

SAU BÀI HỌC NÀY EM SẼ:

- Chỉ ra được một số lĩnh vực của khoa học công nghệ và đời sống đã và đang phát triển mạnh mẽ dựa trên những thành tựu to lớn của Trí tuệ nhân tạo.
- Nêu được ví dụ để thấy một hệ thống Trí tuệ nhân tạo có tri thức, có khả năng suy luận và khả năng học,...
- Nêu được cảnh báo về sự phát triển của Trí tuệ nhân tạo trong tương lai.



Trong các cuộc tranh luận về AI thường có hai quan điểm sau:

- Trong tương lai, AI sẽ có thể thông minh hơn nhiều và thay thế hoàn toàn con người.
- AI có thể làm được nhiều việc nhưng không thể thay thế con người.

Em ủng hộ quan điểm nào trong hai quan điểm trên? Vì sao?

1. TRÍ TUỆ NHÂN TẠO VÀ SỰ PHÁT TRIỂN CỦA MỘT SỐ LĨNH VỰC

Hoạt động 1 Tìm hiểu một số lĩnh vực phát triển mạnh mẽ nhờ ứng dụng của AI

Ngày nay, nhiều lĩnh vực đang có thay đổi lớn lao nhờ ứng dụng AI. Hãy chỉ ra một vài lĩnh vực mà em tìm hiểu được qua các phương tiện thông tin đại chúng và Internet.



Nhiều lĩnh vực khoa học, công nghệ và đời sống đã và đang nhận được rất nhiều lợi ích từ sự phát triển của AI. Dưới đây là một vài lĩnh vực tiêu biểu:

Hệ chuyên gia: Nhờ những thành tựu của AI, lĩnh vực này đã phát triển đáng kể. Ban đầu, hệ chuyên gia là chương trình máy tính được thiết kế dựa trên các luật suy diễn và tri thức của chuyên gia trong từng lĩnh vực cụ thể. Với sự phát triển của AI, đặc biệt là Học máy, nhiều hệ chuyên gia đã có khả năng tự học từ dữ liệu để tự hình thành các luật và tri thức dựa trên dữ liệu.

Y học và chăm sóc sức khỏe: AI được sử dụng để cải thiện chất lượng hình ảnh y tế, làm nổi bật những cấu trúc bất thường bên trong cơ thể, thực hiện đo đạc các chỉ số lâm sàng, hỗ trợ đưa ra các chẩn đoán và hướng điều trị chính xác, kịp thời. Ví dụ, các chuyên gia y tế khẳng định phần mềm IBM Watson for Oncology đã góp phần nâng cao hiệu quả điều trị ung thư.

Giao thông vận tải: AI đã được sử dụng để phát triển các phương tiện tự lái, quản lý giao thông thông minh và định tuyến phương tiện vận tải. Sự phát triển mạnh mẽ của ô tô tự lái, máy bay không người lái,... trong những năm gần đây không thể có được nếu không có AI.

Tài chính, ngân hàng: AI không chỉ hỗ trợ tự động hoá cập nhật chứng từ, hoá đơn vào cơ sở dữ liệu, mà còn giúp phân tích, xử lý dữ liệu một cách hiệu quả để hỗ trợ quyết định đầu tư, phát hiện và ngăn chặn gian lận, nâng cao trải nghiệm khách hàng.

Sản xuất: AI được sử dụng để cải thiện hiệu suất, hiệu quả và sự phát triển bền vững của các lĩnh vực sản xuất. Trong công nghiệp, AI giúp tự động hoá nhiều quá trình, từ chế tạo, lắp ráp, kiểm tra chất lượng đến quản lý chuỗi cung ứng. Các robot và hệ thống tự động hoá được tích hợp AI có khả năng thực hiện nhiều công việc lặp đi lặp lại một cách hiệu quả (Hình 2.1). Trong nông nghiệp, AI được sử dụng trong các trang trại thông minh để theo dõi những yếu tố có ảnh hưởng trực tiếp tới hiệu quả nuôi trồng như điều kiện thời tiết, đất đai, sức đề kháng với dịch bệnh và thời tiết của vật nuôi, cây trồng. AI có thể giúp người nông dân tối ưu hoá quy trình chăm sóc vật nuôi và cây trồng; hợp lý hoá tưới tiêu, dự đoán mùa vụ, xác định thời điểm thu hoạch tối ưu dựa trên dữ liệu về điều kiện chăm sóc, thời tiết, đất đai và cây giống,...



Hình 2.1. Robot và hệ thống tự động hoá

Giáo dục: AI được sử dụng để phát triển các nền tảng học tập được cá nhân hoá và hỗ trợ đánh giá kết quả học tập. Ví dụ, AI được sử dụng để phát triển các nền tảng học trực tuyến thông minh, có khả năng theo dõi tiến trình học tập, đề xuất nội dung học tập phù hợp và cung cấp phản hồi tức thì cho từng cá nhân người học. Các trợ lý học tập ảo dựa trên AI có thể hỗ trợ học sinh và giáo viên bằng cách trả lời câu hỏi, cung cấp hướng dẫn và tài liệu học tập,...

Có thể chỉ ra những ảnh hưởng của AI tới nhiều lĩnh vực khoa học và đời sống khác. Chẳng hạn, trong lĩnh vực khoa học xã hội và nhân văn, AI được sử dụng không chỉ để thu thập và phân tích tự động quan điểm xã hội, mà còn để mô phỏng và mô hình hoá nhiều hiện tượng xã hội và nhân học. Một số ứng dụng AI có khả năng sáng tạo các tác phẩm âm nhạc, hội hoạ, văn học theo nhiều phong cách khác nhau. Sự kết hợp IoT và AI (AIoT) cho phép các nhà khoa học giám sát môi trường tự nhiên và theo dõi tình hình biến đổi khí hậu. Những thành tựu của xử lý ngôn ngữ tự nhiên và thị giác máy tính trong AI đã góp phần phát triển hàng loạt các ứng dụng thiết thực cho đời sống như dịch thuật tự động, hỗ trợ khách hàng bằng ngôn ngữ tự nhiên, các hệ thống nhận dạng hình ảnh và video đa dạng,... Ngày nay, trong nhiều lĩnh vực như công nghiệp trò chơi điện tử hay thám hiểm không gian vũ trụ,... việc ứng dụng AI đã thật sự trở thành một phần không thể thiếu.

Ngày nay, nhiều lĩnh vực của khoa học công nghệ và đời sống đã và đang phát triển mạnh mẽ dựa trên những thành tựu to lớn của AI. Có thể nói AI đã trở thành một phần không thể thiếu của cuộc sống hiện đại.



Hãy chỉ ra một vài lĩnh vực có sự phát triển đột phá nhờ những thành tựu của AI.

2. TRÍ TUỆ NHÂN TẠO VÀ MỘT VÀI CẢNH BÁO

Hoạt động 2 Những hình dung của em về tương lai của AI

Trên cơ sở những thông tin về sự phát triển của AI ngày nay, hãy cho biết những suy nghĩ của em về tương lai của AI.



Sự phát triển của AI đã và đang diễn ra với tốc độ nhanh chóng, có tiềm năng cách mạng hoá nhiều lĩnh vực khoa học và đời sống. Những ứng dụng hết sức đa dạng của AI đã cho ta thấy rõ điều đó. Tuy nhiên, các thành tựu hiện tại của AI vẫn chỉ hạn chế trong phạm vi *Trí tuệ nhân tạo hẹp/Trí tuệ nhân tạo yếu*. Trong tương lai, nhiều chuyên gia kì vọng sẽ phát triển thành công *Trí tuệ nhân tạo tổng quát/ Trí tuệ nhân tạo mạnh*, có năng lực trí tuệ như con người, bao gồm cả khả năng áp dụng tri thức từ lĩnh vực này sang lĩnh vực khác. Có thể nói đây là mục tiêu dài hạn và hết sức phức tạp trong nghiên cứu và phát triển AI.

Tuy vậy, triển vọng thực tế không phải là không có nhiều hứa hẹn. Sự phát triển mạnh mẽ trong những năm gần đây của AI tạo sinh (Generative AI) là một minh chứng. AI tạo sinh tập trung vào việc xây dựng các thuật toán và mô hình có thể tạo nội dung (hình ảnh, âm thanh, văn bản) một cách tự động, được sử dụng trong nhiều lĩnh vực, bao gồm nghệ thuật, giải trí, quảng cáo và trò chơi.

Một ví dụ điển hình có thể nêu ở đây là ChatGPT, một hệ thống xử lí ngôn ngữ tiên tiến do OpenAI phát triển. Nó là một hệ thống có tri thức, có khả năng suy luận và khả năng học,... Cụ thể:

ChatGPT được huấn luyện trên một lượng lớn dữ liệu văn bản, ví dụ, phiên bản GPT-3 được huấn luyện trên hàng vạn văn bản thuộc nhiều lĩnh vực khác nhau trên Internet, bao gồm sách, bài báo, trang web và nhiều nguồn khác.

Một trong những khả năng được đánh giá cao của ChatGPT đó là khả năng hiểu và tạo văn bản ngôn ngữ tự nhiên. Nó có thể đọc hiểu các câu hỏi và yêu cầu của người dùng, sau đó trả lời một cách hợp lí và thông minh. Đây chính là thể hiện khả năng suy luận và tri thức về ngôn ngữ của ChatGPT.

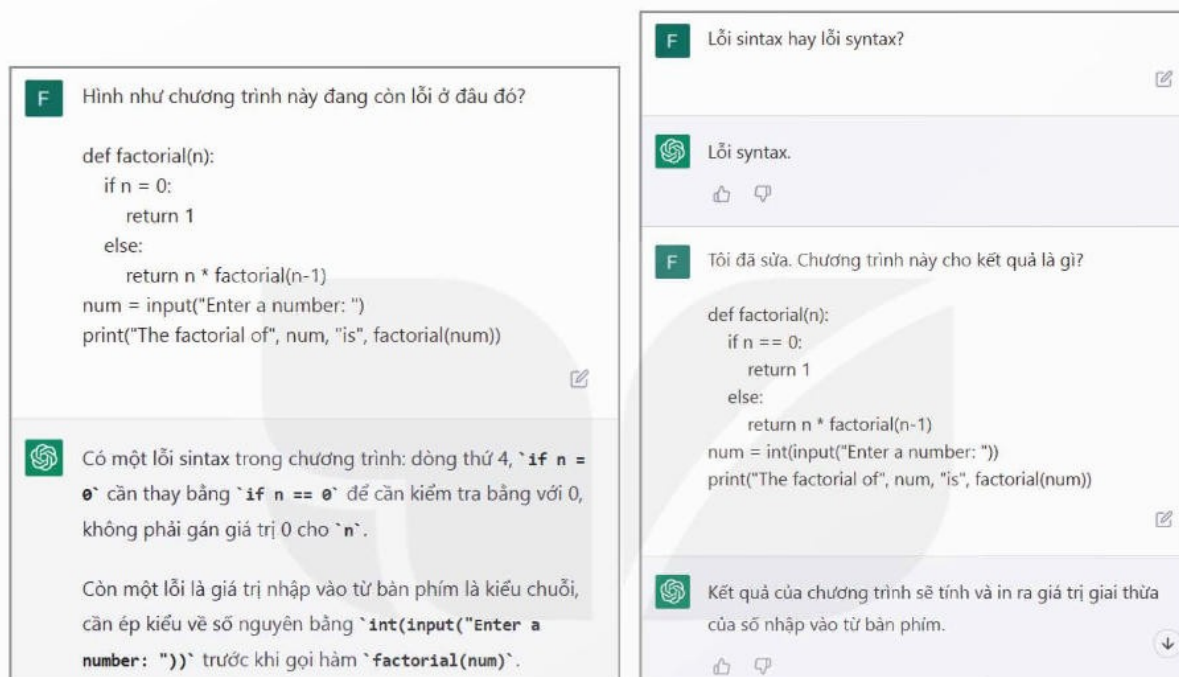
Hơn thế nữa, ChatGPT không chỉ cung cấp những câu trả lời cố định mà còn có khả năng đưa ra các câu trả lời đa dạng dựa trên ngữ cảnh và yêu cầu cụ thể của người dùng. Điều này cho thấy sự linh hoạt trong suy nghĩ và khả năng suy luận của nó.

ChatGPT có khả năng xử lí thông tin phức tạp để trả lời các câu hỏi mang tính phân tích và tổng hợp. Nó có thể xử lí dữ liệu từ nhiều nguồn để cung cấp các giải pháp dựa trên khả năng suy luận và tổng hợp tri thức.

Việc có thể tương tác một cách liên tục và cải thiện nội dung trả lời dựa trên phản hồi của người dùng cho thấy ChatGPT có khả năng nắm bắt và thích nghi với ngữ cảnh cụ thể.

ChatGPT có khả năng làm việc với nhiều thứ tiếng khác nhau, tuy nhiên, mức độ hiểu và khả năng trả lời có thể khác nhau tùy thuộc vào mức độ phát triển và tập trung của nó cho từng ngôn ngữ cụ thể. Nói chung, ChatGPT hoạt động tốt nhất trong các ngôn ngữ phổ biến và có nhiều dữ liệu huấn luyện, chẳng hạn như tiếng Anh, Tây Ban Nha, Pháp, Đức, Ý và một số ngôn ngữ khác. Đối với các ngôn ngữ còn lại, đặc biệt là các ngôn ngữ ít phổ biến hoặc không có nhiều dữ liệu huấn luyện, ChatGPT có thể có khả năng hạn chế hơn. Tuy nhiên, nhà phát triển và cộng đồng có thể tạo ra phiên bản tùy chỉnh của ChatGPT cho các ngôn ngữ cụ thể bằng cách tinh chỉnh và huấn luyện lại mô hình trên dữ liệu ngôn ngữ đó.

Tóm lại, ChatGPT được coi là một hệ thống AI có hiểu biết sâu rộng, có khả năng tương tác với người dùng giống như một cuộc hội thoại giữa người với người. Nó có thể trích rút thông tin từ nhiều nguồn khác nhau và tạo ra văn bản chi tiết, mang tính tương tác, giống như cách con người trả lời các câu hỏi trong mọi lĩnh vực. Bên cạnh đó, ChatGPT còn có thể làm thơ, soạn nhạc, viết thư, thiết kế và thậm chí là cả sửa lỗi trong lập trình (Hình 2.2). Mặc dù còn nhiều hạn chế, song nhiều chuyên gia đánh giá ChatGPT là minh chứng cho sức mạnh của AI, coi ChatGPT là bước đột phá trong lĩnh vực AI của năm 2022, có thể dẫn tới những thay đổi lớn về cách làm việc, cách suy nghĩ, cách dạy và học trong nhà trường.



Hình 2.2. Một đoạn hội thoại được thực hiện với ChatGPT

Bên cạnh những lợi ích to lớn, sự phát triển của AI còn kéo theo những nguy cơ cần được cảnh báo, chẳng hạn:

- **Áp lực thất nghiệp:** AI có khả năng tự động hoá nhiều công việc, dẫn đến tình trạng thất nghiệp và tạo nên nhiều nguy cơ cho sự phát triển xã hội.
- **Ảnh hưởng quyền riêng tư:** Nhiều ứng dụng AI hoạt động dựa vào việc thu thập một lượng lớn dữ liệu cá nhân, điều này làm tăng mối lo ngại về quyền riêng tư có khả năng bị lạm dụng.
- **Khả năng thiếu minh bạch:** Phần lớn các ứng dụng AI hiện nay đều là các "hộp đen", gây khó khăn cho việc hiểu các quyết định được đưa ra như thế nào, dẫn đến việc thiếu trách nhiệm giải trình để đảm bảo tính minh bạch.
- **Rủi ro về an ninh, an toàn:** Nhiều ứng dụng AI được xây dựng và triển khai trực tuyến. Điều này có thể bị lợi dụng để xâm nhập hoặc tấn công thay đổi dữ liệu và mô hình, có thể dẫn đến những quyết định không chính xác do AI đưa ra. Các quyết định sai lầm đó có thể gây nguy hại trực tiếp cho con người, ví dụ chẩn đoán sai về tình trạng bệnh tật hay ra quyết định tấn công các mục tiêu dân sự trong xung đột vũ trang,...

Các cảnh báo trên cũng đặt ra một số khía cạnh đạo đức cần được xem xét để đảm bảo AI phải được xây dựng và sử dụng một cách minh bạch và có trách nhiệm. Thực tế đang đòi hỏi phải có những ràng buộc mang tính pháp lý đối với việc phát triển và ứng dụng AI trong một số lĩnh vực có khả năng ảnh hưởng trực tiếp tới sinh mạng con người. Ví dụ, ngăn cấm việc giao toàn quyền quyết định cho AI điều khiển vũ khí sát thương. Những thử nghiệm liên kết sinh học và công nghệ với nhau theo cách kết hợp bộ não con người và “bộ não” robot trong một chỉnh thể hợp nhất cũng đang là vấn đề gây nhiều tranh cãi. Ngoài khía cạnh đạo đức, những thử nghiệm này mang tính rủi ro cao. Bởi lẽ, trong trường hợp “bộ não” robot được phát triển tới mức thông minh hơn con người, việc công nghệ có thể được sử dụng để kiểm soát hoặc thay đổi ý thức, hành vi của con người theo một cách không mong muốn là nguy cơ hoàn toàn có thực.

Vì vậy, cần phải có các giải pháp đảm bảo an toàn và giám sát chặt chẽ quá trình phát triển AI, cũng như đảm bảo sao cho AI có thể mang lại nhiều cơ hội cho sự tiến bộ, nâng cao chất lượng cuộc sống, phục vụ lợi ích cộng đồng, không gây hậu quả xấu cho xã hội. Đây đang là mối quan tâm lớn lao của các chuyên gia, các nhà khoa học, các chính trị gia, các nhà hoạt động xã hội cho tới người dân trên toàn thế giới.

Các thành tựu hiện tại của AI nói chung đều thuộc phạm vi *Trí tuệ nhân tạo hẹp/ Trí tuệ nhân tạo yếu*. Tuy nhiên, những hệ thống như ChatGPT cũng cho thấy những khả năng không nhỏ của AI. Việc phát triển *Trí tuệ nhân tạo tổng quát/ Trí tuệ nhân tạo mạnh*, có năng lực trí tuệ như con người, bao gồm cả khả năng áp dụng tri thức từ lĩnh vực này sang lĩnh vực khác vẫn đang trong quá trình nghiên cứu và phát triển. Bên cạnh đó, sự phát triển của AI cũng đặt ra nhiều cảnh báo về những nguy cơ liên quan tới các khía cạnh đạo đức và xã hội cần phải được giải quyết.



1. Trên cơ sở các phân tích về khả năng xử lý ngôn ngữ của ChatGPT ở trên, hãy chỉ ra một vài ví dụ ứng dụng AI cũng có khả năng đọc hiểu, trả lời câu hỏi của người dùng.
2. Hãy nêu một số nguy cơ có thể xảy ra liên quan tới việc phát triển AI.



LUYỆN TẬP

1. ChatGPT là hệ thống AI sử dụng ngôn ngữ để tương tác với con người. Hãy nêu một vài ứng dụng AI sử dụng hình ảnh để tương tác.
2. Vì sao cần ngăn cấm việc giao toàn quyền quyết định cho AI điều khiển vũ khí sát thương?



VẬN DỤNG

Hãy tìm hiểu cách thức hoạt động của ứng dụng tìm đường trên Google Maps để nhận biết các biểu hiện “thông minh” của ứng dụng này.