn như nhận dạng khuôn mặt của chủ sở hữu hoặc mối đe dọa dựa trên biểu cảm khuôn mặt. Tuy nhiên, đầu vào có thể bao gồm màn hình, có thể kiểm tra các dấu hiệu sinh tồn của người dùng để tìm các vấn đề tiềm ẩn. Trên thực tế, AI có thể sử dụng một số lượng lớn các đầu vào thông minh, hầu hết trong số đó thậm chí còn chưa được phát minh.

Hiện tại, các ứng dụng thường chỉ xem xét ba cấp độ thân thiện đầu tiên này. Tuy nhiên, khi trí thông minh của AI tăng lên, thì việc AI thể hiện các hành vi Trí tuệ nhân tạo thân thiện (FAI) phù hợp với Trí tuệ nhân tạo tổng quát (AGI) có tác động tích cực đến nhân loại trở nên cần thiết. AI có các mục tiêu, nhưng các mục tiêu đó có thể không phù hợp với đạo đức của con người và khả năng không phù hợp gây ra sự lo lắng ngày nay. Một FAI sẽ bao gồm logic để đảm bảo rằng các mục tiêu của AI vẫn phù hợp với các mục tiêu của nhân loại, tương tự như ba định luật được tìm thấy trong các cuốn sách của Isaac Asimov (xem “Ba định luật về người máy của Isaac Asimov” tại webhome.auburn.edu), mà bạn sẽ thấy được thảo luận chi tiết hơn trong [Chương 12.](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-05-seeing-ai-uses-in-computer-applications/#page310) Tuy nhiên, nhiều người cho rằng ba định luật chỉ là điểm khởi đầu tốt và chúng ta cần các biện pháp bảo vệ hơn nữa (xem “Sau 75 năm, Ba định luật về người máy của Isaac Asimov cần được cập nhật” The Conversation.com.

Tất nhiên, tất cả những thảo luận này về luật pháp và đạo đức có thể khá khó hiểu và khó định nghĩa. Một ví dụ đơn giản về hành vi của FAI là FAI sẽ từ chối tiết lộ thông tin cá nhân của người dùng trừ khi người nhận có nhu cầu biết. Trên thực tế, FAI có thể tiến xa hơn nữa bằng cách khớp mẫu dữ liệu đầu vào của con người và định vị thông tin cá nhân tiềm ẩn trong đó, thông báo cho người dùng về khả năng gây hại trước khi gửi thông tin đi bất kỳ đâu. Vấn đề là AI có thể thay đổi đáng kể cách con người xem các ứng dụng và tương tác với chúng.

***Thực hiện sửa lỗi tự động***

Con người liên tục sửa chữa mọi thứ. Vấn đề không phải là mọi thứ đều sai. Thay vào đó, vấn đề là làm cho mọi thứ tốt hơn một chút (hoặc ít nhất là cố gắng làm cho chúng tốt hơn). Ngay cả khi con người cố gắng đạt được mức độ đúng đắn phù hợp tại một thời điểm cụ thể, một trải nghiệm mới sẽ đặt ra câu hỏi về mức độ đúng đắn đó vì giờ đây, con người có thêm dữ liệu để đánh giá toàn bộ câu hỏi về điều gì cấu thành đúng trong một tình huống cụ thể. Để bắt chước hoàn toàn trí thông minh của con người, AI cũng phải có khả năng liên tục sửa chữa các kết quả mà nó cung cấp, ngay cả khi các kết quả hiện tại sẽ mang lại kết quả tích cực. Các phần sau đây thảo luận về vấn đề chính xác và xem xét cách các bản sửa lỗi tự động đôi khi không thành công.

## Xem xét các loại sửa chữa của AI

Khi hầu hết mọi người nghĩ về AI và sửa lỗi, họ nghĩ về trình kiểm tra chính tả hoặc trình kiểm tra ngữ pháp. Một người mắc lỗi (hoặc ít nhất là AI nghĩ vậy) và AI sửa lỗi này để tài liệu được đánh máy chính xác nhất có thể. Tất nhiên, con người mắc rất nhiều lỗi, vì vậy có một AI để sửa lỗi là một ý tưởng hay.

Việc sửa lỗi có thể có nhiều hình thức khác nhau và không nhất thiết có nghĩa là lỗi đã xảy ra hoặc sẽ xảy ra trong tương lai. Ví dụ, một chiếc xe có thể hỗ trợ người lái xe bằng cách liên tục điều chỉnh vị trí làn đường. Người lái xe có thể vẫn trong giới hạn lái xe an toàn, nhưng AI có thể cung cấp những điều chỉnh nhỏ này để giúp đảm bảo người lái xe vẫn an toàn.

Tiến xa hơn nữa trong toàn bộ kịch bản hiệu chỉnh, hãy tưởng tượng rằng chiếc xe phía trước chiếc xe có AI dừng lại đột ngột vì một con nai trên đường. Người lái xe hiện tại không hề phạm phải bất kỳ lỗi nào.

Tuy nhiên, AI có thể phản ứng nhanh hơn người lái xe và hành động để dừng xe nhanh nhất và an toàn nhất có thể để xử lý chiếc xe đang dừng phía trước.

***Nhìn thấy lợi ích của việc sửa lỗi tự động***

Khi AI thấy cần phải sửa lỗi, nó có thể yêu cầu con người cho phép sửa lỗi hoặc tự động thay đổi. Ví dụ, khi ai đó sử dụng nhận dạng giọng nói để nhập tài liệu và mắc lỗi ngữ pháp, AI nên xin phép trước khi thực hiện thay đổi vì con người có thể thực sự muốn nói đến từ đó hoặc AI có thể hiểu sai ý của con người.

Tuy nhiên, đôi khi điều quan trọng là AI phải cung cấp một quy trình ra quyết định đủ mạnh mẽ để thực hiện các điều chỉnh tự động. Ví dụ, khi xem xét tình huống vị trí làn đường từ phần trước, AI không có thời gian để xin phép; nó phải phanh ngay lập tức nếu không con người có thể tử vong do va chạm. Các điều chỉnh tự động có một vị trí nhất định khi làm việc với AI, giả sử rằng nhu cầu đưa ra quyết định là rất quan trọng và AI là mạnh mẽ.

***Hiểu lý do tại sao các bản sửa lỗi tự động không hiệu quả***

Như đã đề cập trong phần “ [Xem xét lập luận về Phòng Trung Quốc](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-05-seeing-ai-uses-in-computer-applications/#page137) ”, ở đầu chương này, AI thực sự không thể hiểu bất cứ điều gì. Nếu không hiểu, AI không có khả năng bù đắp cho một hoàn cảnh không lường trước được. Trong trường hợp này, hoàn cảnh không lường trước được liên quan đến một sự kiện không được lập trình sẵn, một sự kiện mà AI không thể tích lũy thêm dữ liệu hoặc dựa vào các phương tiện cơ học khác để giải quyết. Con người có thể giải quyết vấn đề vì con người hiểu được cơ sở của vấn đề và thường là đủ các sự kiện xung quanh để xác định một mô hình có thể giúp hình thành giải pháp. Ngoài ra, sự đổi mới và sáng tạo của con người cung cấp các giải pháp mà không có giải pháp nào là hiển nhiên thông qua các phương tiện khác. Do AI hiện tại thiếu cả sự đổi mới và sáng tạo, nên AI ở thế bất lợi trong việc giải quyết các lĩnh vực vấn đề cụ thể.

Để đưa vấn đề này vào đúng bối cảnh, hãy xem xét trường hợp của một trình kiểm tra chính tả. Một con người gõ một từ hoàn toàn hợp lệ nhưng không xuất hiện trong từ điển mà AI sử dụng để sửa lỗi. AI thường thay thế một từ trông gần giống với từ đã chỉ định nhưng vẫn không chính xác. Ngay cả  
sau khi con người kiểm tra tài liệu, nhập lại từ đúng, rồi thêm vào từ điển, AI vẫn có thể mắc lỗi. Ví dụ, AI có thể xử lý từ viết tắt *CPU* khác với *cpu* vì từ trước viết hoa và từ sau viết thường. Con người sẽ thấy rằng hai từ viết tắt này giống nhau và trong trường hợp thứ hai, từ viết tắt là đúng nhưng có thể cần phải viết hoa.

## Xem cách AI đưa ra gợi ý

Gợi ý khác với lệnh. Mặc dù một số người có vẻ như hoàn toàn không hiểu vấn đề, nhưng gợi ý chỉ đơn giản là một ý tưởng được đưa ra như một giải pháp tiềm năng cho một vấn đề. Đưa ra gợi ý ngụ ý rằng có thể có những giải pháp khác và việc chấp nhận một gợi ý không có nghĩa là tự động thực hiện nó. Trên thực tế, gợi ý chỉ là một ý tưởng; nó thậm chí có thể không hiệu quả. Tất nhiên, trong một thế giới hoàn hảo, tất cả các gợi ý đều là những gợi ý tốt — ít nhất là các giải pháp khả thi cho một đầu ra chính xác, điều này hiếm khi xảy ra trong thế giới thực. Các phần sau đây mô tả bản chất của các gợi ý khi chúng áp dụng cho AI.

***Nhận gợi ý dựa trên các hành động trong quá khứ***

Cách phổ biến nhất mà AI sử dụng để tạo ra gợi ý là thu thập các hành động trong quá khứ dưới dạng sự kiện rồi sử dụng các hành động trong quá khứ đó làm tập dữ liệu để đưa ra gợi ý mới. Ví dụ, ai đó mua một Half-Baked Widget mỗi tháng trong ba tháng. Sẽ hợp lý nếu gợi ý mua thêm một cái nữa vào đầu tháng thứ tư. Trên thực tế, một AI thực sự thông minh có thể đưa ra gợi ý vào đúng thời điểm trong tháng. Ví dụ, nếu người dùng mua hàng trong khoảng thời gian từ ngày thứ ba đến ngày thứ năm của tháng trong ba tháng đầu tiên, thì nên bắt đầu đưa ra gợi ý vào ngày thứ ba của tháng rồi chuyển sang thứ khác sau ngày thứ năm.

Con người đưa ra một số lượng lớn các manh mối trong khi thực hiện nhiệm vụ. Không giống như con người, AI thực sự chú ý đến từng manh mối này và có thể ghi lại chúng theo cách nhất quán để tạo *dữ liệu hành động* . Dữ liệu hành động thay đổi tùy theo nhiệm vụ được thực hiện; nó có thể bao gồm những thứ như tương tác với thiết bị, trình tự để đưa ra lựa chọn, tư thế cơ thể, biểu cảm khuôn mặt, cách thể hiện (như thái độ), v.v. Bằng cách thu thập dữ liệu hành động một cách nhất quán, AI có thể đưa ra các gợi ý dựa trên các hành động trong quá khứ với độ chính xác cao trong nhiều trường hợp.

***Nhận gợi ý dựa trên nhóm***

Một cách phổ biến khác để đưa ra gợi ý là dựa vào tư cách thành viên nhóm.

Trong trường hợp này, tư cách thành viên nhóm không nhất thiết phải chính thức. Một nhóm có thể bao gồm một hiệp hội lỏng lẻo của những người có chung một số nhu cầu hoặc hoạt động nhỏ. Ví dụ, một người đốn gỗ, một chủ cửa hàng và một chuyên gia dinh dưỡng đều có thể mua sách bí ẩn. Mặc dù họ không có điểm chung nào khác, thậm chí không có địa điểm, nhưng thực tế là cả ba đều thích bí ẩn khiến họ trở thành một phần của nhóm. Một AI có thể dễ dàng phát hiện ra các mô hình như thế này mà con người có thể không nhận ra, vì vậy nó có thể đưa ra các gợi ý mua hàng tốt dựa trên các mối quan hệ nhóm khá lỏng lẻo này.

Các nhóm có thể bao gồm các kết nối siêu nhiên chỉ mang tính tạm thời. Ví dụ, tất cả những người đã bay chuyến bay 1982 ra khỏi Houston vào một ngày nhất định có thể tạo thành một nhóm. Một lần nữa, không có bất kỳ kết nối nào tồn tại giữa những người này ngoại trừ việc họ xuất hiện trên một chuyến bay cụ thể. Tuy nhiên, bằng cách biết thông tin này, AI có thể thực hiện thêm quá trình lọc để xác định những người trong chuyến bay thích điều bí ẩn. Vấn đề là AI có thể đưa ra những gợi ý tốt dựa trên sự liên kết nhóm ngay cả khi nhóm đó khó (nếu không muốn nói là không thể) xác định theo quan điểm của con người.

***Nhận được những gợi ý sai***

Bất kỳ ai đã dành thời gian mua sắm trực tuyến đều biết rằng các trang web thường đưa ra các gợi ý dựa trên nhiều tiêu chí khác nhau, chẳng hạn như các lần mua trước hoặc thậm chí là các tìm kiếm. Thật không may, những gợi ý này thường sai vì AI cơ bản thiếu sự hiểu biết. Ví dụ, nếu bạn là tác giả của một cuốn sách và bạn xem số liệu thống kê về cuốn sách của mình trên Amazon, AI của Amazon sẽ liên tục đề xuất bạn mua các bản sao của cuốn sách đó bất kể bạn muốn làm gì về nó. Một ví dụ khác, khi ai đó mua một Tiện ích siêu rộng một lần trong đời, thì con người có thể biết rằng giao dịch mua đó thực sự chỉ diễn ra một lần trong đời vì rất khó có khả năng ai đó sẽ cần hai lần. Tuy nhiên, AI không hiểu được sự thật này. Vì vậy, trừ khi một lập trình viên cụ thể tạo ra một quy tắc chỉ định rằng Tiện ích siêu rộng là giao dịch mua một lần trong đời, AI có thể chọn tiếp tục đề xuất sản phẩm vì doanh số bán hàng thấp một cách dễ hiểu. Khi tuân theo một quy tắc phụ về việc quảng bá các sản phẩm có doanh số bán chậm hơn, AI hoạt động theo các đặc điểm mà nhà phát triển cung cấp cho nó, nhưng các gợi ý mà nó đưa ra hoàn toàn sai.

Bên cạnh các lỗi logic hoặc dựa trên quy tắc trong AI, các đề xuất có thể bị hỏng do các vấn đề về dữ liệu. Ví dụ, GPS có thể đưa ra đề xuất dựa trên dữ liệu tốt nhất có thể cho một chuyến đi cụ thể. Tuy nhiên, việc xây dựng đường có thể khiến lộ trình được đề xuất trở nên không khả thi vì đường đã đóng. Tất nhiên, nhiều ứng dụng GPS có xem xét đến việc xây dựng đường, nhưng đôi khi chúng không xem xét các vấn đề khác, chẳng hạn như thay đổi đột ngột về giới hạn tốc độ hoặc điều kiện thời tiết khiến một lộ trình cụ thể trở nên nguy hiểm. Con người có thể khắc phục tình trạng thiếu dữ liệu thông qua sự đổi mới, chẳng hạn như bằng cách sử dụng con đường ít người qua lại hoặc hiểu ý nghĩa của các biển báo đường vòng. Khi AI vượt qua được các vấn đề về logic, quy tắc và dữ liệu, đôi khi nó vẫn đưa ra các đề xuất không tốt vì nó không hiểu mối tương quan giữa các tập dữ liệu nhất định theo cùng cách mà con người hiểu. Ví dụ, AI có thể không biết đề xuất sơn sau khi con người mua kết hợp ống nước và vách thạch cao khi sửa chữa hệ thống ống nước. Nhu cầu sơn vách thạch cao và khu vực xung quanh sau khi sửa chữa là điều hiển nhiên đối với con người vì con người có khiếu thẩm mỹ mà AI không có. Con người tạo ra mối tương quan giữa nhiều sản phẩm khác nhau mà AI không thể nhận ra.

## Xem xét các lỗi dựa trên AI

Một lỗi hoàn toàn xảy ra khi kết quả của một quy trình, với các đầu vào cụ thể, không chính xác ở bất kỳ dạng nào. Câu trả lời không cung cấp phản hồi phù hợp cho truy vấn. Không khó để tìm ví dụ về lỗi dựa trên AI. Ví dụ: "AI image awareness fooled by single pixel change" tại BBC News.com mô tả cách một điểm khác biệt pixel duy nhất trong một hình ảnh đánh lừa một AI cụ thể. Lỗi cũng có thể dẫn đến các cuộc tấn công, chẳng hạn như các cuộc tấn công được mô tả trong thanh bên " [AI Exploits](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-05-seeing-ai-uses-in-computer-applications/#page136) ", ở đầu chương. Bạn có thể đọc thêm về tác động của các cuộc tấn công đối nghịch lên AI trong "Image-scaling attacks highlight danger of adversarial machine learning" tại TechTalks.com và trong "One Pixel Attack for Fooling Deep Neural Networks" tại ResearchGate.net. Vấn đề là AI vẫn có tỷ lệ lỗi cao trong một số trường hợp và các nhà phát triển làm việc với AI thường không chắc chắn tại sao lỗi lại xảy ra.

**Có nhiều nguồn gây ra lỗi trong AI.** Tuy nhiên, như đã lưu ý trong [Chương 1](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-05-seeing-ai-uses-in-computer-applications/#page24) , AI thậm chí không thể mô phỏng cả bảy dạng trí thông minh của con người, vì vậy sai lầm không chỉ có thể xảy ra mà còn không thể tránh khỏi. Phần lớn tài liệu trong [Chương 2](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-05-seeing-ai-uses-in-computer-applications/#page45) tập trung vào dữ liệu và tác động của nó lên AI khi dữ liệu bị lỗi theo một cách nào đó. Trong [Chương 3](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-05-seeing-ai-uses-in-computer-applications/#page80) , bạn cũng thấy rằng ngay cả các thuật toán mà AI sử dụng cũng có giới hạn. [Chương 4](https://bo0kbuzz.wordpress.com/2024/12/11/artificial-intelligence-for-dummies-chapter-05-seeing-ai-uses-in-computer-applications/#page103) chỉ ra rằng AI không có quyền truy cập vào cùng số lượng hoặc loại giác quan của con người. Như bài viết "Trí tuệ nhân tạo không thông minh như bạn (hoặc Elon Musk) nghĩ" trên TechCrunch.com đã lưu ý, nhiều nhiệm vụ có vẻ như không thể mà AI thực hiện ngày nay là kết quả của việc sử dụng các phương pháp ‘sức mạnh cơ bắp’ hơn là bất kỳ thứ gì gần với tư duy thực tế.

Một vấn đề lớn đang ngày càng trở nên rõ ràng hơn là các tập đoàn thường bỏ qua hoặc thậm chí phớt lờ các vấn đề liên quan đến AI. Trọng tâm là sử dụng AI để giảm chi phí và cải thiện năng suất, điều này có thể không đạt được. Bài viết “The Limits of Artificial Intelligence” của Bloomberg.com thảo luận về vấn đề này khá chi tiết. Một trong những ví dụ thú vị hơn nhưng cũng đáng lo ngại hơn về một thực thể công ty đi quá xa với AI là Tay của Microsoft, được đào tạo một cách ác ý để đưa ra những nhận xét phân biệt chủng tộc, phân biệt giới tính và khiêu dâm trước đám đông lớn trong một bài thuyết trình (xem “Trình trò chuyện trực tuyến của Microsoft, Tay, lọt vào danh sách những thất bại công nghệ lớn nhất hàng năm của MIT” tại GeekWire.com).

# Chương 6: Tự động hóa các quy trình phổ biến

**TRONG CHƯƠNG NÀY**

* **Sử dụng AI để đáp ứng nhu cầu của con người**
* **Làm cho ngành công nghiệp hiệu quả hơn**
* **Phát triển các giao thức an toàn động bằng cách sử dụng AI**

[Chương 5](https://bo0kbuzz.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-06-automating-common-processes/#page131) xem xét việc sử dụng AI trong một *ứng dụng,* đó là tình huống mà con người tương tác với AI theo một cách có ý nghĩa nào đó, ngay cả khi con người không biết đến sự hiện diện của AI. Mục tiêu là giúp con người làm một việc gì đó nhanh hơn, dễ dàng hơn hoặc hiệu quả hơn, hoặc đáp ứng một số nhu cầu khác. Một quy *trình* bao gồm AI thì khác vì AI hiện đang hoạt động để hỗ trợ con người hoặc thực hiện một số nhiệm vụ khác mà không cần can thiệp trực tiếp. Phần đầu tiên của chương này đề cập đến cách các quy trình giúp ích cho con người. Vì sự nhàm chán có thể là kịch bản tồi tệ nhất của con người (hãy nghĩ đến tất cả những điều tiêu cực xảy ra khi con người buồn chán), nên chương này xem xét quy trình AI dành cho con người theo góc nhìn của sự nhàm chán.

Một trong những cách AI được sử dụng lâu nhất là sử dụng trong công nghiệp, chẳng hạn như quy trình sản xuất, để cuối cùng cho phép triển khai Công nghiệp 4.0 (xem "Công nghiệp 4.0 là gì?" tại TWI Global.com để biết chi tiết). Hãy xem xét tất cả các rô-bốt hiện đang cung cấp năng lượng cho các nhà máy trên khắp thế giới. Mặc dù tự động hóa do AI cung cấp thay thế con người, nhưng nó cũng giúp con người an toàn hơn bằng cách thực hiện các nhiệm vụ thường được coi là nguy hiểm. Thật kỳ lạ, một trong những nguyên nhân quan trọng nhất gây ra tai nạn công nghiệp và vô số vấn đề khác là sự nhàm chán, như đã giải thích trong "Sự nhàm chán tại nơi làm việc" tại The Psychologist.bps.org.uk. Bài viết "Cách biến sự nhàm chán thành công việc của bạn" tại Fast Company.com cố gắng thay đổi mọi thứ, nhưng sự nhàm chán vẫn có thể và đang nguy hiểm. Rô-bốt có thể thực hiện những công việc lặp đi lặp lại đó một cách nhất quán và không cảm thấy nhàm chán (mặc dù đôi khi bạn có thể thấy ngáp).

Trong trường hợp bạn vẫn chưa đủ chán, bạn cũng có thể đọc một số thông tin về nó trong phần thứ ba của chương, thảo luận về

một số lĩnh vực mới nhất mà AI vượt trội —làm cho mọi môi trường trở nên an toàn hơn. Trên thực tế, chỉ trong ngành công nghiệp ô tô, bạn có thể tìm thấy vô số cách mà việc sử dụng AI đang làm cho mọi thứ trở nên tốt hơn (xem “Trí tuệ nhân tạo định hình lại ngành công nghiệp ô tô” tại Future Bridge.com để biết chi tiết).

\*Mục đích của chương này là AI hoạt động tốt trong các quy trình, đặc biệt là những quy trình mà con người có xu hướng cảm thấy nhàm chán, khiến họ mắc lỗi khi AI có khả năng sẽ không mắc lỗi. Tất nhiên, AI không thể loại bỏ mọi nguồn gây mất hiệu quả, mất hứng thú và vấn đề an toàn. Trước hết, con người có thể chọn cách bỏ qua sự trợ giúp của AI, nhưng bản chất của những hạn chế này còn sâu xa hơn thế nữa. Như đã thảo luận trong các chương trước (đặc biệt là [Chương 5](https://bo0kbuzz.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-06-automating-common-processes/#page131) ), AI không hiểu; nó không thể cung cấp các giải pháp sáng tạo hoặc đổi mới cho các vấn đề, vì vậy một số vấn đề không thể giải quyết được bằng AI, bất kể ai đó nỗ lực tạo ra nó như thế nào.

## 1. Phát triển các giải pháp cho sự nhàm chán

Các cuộc thăm dò thường cho thấy mọi người *nghĩ* họ muốn gì, thay vì những gì họ thực sự muốn, nhưng chúng vẫn hữu ích. Khi những sinh viên mới tốt nghiệp được thăm dò để xem họ muốn cuộc sống như thế nào, không một ai trong số họ nói rằng, "Ôi làm ơn, hãy để tôi buồn chán!" (hãy xem "Bạn muốn sống cuộc sống như thế nào?" tại Huffington Post.com). Trên thực tế, bạn có thể thăm dò bất kỳ nhóm nào và không đưa ra một câu trả lời buồn chán nào. Hầu hết con người (nói "tất cả" có thể dẫn đến một loạt email, với các ví dụ) không muốn buồn chán. Trong một số trường hợp, AI có thể hợp tác với con người để làm cho cuộc sống trở nên thú vị hơn — ít nhất là đối với con người. Các phần sau đây thảo luận về các giải pháp cho sự buồn chán của con người mà AI có thể cung cấp (và một số giải pháp mà AI không thể).

***Làm cho nhiệm vụ thú vị hơn***

Bất kỳ nghề nghiệp nào, dù là nghề cá nhân hay nghề của một tổ chức, đều có những đặc điểm nhất định thu hút mọi người và khiến họ muốn tham gia vào nghề đó. Rõ ràng, một số nghề, chẳng hạn như chăm sóc con cái của bạn, không được trả lương, nhưng sự hài lòng khi làm như vậy có thể cực kỳ cao. Tương tự như vậy, làm nhân viên kế toán có thể được trả lương khá cao nhưng không mang lại nhiều sự hài lòng trong công việc. Nhiều cuộc thăm dò ý kiến ​​(chẳng hạn như cuộc thăm dò này trong “Báo cáo xếp hạng việc làm năm 2019” tại [CareerCast.com](https://www.careercast.com/jobs-rated/2019-jobs-rated-report) ” và các bài viết như “Chìa khóa của hạnh phúc là gì: Mức lương cao hay sự hài lòng trong công việc?” tại Engineering and Technology Jobs.org nói về sự cân bằng giữa tiền bạc và sự hài lòng, nhưng việc đọc chúng thường gây nhầm lẫn vì cơ sở để đưa ra quyết định là mơ hồ. Tuy nhiên, hầu hết các nguồn này đều đồng ý rằng sau khi một người kiếm được một số tiền nhất định, sự hài lòng sẽ trở thành chìa khóa để duy trì sự quan tâm đến nghề nghiệp (bất kể nghề nghiệp đó là gì). Tất nhiên, việc tìm ra những gì tạo nên sự hài lòng trong công việc là gần như không thể, nhưng sự quan tâm vẫn nằm ở vị trí cao trong danh sách. Một nghề nghiệp thú vị sẽ luôn có tiềm năng hài lòng cao hơn.

\*Vấn đề không phải là nhất thiết phải thay đổi công việc, mà là làm cho công việc trở nên thú vị hơn như một phương tiện để tránh sự nhàm chán. AI có thể hỗ trợ hiệu quả quá trình này bằng cách loại bỏ sự lặp lại khỏi các nhiệm vụ.

Tuy nhiên, các ví dụ như Alexa của Amazon và Home của Google lại cung cấp các giải pháp thay thế khác. Cảm giác cô đơn có thể lan tỏa khắp nhà, nơi làm việc, ô tô và các địa điểm khác là tác nhân chính gây ra sự buồn chán. Khi con người bắt đầu cảm thấy cô đơn, chứng trầm cảm sẽ xuất hiện và sự buồn chán thường chỉ cách một bước chân. Việc tạo các ứng dụng sử dụng giao diện Alexa (xem <https://developer.amazon.com/> ) hoặc Hành động trên Google API (xem <https://developers.google.com/actions/> ) để mô phỏng tương tác của con người theo cách phù hợp có thể cải thiện trải nghiệm tại nơi làm việc. Quan trọng hơn, việc phát triển các giao diện thông minh theo cách này có thể giúp con người thực hiện nhanh chóng nhiều nhiệm vụ tầm thường, chẳng hạn như nghiên cứu thông tin và tương tác với các thiết bị thông minh, không chỉ công tắc đèn (xem “Cách điều khiển đèn bằng Amazon Echo” tại iMore.com và <https://store.google.com/product/google_home> để biết chi tiết).

***Giúp con người làm việc hiệu quả hơn***

Hầu hết mọi người, ít nhất là những người có tư duy tiến bộ, đều có một số ý tưởng về cách họ muốn AI làm cho cuộc sống của họ tốt hơn bằng cách loại bỏ các nhiệm vụ mà họ không muốn làm. Cuộc thăm dò trong "Những nhiệm vụ nào trong công việc của bạn mà bạn muốn được AI tự động hóa?" tại blog.devolutions.net cho thấy một số cách thú vị hơn mà họ mong muốn AI có thể cải thiện cuộc sống của họ:. Nhiều trong số chúng là tầm thường, nhưng hãy chú ý đến những cách như phát hiện khi người quan trọng khác không vui và gửi hoa. Có lẽ nó sẽ không hiệu quả, nhưng dù sao thì đó cũng là một ý tưởng thú vị.

Vấn đề là con người có thể sẽ đưa ra những ý tưởng thú vị nhất về cách tạo ra một AI đáp ứng cụ thể nhu cầu của họ. Trong hầu hết các trường hợp, những ý tưởng nghiêm túc cũng sẽ hiệu quả với những người dùng khác. Ví dụ, tự động hóa phiếu yêu cầu hỗ trợ là điều có thể hiệu quả trong một số trường hợp khác nhau

ngành công nghiệp. Nếu ai đó đưa ra một giao diện chung, với phần phụ trợ có thể lập trình để tạo ra các phiếu yêu cầu sửa chữa tùy chỉnh cần thiết, AI có thể giúp người dùng tiết kiệm rất nhiều thời gian và đảm bảo hiệu quả trong tương lai bằng cách đảm bảo rằng các phiếu yêu cầu sửa chữa luôn ghi lại thông tin cần thiết.

|  |
| --- |
| **TRÍ TUỆ PHẢN ĐỐI TRONG CÔNG VIỆC** Ít người thích mọi thứ trở nên khó khăn; hầu hết chúng ta đều muốn bắt tay vào công việc một cách dễ dàng và có được cảm giác hài lòng mỗi ngày. Tuy nhiên, một số bài viết và sách trắng mới dường như chỉ ra rằng việc đưa AI vào nơi làm việc thực sự khiến mọi thứ trở nên khó khăn hơn. Hãy xem xét bài viết này từ *The Atlantic, “* AI đang đến với những công việc tầm thường yêu thích của bạn”). Tuy nhiên, bài viết thực sự không nói về những công việc tầm thường. Nó nói nhiều hơn về việc AI hút hết niềm vui trong công việc của một người và chỉ để lại những yếu tố căng thẳng nhất mà chỉ con người mới có thể giải quyết hiệu quả. Bài viết xem xét mặt trái của vấn đề: những trường hợp khi tự động hóa khiến công việc của một người trở nên khó khăn hơn đáng kể và chắc chắn là kém thỏa mãn hơn, và con người thậm chí còn không được trả nhiều tiền hơn để làm việc đó. Quan trọng hơn, cơ hội đưa ra quyết định đúng đắn của con người vì tất cả các quyết định đều khó khăn cũng giảm xuống, điều này sau đó có thể khiến ban quản lý có ấn tượng rằng một nhân viên đột nhiên mất hứng thú hoặc đơn giản là không tập trung. Đến một lúc nào đó, sẽ phải cân bằng giữa những gì AI làm và những gì con người làm để duy trì sự hài lòng trong công việc. Thiết kế AI hiện tại không hề xem xét đến khía cạnh này của nhu cầu con người, nhưng nó sẽ là một yêu cầu trong tương lai. |

## 2. Hiểu cách AI làm giảm sự nhàm chán

Sự buồn chán có nhiều gói, và con người xem những gói này theo nhiều cách khác nhau. Có sự buồn chán đến từ việc không có đủ nguồn lực, kiến ​​thức hoặc các nhu cầu khác được đáp ứng. Một loại buồn chán khác đến từ việc không biết phải làm gì tiếp theo khi các hoạt động không tuân theo một khuôn mẫu cụ thể. AI có thể giúp giải quyết loại buồn chán đầu tiên; nhưng không thể giải quyết loại buồn chán thứ hai. Phần này sẽ xem xét loại buồn chán đầu tiên. (Phần tiếp theo sẽ xem xét loại buồn chán thứ hai.)

\*Việc tiếp cận các nguồn tài nguyên đủ loại giúp giảm sự nhàm chán bằng cách cho phép con người sáng tạo mà không cần phải có nhu cầu vật liệu cần thiết. Sau đây là một số cách mà AI có thể giúp tiếp cận các nguồn tài nguyên dễ dàng hơn:

 Tìm kiếm các mặt hàng cần thiết trực tuyến

 Tự động đặt hàng các mặt hàng cần thiết

 Thực hiện giám sát cảm biến và thu thập dữ liệu khác

 Quản lý dữ liệu

 Hoàn thành các nhiệm vụ tầm thường hoặc lặp đi lặp lại

***Xem xét cách AI không thể làm giảm sự nhàm chán***

Như đã lưu ý trong các chương trước, đặc biệt là [Chương 4](https://bo0kbuzz.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-06-automating-common-processes/#page103) và Chương [5](https://bo0kbuzz.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-06-automating-common-processes/#page131) , AI không sáng tạo hoặc trực quan. Vì vậy, yêu cầu AI nghĩ ra điều gì đó để bạn làm khó có thể tạo ra kết quả thỏa mãn. Ai đó có thể lập trình AI để theo dõi mười điều hàng đầu mà bạn thích làm và sau đó chọn ngẫu nhiên một trong số chúng, nhưng kết quả vẫn sẽ không thỏa mãn vì AI không thể tính đến các khía cạnh như trạng thái tinh thần hiện tại của bạn. Trên thực tế, ngay cả với phần mềm nhận dạng biểu cảm khuôn mặt tốt nhất, AI sẽ không có khả năng tương tác với bạn theo cách tạo ra bất kỳ kết quả thỏa mãn nào.

AI cũng không thể thúc đẩy bạn. Hãy nghĩ về những gì xảy ra khi một người bạn đến giúp thúc đẩy bạn (hoặc bạn thúc đẩy người bạn đó). Người bạn đó thực sự dựa vào sự kết hợp giữa kiến ​​thức nội tâm (thấu cảm bằng cách cân nhắc cảm giác khi ở trong hoàn cảnh của bạn) và kiến ​​thức giữa các cá nhân (chiếu những ý tưởng sáng tạo về cách để có được phản ứng cảm xúc tích cực từ bạn). AI sẽ không có bất kỳ kiến ​​thức nào thuộc loại đầu tiên và chỉ có một lượng cực kỳ hạn chế kiến ​​thức thuộc loại thứ hai, như đã mô tả trong [Chương 1.](https://bo0kbuzz.wordpress.com/2024/12/12/artificial-intelligence-for-dummies-chapter-06-automating-common-processes/#page24) Do đó, AI không thể giảm sự nhàm chán của bạn thông qua các kỹ thuật thúc đẩy.

\*Dù sao thì buồn chán không hẳn lúc nào cũng là điều xấu. Một số nghiên cứu gần đây đã chỉ ra rằng buồn chán thực sự giúp thúc đẩy tư duy sáng tạo, đó là hướng mà con người cần hướng tới (xem “Buồn chán có thể tốt cho bạn—Nếu bạn làm đúng” tại Time.com và “Khoa học đằng sau lợi ích của việc buồn chán đối với tư duy sáng tạo” tại Fast Company.com làm ví dụ). Bất chấp vô số bài viết về cách AI sẽ lấy đi công việc, điều quan trọng là phải xem xét rằng bản thân những công việc mà AI đang lấy đi thường rất nhàm chán và không để lại cho con người thời gian để sáng tạo. Ngay cả ngày nay, con người vẫn có thể tìm thấy những công việc sáng tạo, hiệu quả để làm nếu họ thực sự nghĩ về điều đó. Bài viết “7 sự thật đáng ngạc nhiên về sự sáng tạo, theo

Science” tại Fast Company.com thảo luận về vai trò của việc mơ mộng khi buồn chán trong việc tăng cường khả năng sáng tạo. Trong tương lai, nếu con người thực sự muốn vươn tới các vì sao và làm những điều tuyệt vời khác, thì khả năng sáng tạo sẽ là điều cần thiết, vì vậy thực tế là AI không thể làm giảm sự buồn chán của bạn thực sự là một điều tốt.

## 3. Làm việc trong môi trường công nghiệp

Bất kỳ bối cảnh công nghiệp nào cũng có thể có nguy cơ về an toàn, bất kể có bao nhiêu thời gian, công sức và tiền bạc được đầu tư vào vấn đề này. Bạn có thể dễ dàng tìm thấy các bài viết như bài viết này, “Hướng dẫn về các mối nguy hiểm phổ biến nhất tại nơi làm việc” tại High Speed [​​Training.co.uk](https://www.highspeedtraining.co.uk/hub/hazards-in-the-workplace/) , trong đó mô tả các mối nguy hiểm phổ biến về an toàn trong bối cảnh công nghiệp. Mặc dù con người gây ra nhiều vấn đề này và sự nhàm chán khiến chúng trở nên tồi tệ hơn, nhưng môi trường thực tế mà con người đang làm việc lại gây ra rất nhiều vấn đề. Các phần sau đây mô tả cách tự động hóa có thể giúp con người sống lâu hơn và tốt hơn.

***Phát triển nhiều cấp độ tự động hóa khác nhau***

Tự động hóa trong các thiết lập công nghiệp đã có từ lâu đời hơn bạn nghĩ. Nhiều người nghĩ rằng dây chuyền lắp ráp của Henry Ford là điểm khởi đầu của tự động hóa (xem “Dây chuyền lắp ráp của Ford bắt đầu lăn bánh” tại History.com). Trên thực tế, những điều cơ bản của tự động hóa bắt đầu vào năm 1104 sau Công nguyên tại Venice (xem “Xu hướng tự động hóa nhà máy thế kỷ 21” tại Mouser.com), nơi 16.000 công nhân có thể đóng toàn bộ một tàu chiến chỉ trong một ngày. Người Mỹ đã lặp lại kỳ tích đóng tàu chiến cực kỳ nhanh chóng với các tàu hiện đại trong Thế chiến II (đọc về nó trong “Đóng tàu trong Thế chiến II tại Khu vực Vịnh San Francisco” tại nps.gov) bằng cách dựa nhiều vào tự động hóa. Trên thực tế, cho đến nay đã có bốn cuộc cách mạng công nghiệp theo Viện Phát triển Doanh nhân (“Bốn cuộc cách mạng công nghiệp”). Vì vậy, tự động hóa đã tồn tại trong một thời gian dài.

Điều đã không xuất hiện trong một thời gian dài là một AI thực sự có thể giúp con người trong quá trình tự động hóa. Trong nhiều trường hợp ngày nay, một người vận hành bắt đầu bằng cách phác thảo cách thực hiện nhiệm vụ, tạo *công việc* và sau đó chuyển giao công việc cho máy tính. Một ví dụ về một trong số nhiều loại công việc khá mới là Tự động hóa quy trình bằng robot (RPA), cho phép con người đào tạo phần mềm để hành động thay cho con người khi làm việc với các ứng dụng (xem “Các công cụ của tương lai ngày nay” tại Valamis.com). Nhiều công ty hiện đang cung cấp dịch vụ RPA, chẳng hạn như UiPath

( <https://www.uipath.com/rpa/robotic-process-automation> ). Quá trình này khác với việc viết kịch bản, chẳng hạn như sử dụng Visual Basic for Applications (VBA) trong Microsoft Office, ở chỗ RPA không dành riêng cho ứng dụng và không yêu cầu mã hóa. Nhiều người thấy ngạc nhiên khi thực sự có mười cấp độ tự động hóa, chín trong số đó có thể dựa vào AI. Cấp độ bạn chọn phụ thuộc vào ứng dụng của bạn:

1. Người vận hành tạo ra một công việc và chuyển giao cho máy tính thực hiện.

* AI giúp con người xác định các lựa chọn công việc.
* AI xác định các lựa chọn công việc tốt nhất và sau đó cho phép con người chấp nhận hoặc từ chối đề xuất.
* AI xác định các tùy chọn, sử dụng chúng để xác định một loạt các hành động, sau đó chuyển danh sách các hành động cho con người để chấp nhận hoặc từ chối từng hành động trước khi thực hiện.
* AI xác định các tùy chọn, định nghĩa một loạt hành động, tạo công việc và sau đó yêu cầu con người chấp thuận trước khi gửi công việc đó cho máy tính.
* AI sẽ tự động tạo công việc và gửi đến hàng đợi công việc của máy tính, trong đó người vận hành đóng vai trò trung gian trong trường hợp công việc đã chọn cần chấm dứt trước khi triển khai thực tế.
* AI tạo ra và thực hiện công việc, sau đó thông báo cho người vận hành biết công việc đã làm trong trường hợp công việc cần được sửa chữa hoặc đảo ngược.
* AI tạo ra và thực hiện công việc, chỉ cho con người biết nó đã làm gì khi con người yêu cầu.
* AI tạo ra và thực hiện công việc mà không cung cấp bất kỳ phản hồi nào trừ khi con người cần can thiệp, chẳng hạn như khi xảy ra lỗi hoặc kết quả không như mong đợi.

1. AI khởi tạo nhu cầu cho công việc, thay vì chờ con người ra lệnh để tạo công việc. AI chỉ cung cấp phản hồi khi con người phải can thiệp, chẳng hạn như khi xảy ra lỗi. AI có thể cung cấp một mức độ sửa lỗi và tự quản lý các kết quả không mong muốn.

***Sử dụng nhiều hơn là chỉ robot***

Khi nghĩ về công nghiệp, hầu hết mọi người nghĩ về tự động hóa: robot tạo ra đồ vật. Tuy nhiên, xã hội thực sự đang ở ít nhất là cuộc cách mạng công nghiệp lần thứ tư; chúng ta đã có hơi nước, sản xuất hàng loạt, tự động hóa và bây giờ là giao tiếp (xem “Cách mạng công nghiệp – Từ Công nghiệp 1.0 đến Công nghiệp 4.0” tại Desouttertools.com để biết chi tiết). (Một số người đã nói về cấp độ thứ năm, cá nhân hóa; hãy xem bài đăng trên LinkedIn này, “Công nghiệp 5.0 - Tương lai của cá nhân hóa.”) Một AI cần thông tin từ nhiều nguồn khác nhau để thực hiện các nhiệm vụ một cách hiệu quả. Theo đó, một bối cảnh công nghiệp có thể thu thập càng nhiều thông tin từ nhiều nguồn khác nhau thì AI có thể hoạt động càng tốt (giả sử rằng dữ liệu cũng được quản lý đúng cách). Với ý tưởng đa nguồn này, các bối cảnh công nghiệp thuộc mọi loại hiện dựa vào Công cụ truyền thông công nghiệp (ICE) để phối hợp giao tiếp giữa tất cả các nguồn khác nhau mà AI yêu cầu.

Robot thực hiện phần lớn công việc thực tế trong môi trường công nghiệp, nhưng bạn cũng cần các cảm biến để đánh giá các rủi ro tiềm ẩn, chẳng hạn như bão. Tuy nhiên, sự phối hợp đang ngày càng trở nên quan trọng hơn để đảm bảo các hoạt động vẫn hiệu quả. Ví dụ, đảm bảo rằng các xe tải chở nguyên liệu đến đúng thời điểm, trong khi các xe tải khác chở hàng thành phẩm có sẵn khi cần, là những nhiệm vụ thiết yếu để duy trì hiệu quả hoạt động của sàn kho. AI cần biết về tình trạng bảo trì của tất cả các thiết bị để đảm bảo rằng các thiết bị nhận được sự chăm sóc tốt nhất có thể (để cải thiện độ tin cậy); AI cũng cần biết thời điểm thiết bị ít cần thiết nhất (để cải thiện hiệu quả). AI cũng cần xem xét các vấn đề như chi phí tài nguyên. Có lẽ có thể đạt được lợi thế bằng cách chạy một số thiết bị vào buổi tối khi điện rẻ hơn.

***Chỉ dựa vào tự động hóa***

Những ví dụ ban đầu về các nhà máy không có con người bao gồm các thiết lập đặc biệt, chẳng hạn như các nhà máy sản xuất chip đòi hỏi môi trường cực kỳ sạch sẽ. Tuy nhiên, kể từ khi bắt đầu sớm như vậy, tự động hóa đã lan rộng. Do những nguy hiểm đối với con người và chi phí sử dụng con người để thực hiện một số loại

nhiệm vụ công nghiệp, ngày nay bạn có thể tìm thấy nhiều ví dụ về các nhà máy thông thường không cần sự can thiệp của con người (xem “Không có con người, chỉ có robot” tại Singularity Hub.com để biết ví dụ). Thuật ngữ cho loại hình công nghiệp đó là *sản xuất tắt đèn,* được trình bày chi tiết trong “Lights outManufacturing…Is it Possible?” tại blog Syscon Plantstar.

\*Một số công nghệ tại một thời điểm nào đó sẽ cho phép thực hiện tất cả các nhiệm vụ liên quan đến nhà máy mà không cần sự can thiệp của con người [(xem https://waypointrobotics.com/blog/manufacturing-trends/ để biết ví dụ). Vấn đề là cuối cùng xã hội](https://waypointrobotics.com/blog/manufacturing-trends/) sẽ cần tìm việc làm, ngoài các công việc lặp đi lặp lại trong nhà máy, để con người thực hiện.

## 4. Tạo ra một môi trường an toàn

Một trong những vai trò thường được nêu nhất của AI, bên cạnh việc tự động hóa các tác vụ, là giữ an toàn cho con người theo nhiều cách khác nhau. Các bài viết như “7 lý do bạn nên chấp nhận, không phải sợ hãi, Trí tuệ nhân tạo” tại Futurism.com mô tả một môi trường mà AI đóng vai trò trung gian, chịu tác động mà con người thường phải chịu khi xảy ra sự cố an toàn. An toàn có nhiều hình thức. Đúng vậy, AI sẽ giúp làm việc trong nhiều môi trường an toàn hơn, nhưng nó cũng sẽ giúp tạo ra một môi trường lành mạnh hơn và giảm thiểu rủi ro liên quan đến các tác vụ thông thường, bao gồm cả lướt Internet. Các phần sau đây cung cấp tổng quan về những cách mà AI có thể cung cấp một môi trường an toàn hơn.

***Xem xét vai trò của sự nhàm chán trong các vụ tai nạn***

Từ việc lái xe đến làm việc, sự buồn chán làm tăng các loại tai nạn (xem “Khảo sát về lái xe mất tập trung năm 2021: Người lái xe thú nhận hành vi xấu” tại Insurance.com và “Buồn chán trong công việc có phải là mối nguy hiểm tại nơi làm việc không?” tại Risk and Insurance.com). Trên thực tế, bất cứ khi nào ai đó được cho là thực hiện một nhiệm vụ đòi hỏi bất kỳ mức độ tập trung nào và thay vào đó lại hành động như thể họ đang nửa tỉnh nửa mê, thì kết quả hiếm khi tốt. Vấn đề này nghiêm trọng và đáng kể đến mức bạn có thể tìm thấy rất nhiều bài viết về chủ đề này, chẳng hạn như “Mô hình hóa sự buồn chán của con người khi làm việc: công thức toán học và khuôn khổ xác suất” tại Emerald Insight.com. Các giải pháp có dạng các bài viết như “Mô hình hóa sự luân chuyển công việc trong các hệ thống sản xuất: Nghiên cứu về sự buồn chán của nhân viên và các biến thể kỹ năng” tại ResearchGate.net. Việc tai nạn có thực sự xảy ra (hay là suýt xảy ra) phụ thuộc vào cơ hội ngẫu nhiên. Hãy tưởng tượng thực sự phát triển các thuật toán giúp xác định xác suất xảy ra tai nạn do buồn chán trong một số điều kiện nhất định.

***Sử dụng AI để tránh các vấn đề an toàn***

Không có AI nào có thể ngăn ngừa tai nạn do con người gây ra, chẳng hạn như sự buồn chán. Trong trường hợp tốt nhất, khi con người quyết định thực sự tuân theo các quy tắc mà AI giúp tạo ra, AI chỉ có thể giúp tránh các vấn đề tiềm ẩn. Không giống như robot của Asimov, không có biện pháp bảo vệ ba luật nào được áp dụng trong bất kỳ môi trường nào; con người phải lựa chọn để duy trì sự an toàn. Với thực tế này, AI có thể giúp theo những cách sau:

 Đề xuất luân phiên công việc (cho dù ở nơi làm việc, trên ô tô hay thậm chí ở nhà) để công việc luôn thú vị

 Theo dõi hiệu suất của con người để đề xuất tốt hơn thời gian ngừng hoạt động do mệt mỏi hoặc các yếu tố khác

 Hỗ trợ con người thực hiện các nhiệm vụ để kết hợp trí thông minh mà con người cung cấp với thời gian phản ứng nhanh của AI

 Tăng cường khả năng phát hiện của con người để các vấn đề an toàn tiềm ẩn trở nên rõ ràng hơn

 Đảm nhận các công việc lặp đi lặp lại để con người ít có khả năng bị mệt mỏi và có thể tham gia vào các khía cạnh thú vị của bất kỳ công việc nào

***Hiểu rằng AI không thể loại bỏ các vấn đề về an toàn***

Đảm bảo an toàn hoàn toàn ngụ ý khả năng nhìn thấy tương lai. Vì tương lai là điều chưa biết, nên những rủi ro tiềm ẩn đối với con người tại bất kỳ thời điểm nào cũng không thể biết được vì những tình huống bất ngờ có thể xảy ra. Một tình huống bất ngờ là tình huống mà những người phát triển ban đầu của một chiến lược an toàn cụ thể không hình dung ra. Con người rất giỏi trong việc tìm ra những cách mới để rơi vào tình thế khó khăn, một phần vì chúng ta vừa tò mò vừa sáng tạo. Việc tìm ra phương pháp để vượt qua sự an toàn do AI cung cấp nằm trong bản chất của con người vì con người tò mò; chúng ta muốn xem điều gì sẽ xảy ra nếu chúng ta thử một điều gì đó — thường là điều gì đó ngu ngốc. Những tình huống không thể đoán trước không phải là vấn đề duy nhất mà AI phải đối mặt. Ngay cả khi ai đó tìm ra mọi cách có thể khiến con người trở nên không an toàn, thì sức mạnh xử lý cần thiết để phát hiện sự kiện và xác định phương hướng hành động sẽ là vô cùng lớn. AI sẽ hoạt động rất chậm đến nỗi phản ứng của nó luôn diễn ra quá muộn để tạo ra bất kỳ sự khác biệt nào. Do đó, các nhà phát triển thiết bị an toàn thực sự yêu cầu AI thực hiện mức độ an toàn cần thiết phải xử lý các xác suất và sau đó bảo vệ chống lại các tình huống có khả năng xảy ra cao nhất.