

MACHINE LEARNING

Nguyễn Văn Đạt – 51800364
2/5/2021

1. Giới thiệu tổng quát:

Machine Learning (Học Máy) là một lĩnh vực nhỏ của Khoa Học Máy Tính, nó có khả năng tự học hỏi dựa trên dữ liệu đưa vào mà không cần phải được lập trình cụ thể. Trong những năm gần đây machine learning thu hút rất nhiều người quan tâm, từ các công ty lớn như Google, Facebook, Microsoft đến các công ty khởi nghiệp đều đầu tư vào machine learning. Hàng loạt các ứng dụng sử dụng machine learning ra đời trên mọi lĩnh vực của cuộc sống, tiêu biểu như xe tự vận hành của Google và Tesla, hệ thống tag khuôn mặt trong ảnh của Facebook, trợ lý ảo Siri của Apple, hệ thống gợi ý sản phẩm của Amazon, hệ thống gợi ý phim của Netflix, máy chơi cờ vây AlphaGo của Google DeepMind, ..., ngoài ra còn nhiều ứng dụng khác bao trùm từ khoa học máy tính đến những ngành tưởng như không liên quan là vật lý, hóa học, y học, chính trị.^[1]

Machine learning có 3 mối quan hệ sau:

- Machine learning và Big data.
- Machine learning và trí tuệ nhân tạo.
- Machine learning và dự đoán tương lai.

1.1. Big data:

Big data là một thuật ngữ một tập hợp dữ liệu rất lớn và phức tạp mà các ứng dụng xử lý dữ liệu truyền thống sẽ khó xử lý được hoặc mất rất nhiều thời gian. Những tập dữ liệu lớn này có thể bao gồm các dữ liệu có cấu trúc, không có cấu trúc và bán cấu trúc.^[2]

Big Data thực chất không phải là một ngành khoa học chính thống. Đó là một cụm từ để nêu nên thời kỳ bùng nổ của dữ liệu số hiện nay. Big Data là một kết quả tất yếu của việc mạng Internet ngày càng phát triển. Machine Learning và Big Data có một quan hệ tương hỗ với nhau: machine learning phát triển hơn nhờ sự gia tăng của khối lượng dữ liệu của Big Data; ngược lại, giá trị của Big Data phụ thuộc vào khả năng khai thác tri thức từ dữ liệu của machine learning.



Hình 1. Minh họa Big data

1.2. Artificial Intelligence/AI

Trí tuệ nhân tạo (AI) đề cập đến việc mô phỏng trí thông minh của con người trong máy móc được lập trình để suy nghĩ giống như con người và bắt chước hành động của họ. Thuật ngữ này cũng có thể được áp dụng cho bất kỳ máy nào thể hiện các đặc điểm liên quan đến trí óc con người như học tập và giải quyết vấn đề.

Machine learning là một phương tiện được kỳ vọng sẽ giúp con người đạt được mục tiêu chinh phục AI. Machine learning và AI có mối quan hệ chặt chẽ với nhau nhưng không hẳn là trùng khớp vì một bên là mục tiêu (AI), một bên là phương tiện (machine learning). Chinh phục AI mặc dù vẫn là mục đích cao cấp của machine learning, nhưng hiện tại machine learning chỉ tập trung vào hai mục tiêu:

- Trợ giúp con người trong việc xử lý một khối lượng thông tin khổng lồ mà chúng ta phải đối mặt hàng ngày - Big Data.
- Làm cho máy tính có những khả năng nhận thức cơ bản của con người (ý tưởng robot).

1.3. Dự đoán tương lai:

Dự đoán luôn đi liền với machine learning, khi mà các hệ thống tiên đoán và mô phỏng luôn trợ giúp rất nhiều trong việc giải quyết các vấn đề của machine learning.

2. Các thuật toán machine learning:

Machine learning có 2 thuật toán cơ bản: Học có giám sát (Supervised learning) và Học không giám sát (Unsupervised learning). Sự khác nhau giữa hai thuật toán này là cách cung cấp tập dữ liệu cho mô hình, cách thuật toán sử dụng dữ liệu và loại vấn đề mà chúng giải quyết.

2.1. Support Vector Machines:

Trong học máy, máy vector hỗ trợ (SVM, cũng là mạng vector hỗ trợ [1]) là mô hình học có giám sát với các thuật toán học liên quan phân tích dữ liệu để phân loại và phân tích hồi quy.^[3]

2.2. Mô hình xác suất (Probabilistic Models):

Đây là mô hình cố gắng giải quyết bài toán bằng phân bố xác suất. Điểm mạnh của mô hình xác suất là đơn giản nhưng hiệu quả. Đầu ra của nó không chỉ là label mà còn đi kèm xác suất thể hiện độ chính xác cho kết quả đó.^[4]

2.3. Deep learning:

Là một phần của họ các phương pháp học máy rộng hơn dựa trên mạng nơ-ron nhân tạo với học đại diện.^[5]

3. Ứng dụng machine learning trong thực tiễn:

*Dưới đây chỉ là một phần trong vô vàn ứng dụng nổi bật.

3.1. Lòng tiếng cho phim câm:

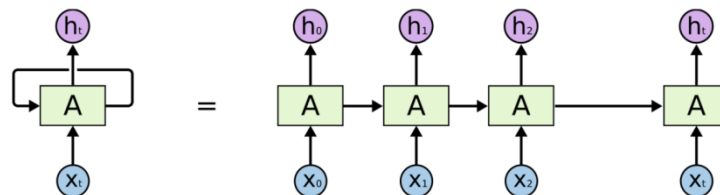
Ứng dụng này sử dụng cả Convolutional Neural Network và Recurrent Neural Network, trong đó Convolutional Neural Network được dùng để nhận dạng cảnh phim, sau đó Recurrent Neural Network được sử dụng để tạo ra chuỗi âm thanh (tiếng trống) phù hợp với cảnh phim vừa được nhận diện. Khi một chiếc máy tính sử dụng cả hai thuật toán trên và một chuyên gia con người cùng lòng tiếng cho một cảnh phim, không ai có thể phân biệt được âm thanh được tạo ra là do máy hay do người (bài kiểm tra này gọi là Turing test, bài kiểm tra độ thông minh của trí tuệ nhân tạo).

3.2. Nhận diện chữ viết tay

Sử dụng Recurrent Neural Network, thuật toán có khả năng học cách viết chữ từ nét chữ của người. Sau đó, thuật toán có thể viết ra bất kỳ dòng chữ nào theo yêu cầu sử dụng nét viết tương tự.

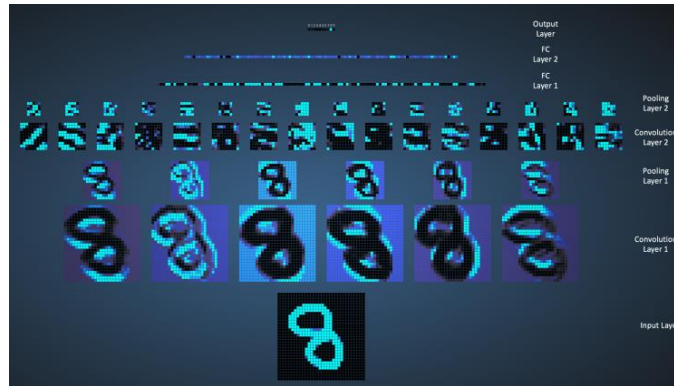
4. Một số học thuật:

Recurrent Neural Network: là một dạng Neural Network tái diễn, được sử dụng để dự đoán các dữ liệu dạng chuỗi, ví dụ như dự đoán từ tiếp theo trong một câu, hoặc dự đoán thời tiết của ngày hôm sau. ^[6]



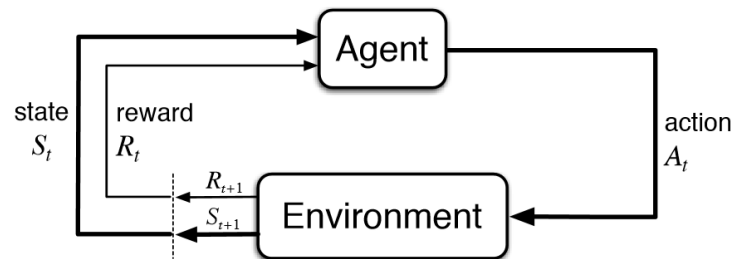
Hình 2. Minh họa Recurrent Neural Network

Convolutional Neural Network: là một dạng Neural Network trong đó các lớp neuron đầu tiên có chức năng tìm tập hợp các điểm đặc trưng (feature) của hình ảnh, và sử dụng các đặc điểm (feature) này để các lớp neuron phía sau đưa ra dự đoán.



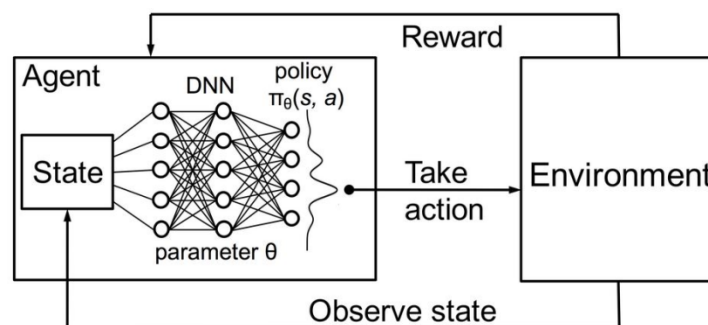
Hình 3. Minh họa Convolutional Neural Network

Reinforcement Learning: Thuật toán này lấy cảm hứng từ quá trình học hỏi theo cơ chế thưởng phạt ở con người. Reinforcement learning học và tích lũy kinh nghiệm để đưa ra hành động (action) sao cho với trạng thái môi trường hiện tại (state), hành động sẽ tối đa hóa phần thưởng (reward) nhận được.



Hình 4. Minh họa Reinforcement Learning

Deep Reinforcement Learning: “Học sâu và tăng cường”. Đây là một kết hợp mạnh mẽ giữa Deep Learning và Reinforcement Learning, trong đó Neural Network tham gia vào quá trình quyết định hành động của Reinforcement Learning.



Hình 5. Minh họa Deep Reinforcement Learning

[1] : <URL : <https://machinelearningcoban.com/2016/12/26/introduce/>>

[2]: <URL:

https://vi.wikipedia.org/wiki/Tr%C3%ADtu%E1%BB%87_nh%C3%A2n_t%E1%BA%A1o>

[3]: <URL: https://en.wikipedia.org/wiki/Support-vector_machine>

[4]: <URL: https://en.wikipedia.org/wiki/Category:Probabilistic_models>

[5]: <URL: https://en.wikipedia.org/wiki/Deep_learning>

[6]: <URL: <https://techinsight.com.vn/ung-dung-thuc-tien-cua-machine-learning/>>