# CHƯƠNG II. THUẬT TOÁN K-MEANS

## 2.1. Tổng quan về Machine Learning

2.1.1.  Lịch sử hình thành và phát triển.

Nguồn gốc của học máy có thể bắt nguồn từ những năm 1940 , khi các nhà nghiên cứu bắt đầu khám phá những bài toán nhận dạng mẫu cơ bản đầu tiên và nghiên cứu những mạng lưới thần kinh. Giai đoạn đầu của lịch sử Machine Learning - ML này được đánh dấu bằng những ý tưởng đột phá và những nỗ lực không ngừng nghỉ trong việc tạo ra những chiếc máy tính có thể bắt chước quá trình suy nghĩ của con người. Năm 1943, Walter Pitts và Warren McCulloch đã nghĩ ra mô hình toán học đầu tiên của mạng lưới thần kinh nhân tạo, đặt nền móng cho mạng lưới thần kinh hiện đại và phát triển các bộ công cụ học máy phân tán McCulloch & Pitts, 1943).

Những người tiên phong trong thời kỳ mở đầu của ML có thể kể đến như Donald Hebb , Alan Turing và Arthur Samuel , mặc dù họ không phải là người khởi xướng duy nhất, tuy nhiên những nghiên cứu và công sự của họ đã đóng góp đáng kể vào sự phát triển của ngành học máy . Công trình của Hebb về sự giao tiếp trong hệ thống thần kinh, bài kiểm tra trí tuệ nhân tạo của Turing và việc Samuel đặt ra thuật ngữ “học máy” đều góp phần vào lĩnh vực trí tuệ nhân tạo (AI) đang phát triển và đặt nền móng cho vô số thuật toán học máy mà chúng ta biết ngày nay (Hebb, 1949; Turing, 1950; Samuel, 1959).

Hình 2. 3: Đại kiện tướng Ke Jie thua bẽ bàng 2 trận liên tiếp trước trí tuệ nhân tạo của Google.

Một vài sự kiện tiêu biểu trong quá trình hình thành và phát triển của học máy có thể kể đến những mốc thời gian sau:

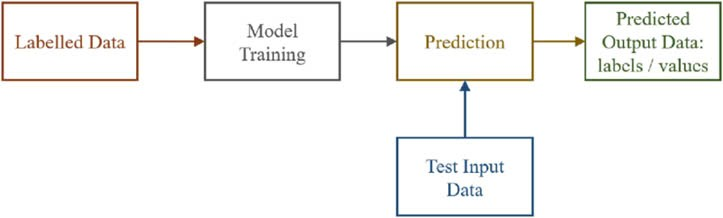
* Năm 1943, Walter Pitts và Warren McCulloch phát triển mô hình học máy đầu tiên để giải quyết thử thách do John von Neumann đặt ra:làm thế nào để các máy tính có thể giao tiếp với nhau.
* Năm 1949, Donald Hebb đưa ra khái niệm về giao tiếp giữa các nơ-ron trong hệ thống thần kinh.
* Năm 1950, Alan Turing giới thiệu bài kiểm tra Turing đánh giấu cột mốc quan trọng trong lĩnh vực AI.
* Năm 1951, Marvin Lee Minsky phát minh ra máy tính tăng cường tương tự thần kinh (SNARC).
* Năm 1967, Cover cùng các cộng sự của mình tạo ra thuật toán kNN.
* Năm 1980, phát hiện ra neocognitron – một mạng lưới thần kinh nhân tạo nhiều lớp, là tiền đề cho mạng tích chập CNN sau này (Fukushima, 1980).
* Năm 1997, Deep Blue của IBM gây sốc cả thế giới khi đánh bại nhà vô địch cờ vua thời bấy giờ.
* Năm 2006, thuật ngữ “Deep Learning” được Geoffrey Hinton đặt ra để mô tả các thuật toán mới cho phép máy tính “nhìn” và phân biệt các đối tượng cững như văn bản trong hình ảnh, video.
* Năm 2017, Google đã công bố nghiên cứu đầu tiên về kiến trúc deep learning mang tên Transformers (Vaswani et al., 2017).
* Năm 2023, OpenAI phát hành ChatGPT.

Từ những ngày đầu nhận dạng mẫu đơn giản cho đến các mô hình học tập phức tạp ngày nay, lịch sử của học máy (ML) đã là một **hành trình hấp dẫn** . Đó là câu chuyện về con người nỗ lực tạo ra những chiếc máy tính có thể học hỏi, thích ứng và đưa ra những quyết định sáng suốt, giống như quá trình nhận thức của chính chúng ta. Chuyến đi này đã định hình lại các ngành công nghiệp, xác định lại sự tương tác giữa con người và máy tính và mở ra một thế giới tiềm năng chưa được khai thác.

2.1.2.  Phân loại.

2.1.2.1. Học có giám sát.

Học có giám sát (Supervised Learning) là một phương pháp quan trọng trong lĩnh vực Machine Learning, trong đó mô hình được huấn luyện trên một tập dữ liệu đã được gán nhãn. Đây là quá trình huấn luyện một mô hình ML bằng cách sử dụng dữ liệu đầu vào (input) và đầu ra mong muốn tương ứng (output) đã biết trước, nhằm mục đích tạo ra một mô hình có khả năng dự đoán chính xác output cho các input mới mà nó chưa từng thấy trước đó. Trong học có giám sát, tập dữ liệu huấn luyện bao gồm các cặp (input, output) đã được gán nhãn, trong đó input là các đặc trưng của dữ liệu, còn output là nhãn hoặc giá trị mục tiêu. Mô hình được huấn luyện để tìm ra mối quan hệ giữa input và output, điều chỉnh các tham số nội bộ để dự đoán output càng chính xác càng tốt. Học có giám sát thường được áp dụng cho hai loại bài toán chính: phân loại (classification), dự đoán một nhãn rời rạc, và hồi quy (regression), dự đoán một giá trị liên tục. Ví dụ điển hình của học có giám sát bao gồm nhận dạng chữ viết tay và dự đoán giá nhà. Phương pháp này có ưu điểm là hiệu suất cao khi có đủ dữ liệu được gán nhãn và dễ đánh giá hiệu quả của mô hình.



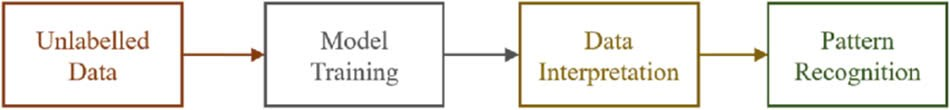
Hình 2.4: Sơ đồ tổng quát của học máy có giám sát.

Tuy nhiên, nó cũng có nhược điểm là cần một lượng lớn dữ liệu đã được gán nhãn, có thể tốn kém và mất thời gian, đồng thời có thể gặp vấn đề overfitting nếu mô hình quá phức tạp hoặc dữ liệu huấn luyện không đại diện. Các thuật toán phổ biến trong học có giám sát bao gồm hồi quy tuyến tính và logistic, cây quyết định và Random Forest, Support Vector Machines (SVM), và Neural Networks. Học có giám sát là một trong những phương pháp cơ bản và quan trọng nhất trong Machine Learning, được ứng dụng rộng rãi trong nhiều lĩnh vực như xử lý ngôn ngữ tự nhiên, thị giác máy tính, và phân tích dự đoán.

2.5.2.2. Học không giám sát.

Học không giám sát (Unsupervised Learning) là một phương pháp quan trọng trong Machine Learning, trong đó mô hình được huấn luyện trên dữ liệu không được gán nhãn. Khác với học có giám sát, phương pháp này không có sẵn thông tin về kết quả mong muốn, mà thay vào đó, mục tiêu là tìm ra cấu trúc hoặc mẫu ẩn trong dữ liệu. Học không giám sát tập trung vào việc khám phá và mô tả các đặc điểm thú vị của dữ liệu mà không cần hướng dẫn cụ thể.

Trong học không giám sát, mô hình phải tự tìm ra các mối quan hệ và cấu trúc trong dữ liệu. Điều này có thể bao gồm việc nhóm các điểm dữ liệu tương tự nhau (clustering), giảm số chiều của dữ liệu để tìm ra các đặc trưng quan trọng nhất (dimensionality reduction), hoặc phát hiện các mẫu bất thường (anomaly detection). Ví dụ phổ biến của học không giám sát bao gồm phân khúc khách hàng trong tiếp thị, phát hiện gian lận trong giao dịch tài chính, và tổ chức các bài báo tin tức theo chủ đề.



Hình 2.5: Sơ đồ tổng quát học máy không giám sát.

Các thuật toán phổ biến trong học không giám sát bao gồm K-means clustering, Hierarchical clustering, Principal Component Analysis (PCA), và các mô hình mạng neural tự mã hóa (autoencoders). Ưu điểm chính của học không giám sát là khả năng xử lý dữ liệu chưa được gán nhãn, cho phép khám phá các insight mới và không được biết trước từ dữ liệu. Nó cũng có thể giúp tiết kiệm thời gian và chi phí liên quan đến việc gán nhãn dữ liệu.

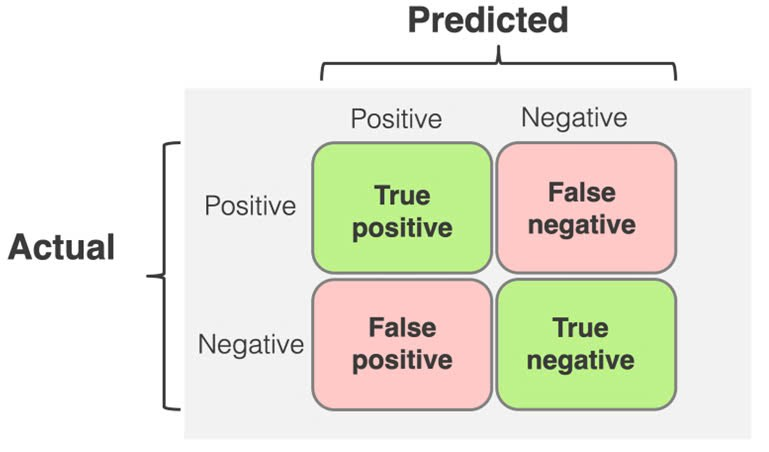
Tuy nhiên, học không giám sát cũng có những thách thức riêng. Việc đánh giá hiệu quả của mô hình có thể khó khăn hơn so với học có giám sát, vì không có tiêu chuẩn "đúng" rõ ràng để so sánh. Ngoài ra, kết quả của học không giám sát có thể khó giải thích hoặc ít có ý nghĩa thực tế nếu không được diễn giải cẩn thận.

Mặc dù có những thách thức, học không giám sát vẫn là một công cụ mạnh mẽ trong kho vũ khí của các nhà khoa học dữ liệu và chuyên gia ML. Nó cho phép khám phá những hiểu biết mới từ dữ liệu, tìm ra các mẫu ẩn, và có thể được sử dụng như một bước tiền xử lý cho các tác vụ học có giám sát phức tạp hơn. Khi kết hợp với các phương pháp khác, học không giám sát đóng vai trò quan trọng trong việc khai thác giá trị từ khối lượng dữ liệu ngày càng tăng trong thế giới hiện đại.

2.1.3.  Các phương pháp đánh giá.

Mô hình học máy thường là các mô hình dự đoán, vì vậy mô hình không có thể dự đoán chính xác 100% những gì chúng ta cung cấp. Với các bài toán như phân loại, phân lớp hay nhận dạng; các phương pháp đánh giá xem một mô hình có hiệu năng cao hay thấp thường được đánh giá thông qua confusion matrix và accuracy.

Confusion Matrix (Ma trận nhầm lẫn) là một bảng thể hiện hiệu suất của mô hình phân loại bằng cách so sánh các dự đoán với giá trị thực tế. Trong trường hợp phân loại nhị phân, ma trận này gồm bốn ô: True Positives (TP), True Negatives (TN), False Positives (FP), và False Negatives (FN). TP và TN là số lượng dự đoán đúng, trong khi FP và FN là số lượng dự đoán sai. Ma trận này cung cấp cái nhìn tổng quan về các loại lỗi mà mô hình mắc phải, giúp hiểu rõ hơn về hiệu suất của mô hình.



Hình 3. 1: Confusion matrix trong mô hình học máy.

Accuracy (Độ chính xác) là tỷ lệ các dự đoán đúng trên tổng số dự đoán. Nó được tính bằng công thức:

Accuracy cung cấp một số đo tổng quát về hiệu suất của mô hình, nhưng có thể gây hiểu nhầm trong trường hợp dữ liệu không cân bằng. Ví dụ, nếu 95% dữ liệu thuộc một lớp, một mô hình luôn dự đoán lớp đó sẽ có accuracy 95%, nhưng không hữu ích trong thực tế.

Mỗi phương pháp này có ưu điểm và hạn chế riêng. Confusion matrix cung cấp thông tin chi tiết nhất nhưng có thể khó diễn giải với dữ liệu lớn. Accuracy dễ hiểu nhưng có thể gây hiểu nhầm với dữ liệu không cân bằng. Trong thực tế, thường sử dụng kết hợp nhiều phương pháp để đánh giá toàn diện hiệu suất của mô hình phân loại.

2.1.4.  Những đột phá và ứng dụng.

Trong những năm gần đây, học máy đã trải qua một loạt đột phá và đổi mới, tạo nên một cuộc cách mạng trong lĩnh vực khoa học và công nghệ. Một số tiến bộ đáng chú ý bao gồm.

* Cuộc cách mạng deep learning vào năm 2012.
* Sự phát triển của các thuật toán học tăng cường, điển hình là AlphaGo của DeepMind.
* Những tiến bộ trong xử lý ngôn ngữ tự nhiên, một trong số đó là Chat GPT của OpenAI.

Những tiến bộ này đã cải thiện đáng kể khả năng và ứng dụng của trí tuệ nhân tạo. Họ cũng cho thấy tiềm năng của máy móc trong việc hiểu và tạo ra ngôn ngữ giống con người, mở đường cho các hệ thống AI phát triển tiên tiến hơn.

*Cuộc các mạng deep learning (2012)*

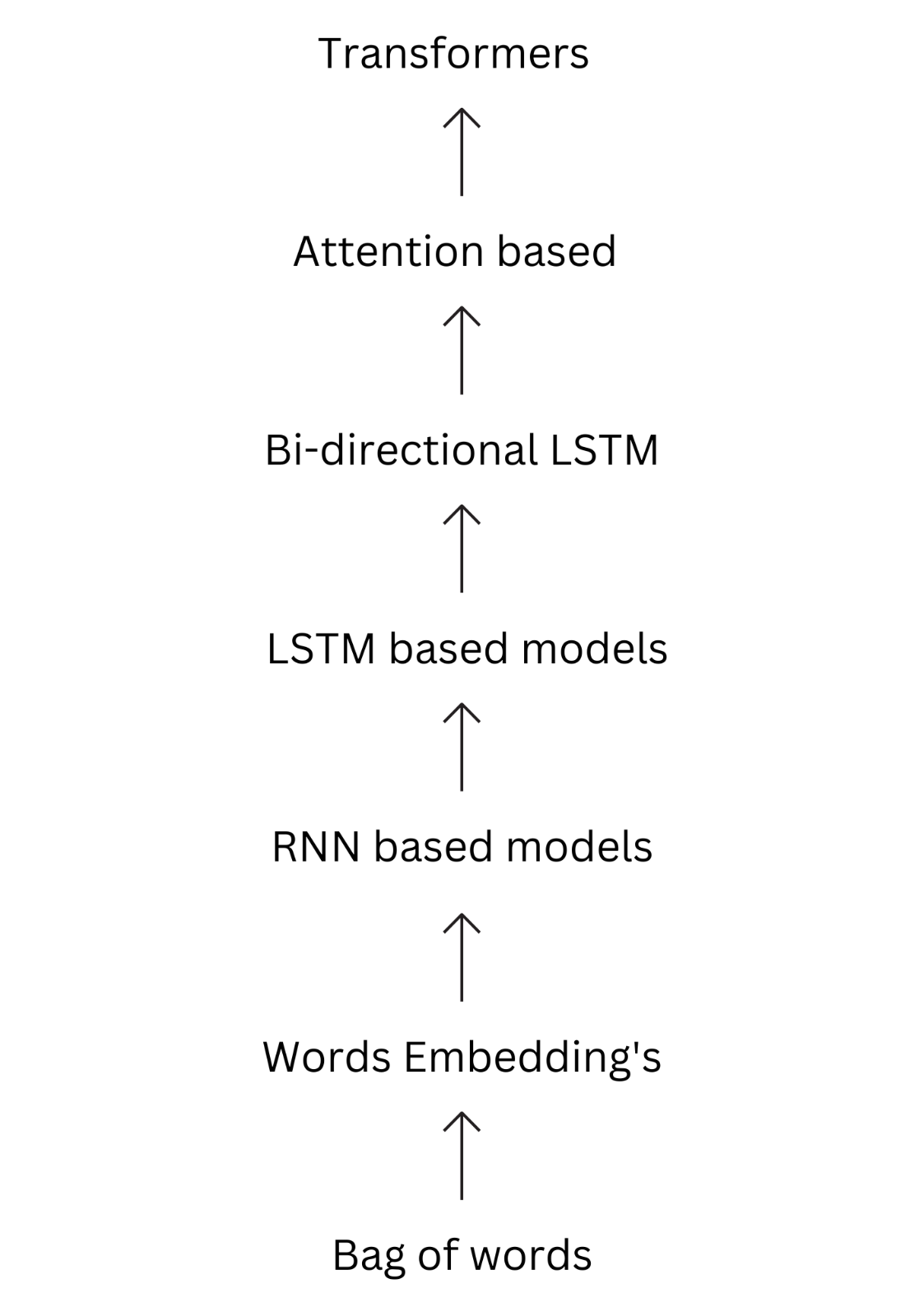
Cuộc cách mạng deep learning (2012) đánh dấu một bước ngoặt quan trọng trong lịch sử của machine learning. Cuộc các mạng này đặc biệt bởi sự ra đời của các kỹ thuật học sâu cho phép khả năng giải quyết vấn đề và nhận dạng mẫu phức tạp hết các mô hình trước đó. Một trong những bước phát triển quan trọng trong cuộc cách mạng này là việc tạo ra AlexNet, một mạng lưới thần kinh sâu giúp cải thiện đáng kể độ chính xác của hệ thống nhận dạng hình ảnh. Sự ra đời của AlexNet đã cho phép máy móc giải quyết những thách thức ngày càng phức tạp hơn, cuối cùng dẫn đến việc tạo ra:

* Hệ thống nhận dạng giọng nói chính xác hơn
* Cải thiện khả năng nhận dạng hình ảnh và video.
* Thuật toán xử lý ngôn ngữ tự nhiên tốt hơn.
* Hệ thống đề xuất nâng cao.
* Xe tự hành hiệu quả hơn

Những hệ thống phức tạp này dựa vào việc phân tích dữ liệu trong quá trình đào tạo của máy tính. Hiệu quả phần lớn phụ thuộc vào các mô hình học máy được sử dụng trong nền tảng học máy của chính chúng.

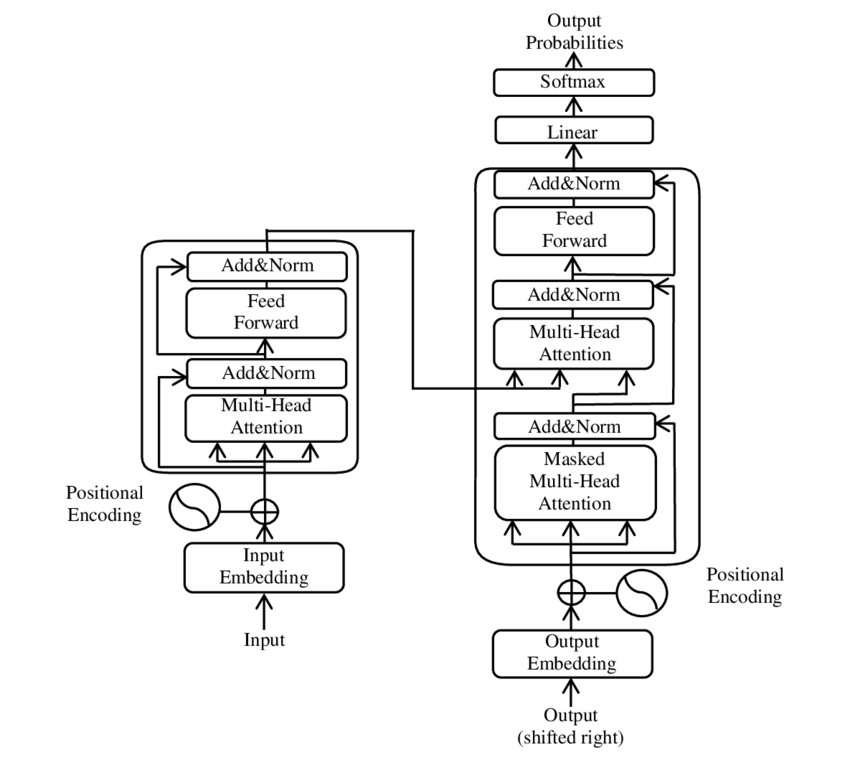
*Mô hình Transformer của Google (2017)*

Transformer là một kiến trúc deep learning dựa trên cơ chế chú ý nhiều đầu vào song song được Google đề xuất vào năm 2017.



Hình 2. 9: Dòng thời gian xuất hiện của các mô hình xử lý ngôn ngữ tự nhiên.

Trước khi có Transformer, các mô hình như RNN (Recurrent Neural Networks) và LSTM (Long Short-Term Memory) đã được sử dụng rộng rãi trong NLP. Tuy nhiên, những mô hình này gặp nhiều khó khăn trong việc xử lý các chuỗi dữ liệu dài do tính chất tuần tự của chúng. Transformer khắc phục những hạn chế này bằng cách sử dụng cơ chế attention, cho phép mô hình có thể xử lý các phần khác nhau của chuỗi đầu vào một cách song song.



Hình 2.10: Cấu trúc chính của mô hình Transformer.

Transformer bao gồm hai phần chính: bộ mã hóa (encoder) và bộ giải mã (decoder).

Bộ mã hóa của Transformer bao gồm một chuỗi các lớp mã hóa giống nhau. Mỗi lớp mã hóa bao gồm hai thành phần chính:

* **Multi-Head Self-Attention Mechanism**: Cho phép mô hình chú ý đến các phần khác nhau của chuỗi đầu vào một cách đồng thời. Cơ chế self-attention tính toán trọng số cho từng phần của đầu vào dựa trên mức độ liên quan của chúng với nhau.
* **Feed-Forward Neural Network**: Một mạng neural đơn giản được áp dụng độc lập cho từng vị trí trong chuỗi đầu vào.
* Mỗi lớp mã hóa còn bao gồm các lớp chuẩn hóa (layer normalization) và các cơ chế bỏ lớp (residual connections) để giúp huấn luyện mô hình hiệu quả hơn.
* Bộ giải mã cũng bao gồm một chuỗi các lớp giải mã giống nhau, với cấu trúc tương tự bộ mã hóa nhưng thêm một thành phần quan trọng:
* **Masked Multi-Head Self-Attention**: Giúp đảm bảo rằng mỗi vị trí trong chuỗi đầu ra chỉ có thể chú ý đến các vị trí trước đó trong chuỗi, để không tiết lộ thông tin tương lai.
* **Multi-Head Attention Over Encoder's Output**: Cho phép bộ giải mã chú ý đến các thông tin từ bộ mã hóa, kết hợp thông tin từ chuỗi đầu vào và chuỗi đầu ra để tạo ra dự đoán.

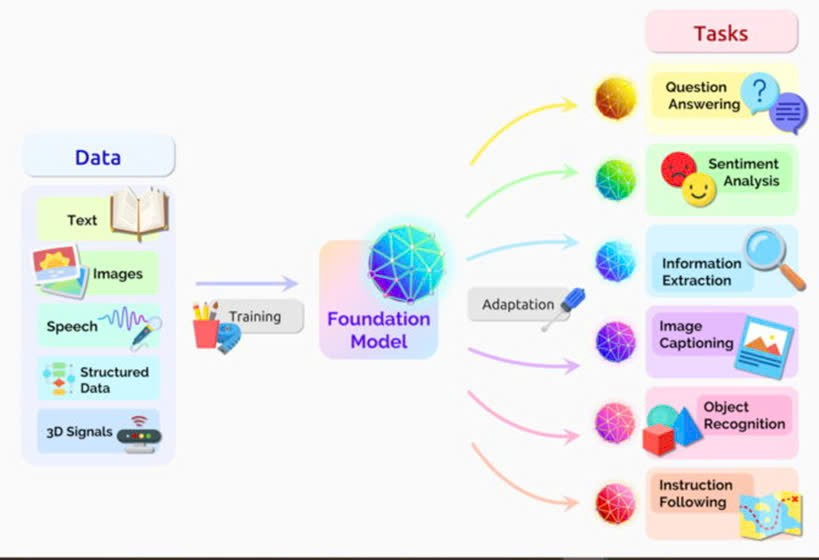
Cơ chế attention là trọng tâm của mô hình Transformer. Nó cho phép mô hình tập trung vào các phần quan trọng của chuỗi đầu vào để tạo ra biểu diễn ngữ nghĩa tốt hơn. Cơ chế này được thực hiện qua các bước:

* **Scaled Dot-Product Attention**: Tính toán trọng số chú ý bằng cách nhân các vector truy vấn (query) với các vector khóa (key), chia cho căn bậc hai của chiều dài vector, và áp dụng hàm softmax để chuẩn hóa.
* **Multi-Head Attention**: Sử dụng nhiều “đầu” attention độc lập để cho phép mô hình học được các khía cạnh khác nhau của thông tin ngữ nghĩa

Mô hình Transformer đã tạo ra bước đột phá lớn trong lĩnh vực xử lý ngôn ngữ tự nhiên nhờ vào khả năng xử lý song song và cơ chế attention mạnh mẽ. Nó đã trở thành nền tảng cho nhiều tiến bộ trong NLP và tiếp tục là một công cụ quan trọng cho các nghiên cứu và ứng dụng tương lai. Cho phép phát triển các mô hình xử lý ngôn ngữ quy mô lớn, chẳng hạn như GPT-3.

* Cải thiện đáng kể hiệu suất và hiệu quả các các tác vụ xử lý ngôn ngữ tự nhiên, ví dụ như dịch máy và mô hình hóa ngôn ngữ.
* Mở đường cho sự phát triển của các mô hình học sâu tiên tiến khác như Switch Transformer, có thể mở rộng thành các mô hình có nhiều tham số với độ chính xác cao và hiện quả hơn.

Transformer đã được sử dụng trong nhiều ứng dụng khác nhau, bao gồm chatbot, trợ lý ảo và hệ thống trả lời câu hỏi.



Hình 2.11: Ứng dụng của mô hình Transformer trên nhiều lĩnh vực khác nhau.

*Xử lý ngôn ngữ tự nhiên và GPT-3 (2020)*

Vào năm 2020, GPT-3 của OpenAI mở ra một chương mới của lĩnh vực trí tuệ nhân tạo khi đã thể hiện tiềm năng của xử lý ngôn ngữ tự nhiên (NLP) trong học máy, cho phép tạo và hiểu ngôn ngữ nâng cao, trong đó có cả ngôn ngữ lập trình. NLP đề cập đến khả năng hiểu và phân tích ngôn ngữ của con người khi nó được nói và viết, cho phép máy móc hiểu và tạo ra ngôn ngữ giống con người. Các khả năng NLP tiên tiến của nó có khả năng cách mạng hóa cách máy móc tương tác với con người, dẫn đến các ứng dụng học máy hiệu quả và chính xác hơn.

*GPT-4, Claude AI và Machine Learning hiện đại (2023)*

Những tiến bộ gần đây trong lĩnh vực học máy bao gồm GPT-4 và Claude AI. Chúng liên tục thách thức giới hạn của những gì có thể làm được trong lĩnh vực trí tuệ nhân tạo. Những tiến bộ mới trong ML cũng dẫn đến các nền tảng tạo hình ảnh, đặc biệt là Midjourney và DALL-E.

**GPT-4** , do OpenAI tạo ra, là một mô hình đa phương thức lớn có thể chấp nhận cả đầu vào văn bản và hình ảnh, đồng thời tạo ra đầu ra văn bản giống với ngôn ngữ của con người. Trí tưởng tượng và khả năng hợp tác của nó khiến nó trở thành một công cụ mạnh mẽ cho các ứng dụng học máy. Rất nhiều ứng dụng phụ thuộc vào GPT-4, bao gồm [Bing AI](https://bing.com/?/ai), [Duolingo](https://duolingo.com/), tác giả AI như [Jasper](https://jasper.ai/) , nền tảng phân tích dữ liệu như [Akkio](https://akkio.com/) và mạng xã hội như Snapchat. [Microsoft đã đầu tư 10 tỷ USD vào OpenAI để phát triển hơn](https://www.nytimes.com/2023/01/23/business/microsoft-chatgpt-artificial-intelligence.html) nữa

**Claude AI** là một nền tảng học sâu tạo điều kiện cho việc phát triển và triển khai nhanh chóng các ứng dụng AI, cung cấp nhiều tính năng, bao gồm xử lý ngôn ngữ tự nhiên, thị giác máy tính và học tăng cường.

[**Midjourney**](https://www.midjourney.com/) là một phòng thí nghiệm nghiên cứu độc lập cung cấp nền tảng cho các công cụ sáng tạo được hỗ trợ bởi AI để tạo hình ảnh trực quan với giao diện ngôn ngữ tự nhiên. Trình tạo hình ảnh AI của Midjourney có thể tạo hình ảnh tùy chỉnh cho các chiến dịch tiếp thị và quảng cáo, chẳng hạn như bài đăng trên mạng xã hội, quảng cáo biểu ngữ và tiếp thị qua email. Midjourney đã sử dụng [**bộ tăng tốc AI được phát triển tùy chỉnh của Google Cloud**](https://www.prnewswire.com/news-releases/midjourney-selects-google-cloud-to-power-ai-generated-creative-platform-301771558.html) , Bộ xử lý Tensor (TPU), để huấn luyện mô hình AI thế hệ thứ tư, có thể hiển thị các hình ảnh được tạo trên nền tảng với tốc độ ngoạn mục. Nền tảng sáng tạo do AI tạo ra của Midjourney đã được hơn **11 triệu thành viên** sử dụng và công ty gần đây đã [hợp tác với Google Cloud để hỗ trợ nền tảng của mình](https://www.prnewswire.com/news-releases/midjourney-selects-google-cloud-to-power-ai-generated-creative-platform-301771558.html) .



Hình 2.12: Một số hình ảnh được tạo ra từ Midjourney.

[**DALL-E**](https://www.axelerant.com/blog/how-to-use-ai-and-ml-to-enhance-dxps) là mô hình AI được phát triển bởi OpenAI, có thể **tạo ra hình ảnh từ các mô tả văn bản** . DALL-E có thể tạo ra hình ảnh của những vật thể không tồn tại trong thế giới thực, chẳng hạn như con ốc sên làm bằng dây đàn hạc hay chiếc ấm trà có hình dạng giống con ốc sên. Tác động của DALL-E đối với lĩnh vực công cụ sáng tạo do AI tạo ra là rất đáng kể và nó đã thu hút sự chú ý của các nhà nghiên cứu, nghệ sĩ cũng như nhà thiết kế.

Những tiến bộ liên tục trong học máy, chẳng hạn như GPT-4 và Claude AI, đang chuyển đổi lĩnh vực này và tạo điều kiện cho các ứng dụng mới và sáng tạo trong nhiều ngành công nghiệp. Khi học máy tiếp tục phát triển và thích ứng, nó được kỳ vọng sẽ đóng vai trò ngày càng quan trọng trong **việc định hình tương lai của công nghệ và xã hội** , chuyển đổi các ngành công nghiệp và xác định lại sự tương tác giữa con người với máy tính

2.1.5. Xu hướng phát triển.

Trong những năm tới, Machine Learning (ML) dự kiến sẽ trải qua một giai đoạn phát triển mạnh mẽ và đa dạng, mang lại những thay đổi quan trọng trong cách chúng ta tận dụng công nghệ để giải quyết các vấn đề phức tạp. Xu hướng phát triển của ML không chỉ tập trung vào việc cải thiện hiệu suất và độ chính xác của các mô hình, mà còn hướng tới việc mở rộng phạm vi ứng dụng và tăng cường khả năng tiếp cận của công nghệ này.

Một trong những xu hướng nổi bật nhất là sự tiến bộ liên tục của Deep Learning. Các kiến trúc mạng neural ngày càng trở nên phức tạp và tinh vi hơn, cho phép xử lý các tác vụ phức tạp với độ chính xác cao hơn. Chúng ta có thể kỳ vọng vào sự xuất hiện của các mô hình ngôn ngữ lớn tiên tiến hơn, có khả năng hiểu và tạo ra ngôn ngữ tự nhiên gần với con người hơn. Trong lĩnh vực thị giác máy tính, các mô hình deep learning sẽ tiếp tục cải thiện khả năng nhận dạng và phân tích hình ảnh, mở ra những ứng dụng mới trong y tế, an ninh, và tự động hóa.

Song song với đó, Reinforcement Learning (RL) đang được kỳ vọng sẽ đạt được những bước tiến quan trọng. RL có tiềm năng to lớn trong việc đào tạo các hệ thống tự động học hỏi và thích nghi với môi trường phức tạp. Trong lĩnh vực robotics, RL có thể giúp các robot thực hiện các nhiệm vụ phức tạp và linh hoạt hơn trong môi trường thực tế. Ngoài ra, RL cũng đang được áp dụng rộng rãi trong tối ưu hóa quyết định, từ quản lý chuỗi cung ứng đến giao dịch tài chính tự động.

Xu hướng học liên tục (Continual Learning) đang nổi lên như một giải pháp cho vấn đề "quên lãng thảm họa" của các mô hình ML. Kỹ thuật này cho phép các mô hình tiếp tục học và thích nghi với dữ liệu mới mà không làm mất đi kiến thức đã học trước đó. Điều này đặc biệt quan trọng trong các ứng dụng thực tế, nơi dữ liệu và môi trường liên tục thay đổi. Continual Learning hứa hẹn sẽ tạo ra các hệ thống AI linh hoạt hơn, có khả năng thích ứng tốt hơn với thế giới thực.

Bên cạnh đó, kỹ thuật học ít mẫu (Few-shot Learning) đang được phát triển mạnh mẽ. Xu hướng này nhắm đến việc tạo ra các mô hình có khả năng học hiệu quả từ lượng dữ liệu huấn luyện hạn chế. Điều này đặc biệt quan trọng trong các lĩnh vực mà việc thu thập dữ liệu là khó khăn hoặc tốn kém, như y học cá nhân hóa hoặc nhận dạng các loài động vật quý hiếm. Few-shot Learning có thể mở rộng đáng kể phạm vi ứng dụng của ML trong các lĩnh vực có dữ liệu hạn chế.

Xu hướng AutoML (Automated Machine Learning) đang dần trở nên phổ biến, hướng tới việc tự động hóa quá trình thiết kế và tối ưu hóa mô hình ML. AutoML không chỉ giúp giảm thời gian và công sức cần thiết để phát triển các mô hình ML hiệu quả, mà còn góp phần democratize ML bằng cách mở rộng khả năng tiếp cận công nghệ này cho nhiều đối tượng hơn, kể cả những người không có chuyên môn sâu về ML.

Cuối cùng, xu hướng hướng tới AI có thể giải thích được (Explainable AI) và công bằng trong ML đang ngày càng được chú trọng. Khi các hệ thống ML ngày càng được áp dụng rộng rãi trong các quyết định quan trọng, việc hiểu được cách thức hoạt động của chúng và đảm bảo tính công bằng trở nên cực kỳ quan trọng. Các nhà nghiên cứu đang phát triển các kỹ thuật mới để làm cho các quyết định của mô hình ML trở nên minh bạch và dễ hiểu hơn, đồng thời phát triển các phương pháp để phát hiện và giảm thiểu thiên kiến trong các mô hình.