

MẠNG NEURAL NHÂN TẠO và GIẢI THUẬT DI TRUYỀN

Neural Network & Genetic Algorithm



Biên soạn: ThS. Phạm Đình Tài
pdtai@ntt.edu.vn
0985.73.39.39

CHƯƠNG 3

Mạng Perceptron



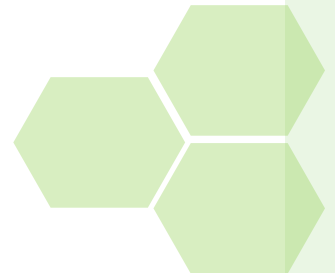
- ✓ **Mô tả giải thuật huấn luyện mạng**
- ✓ **Định nghĩa thể nào là luật học**
- ✓ **Phát triển các luật học cho mạng**
- ✓ **Thảo luận về ưu nhược điểm của mạng Perceptron một lớp**

GIỚI THIỆU CHUNG

- Năm 1943, hai nhà khoa học **Warren McCulloch** and **Walter Pitts** đã giới thiệu mô hình neural nhân tạo lần đầu tiên
- Đặc trưng cơ bản của mô hình neuron là
 - Tổng trọng số của tín hiệu đầu vào được so sánh với một giá trị ngưỡng
 - Nếu tổng lớn hơn hoặc bằng ngưỡng thì đầu ra là 1
 - Nếu tổng nhỏ hơn ngưỡng thì đầu ra là 0
- Họ cho rằng:
 - Bất kỳ một mạng các neural như vậy đều có thể tạo ra bất kỳ một hàm toán hay logic
 - Không giống neu neural sinh học, các tham số của mạng phải được thiết kế vì không có phương pháp học nào

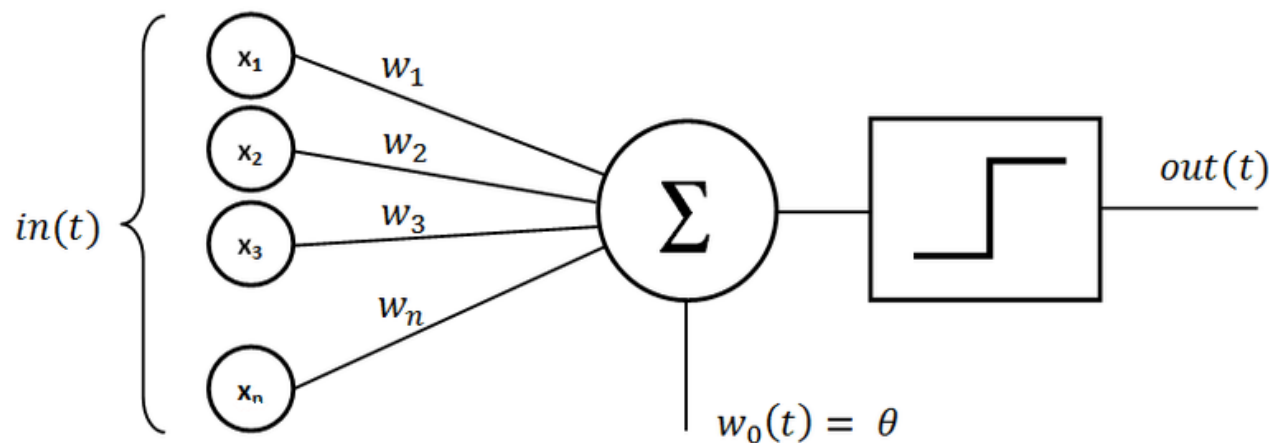
GIỚI THIỆU CHUNG

- Năm 1950, **Frank Rosenblatt** và một số nhà khoa học đã phát triển một lớp các mạng neuron gọi là ***Perceptron***
- Các neuron trong mạng này giống như các neural của Warren McCulloch and Walter Pitts
- Đóng góp chính
 - Đưa vào luật học để huấn luyện các mạng perceptron
 - Đã chứng minh rằng với luật học như vậy có thể hội tụ về một tập các tham số đúng nếu như tồn tại các trọng số như thể của mạng.
 - Luật học đơn giản, tự động
 - Có thể học với các bộ tham số khởi tạo ngẫu nhiên



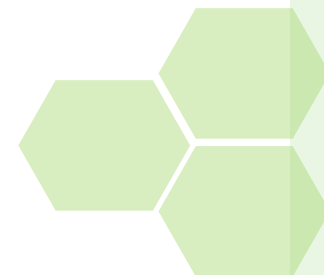
GIỚI THIỆU CHUNG

- Tuy nhiên mạng perceptron có một số hạn chế
 - *Không có khả năng cài đặt được một số hàm cơ bản*
 - *Nó được giải quyết với khi được thiết kế thành nhiều lớp*
- Ngày nay mạng **Perceptron** vẫn là một mạng có ý nghĩa quan trọng vì nó nhanh và tin cậy
- Việc hiểu các thao tác của mạng **Perceptron** sẽ giúp hiểu các mạng phức tạp hơn



LUẬT HỌC

- Là một thủ tục để thay đổi các trọng số và bias của mạng (giải thuật huấn luyện mạng)
- Có nhiều luật học khác nhau
 - *Học có giám sát*
 - *Học không có giám sát*
 - *Học tăng cường*

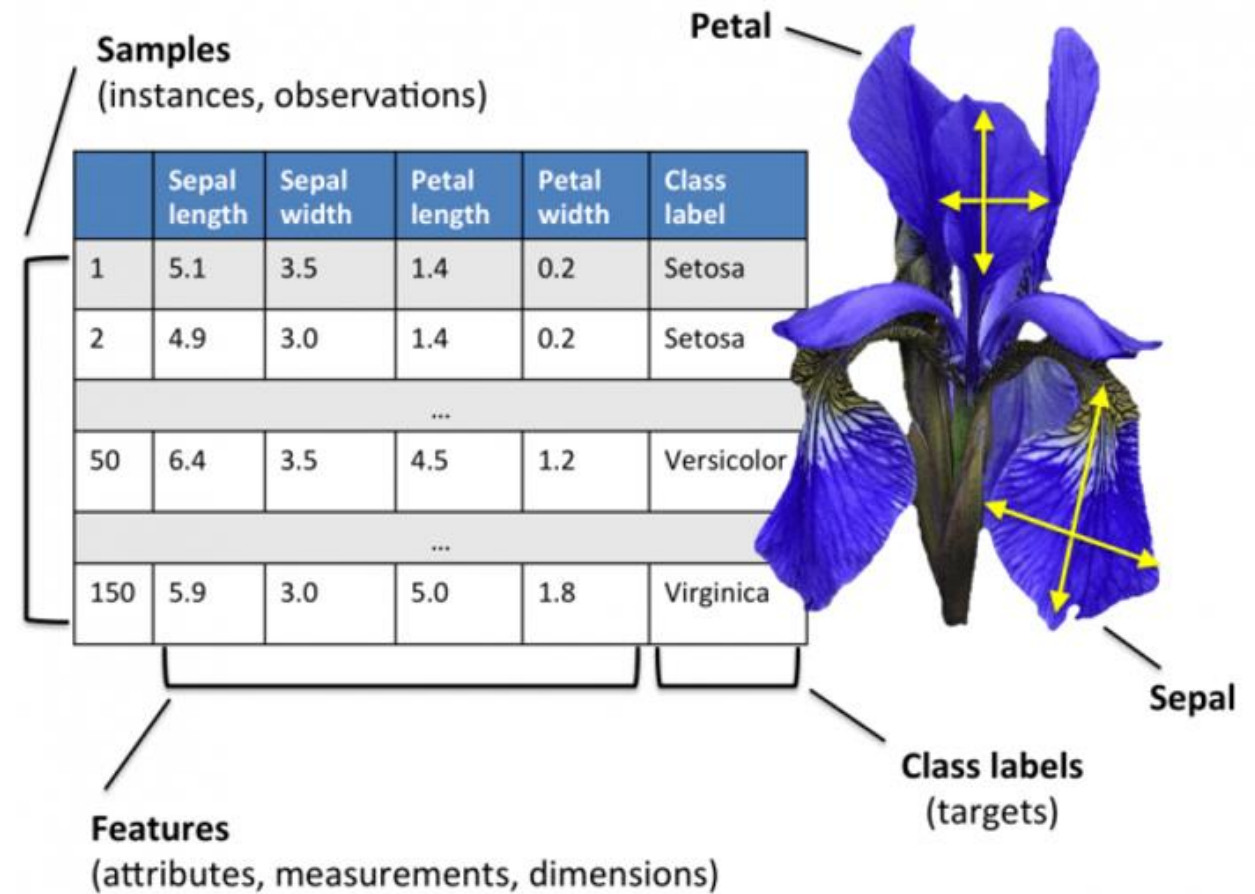


1. HỌC CÓ GIÁM SÁT

- Mô hình học có giám sát (**Supervised learning**) là mô hình học trên dữ liệu có dẫn nhãn, tức là mục tiêu của bài toán machine learning cần học đã được gán nhãn sẵn trong dữ liệu huấn luyện.
- Dữ liệu đầu vào của quá trình học bao gồm cả vector đầu vào chứa các thuộc tính của dữ liệu lẫn giá trị đầu ra mục tiêu (gọi là nhãn của dữ liệu).
- Học có giám sát cho phép dự đoán đầu ra (outcome) của một dữ liệu mới dựa trên các cặp (đầu vào, đầu ra) đã biết từ trước thu được từ bộ dữ liệu huấn luyện.

1. HỌC CÓ GIÁM SÁT

- Bộ dữ liệu huấn luyện bao gồm các cặp (data, label), tức (dữ liệu, nhãn).
- Ví dụ, bộ dữ liệu hoa tử đằng chứa các thuộc tính là chiều dài và chiều rộng của cánh hoa và đài hoa, các thuộc tính này tạo thành dữ liệu đầu vào (data). Đồng thời, nó cũng chứa cả nhãn (class label) của mục tiêu dự đoán (dòng hoa là một trong 3 loại: setosa, versicolor và virginica).

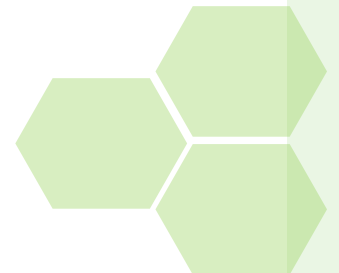


1. HỌC CÓ GIÁM SÁT

- Được cung cấp một tập dữ liệu thể hiện đáp ứng của mạng với các dữ liệu đầu vào

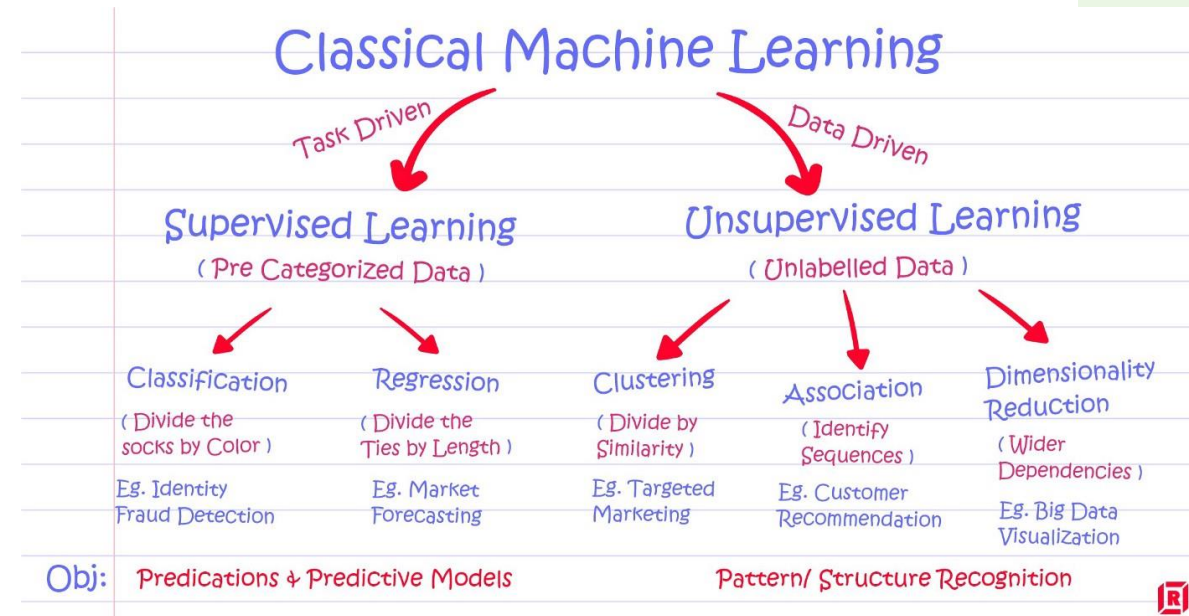
$$\{\mathbf{p}_1, \mathbf{t}_1\}, \{\mathbf{p}_2, \mathbf{t}_2\}, \dots, \{\mathbf{p}_Q, \mathbf{t}_Q\}$$

- Trong đó:
 - \mathbf{p}_i là đầu vào, \mathbf{t}_i là **target** đầu ra
 - Do đầu vào được đưa vào mạng, **đáp ứng** của mạng được so sánh với đầu ra \mathbf{t}_i
- **Luật học**: làm thay đổi trọng số và **bias** của mạng để cho đáp ứng đầu ra gần với **target** nhất



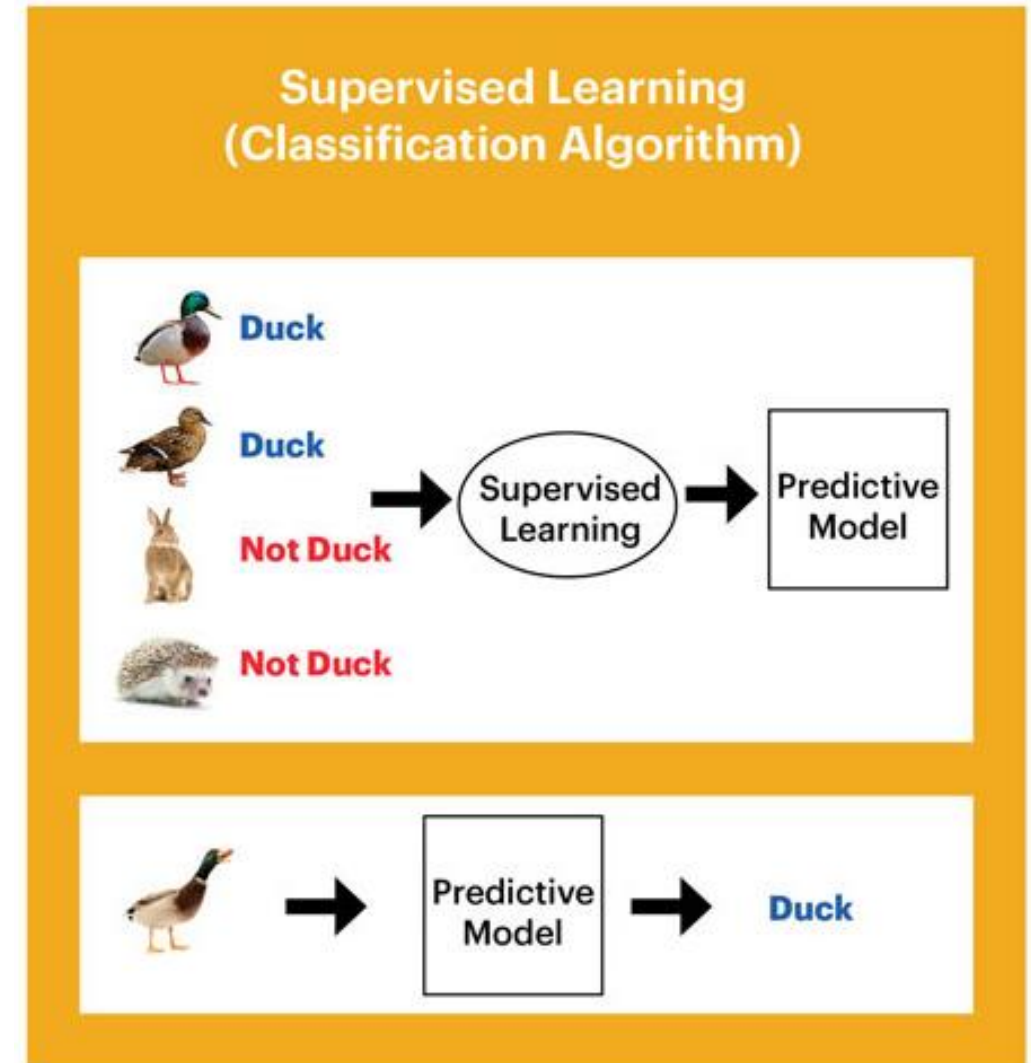
HỌC CÓ GIÁM SÁT

- Hai nhóm bài toán cơ bản trong học có giám sát là **classification** (phân loại) và **regression** (hồi quy), trong đó biến đầu ra của bài toán phân loại có các giá trị rời rạc trong khi biến đầu ra của bài toán hồi quy có các giá trị liên tục.



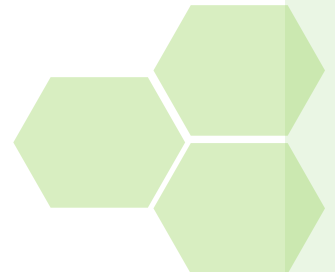
■ **Classification** (phân loại)

- VD Bài toán phân lớp hoa tử đằng ở trên là một ví dụ về bài toán phân lớp. Trong đó mô hình nhận đầu vào là kích thước cánh hoa và nhụy hoa và cho đầu ra là loại hoa tử đằng (setosa, versicolor hoặc virginica).
- Việc phân loại diễn ra khi biến đầu ra là một thể loại nào đó, chẳng hạn như "đỏ" hoặc "xanh" hoặc "Giống Duck" và "không Duck".



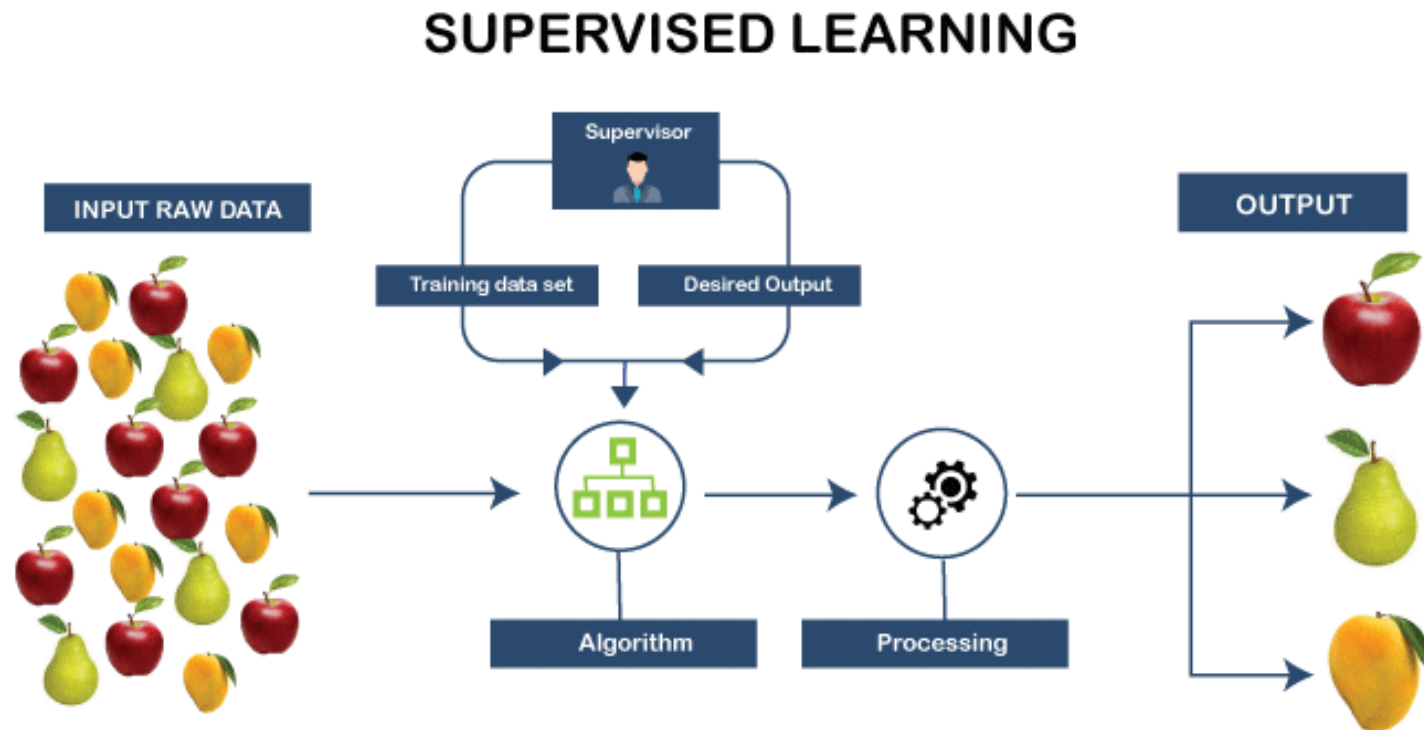
■ **Regression** (hồi quy)

- Việc hồi quy xảy ra là khi biến đầu ra là một giá trị thực, chẳng hạn như “đô la” hay “trọng lượng”.
- Ví dụ: thông qua dự đoán mức độ hài lòng với công việc (thang điểm từ 0-100) theo mức độ thu nhập của nhân viên. Hoặc bài toán dự đoán giá nhà ở theo rất nhiều các thuộc tính số đầu vào như diện tích, thuế,...
- Một số thuật toán được thiết kế chỉ dùng cho bài toán hồi quy (***hồi quy tuyến tính***) hoặc phân lớp (***hồi quy logistic***), nhưng cũng có một số thuật toán có thể sử dụng được cho cả hai bài toán với một chút điều chỉnh nhỏ (mạng neural)



HỌC CÓ GIÁM SÁT

- Với Supervised Learning, bên cạnh xây dựng các mô hình mạnh, việc thu thập và gán nhãn dữ liệu tốt và hợp lý cũng đóng vai trò then chốt để giải quyết các bài toán trong thực tế.

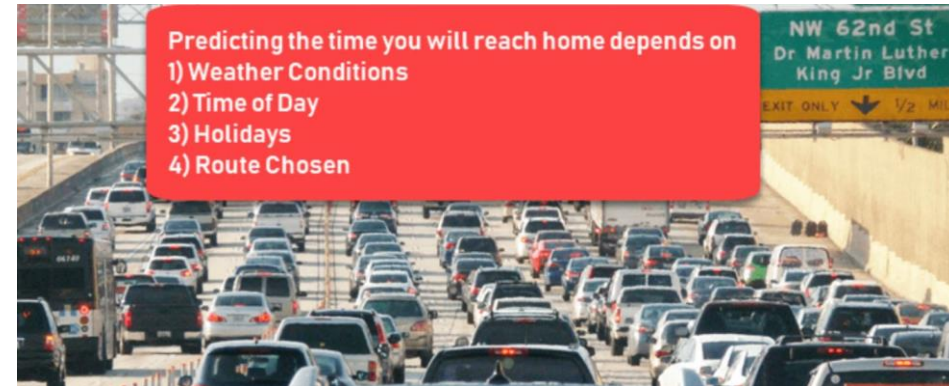


HỌC CÓ GIÁM SÁT

- Theo bản năng, chúng ta biết rằng nếu trời mưa, thì ta sẽ mất nhiều thời gian hơn để lái xe về nhà. Nhưng máy thì sẽ cần dữ liệu và số liệu thống kê.
- Hãy xem làm thế nào ta có thể phát triển một mô hình học tập có giám sát của ví dụ này để giúp người dùng xác định thời gian đi làm.
- Đầu tiên bạn cần tạo là một bộ huấn luyện. Tập huấn luyện này sẽ chứa tổng thời gian đi lại và các yếu tố tương ứng như thời tiết, thời gian, v.v.
- Dựa trên tập huấn luyện này, máy của bạn có thể thấy có mối quan hệ trực tiếp giữa lượng mưa và thời gian bạn sẽ về nhà.

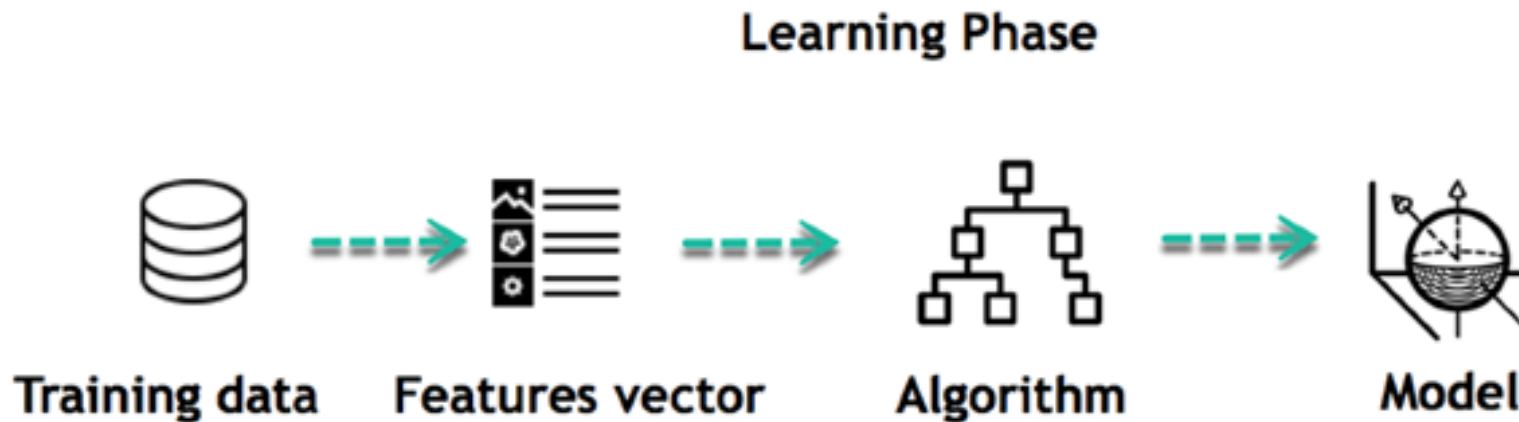
HỌC CÓ GIÁM SÁT

- **Ví dụ:** Bạn muốn đào tạo một cỗ máy để giúp bạn dự đoán bạn sẽ mất bao lâu để lái xe về nhà từ nơi làm việc. Tại đây, bạn bắt đầu bằng cách tạo một tập hợp dữ liệu được dán nhãn. Dữ liệu này bao gồm:
 - *Điều kiện thời tiết*
 - *Thời gian trong ngày*
 - *Ngày lễ*
 - *Chọn đường đi*
- Tất cả những chi tiết này là đầu vào. Đầu ra là lượng thời gian cần thiết để lái xe trở về nhà vào ngày cụ thể đó.



HỌC CÓ GIÁM SÁT

- → nó càng khẳng định rằng trời càng mưa, bạn sẽ lái xe càng lâu để trở về nhà. Nó cũng có thể thấy kết nối giữa thời gian bạn nghỉ làm và thời gian bạn sẽ đi trên đường.
- Bạn càng gần 6 giờ tối. bạn càng mất nhiều thời gian để về nhà. Máy của bạn có thể tìm thấy một số mối quan hệ với dữ liệu được dán nhãn của bạn.



2. HỌC TĂNG CƯỜNG

- Tương tự như học có giám sát
- Tuy nhiên thay vì đưa ra target, cung cấp một giá trị score thể hiện hiệu năng của mạng với đầu vào cho trước
- Học này kém thông dụng hơn học giám sát, thường được áp dụng trong các hệ thống điều khiển

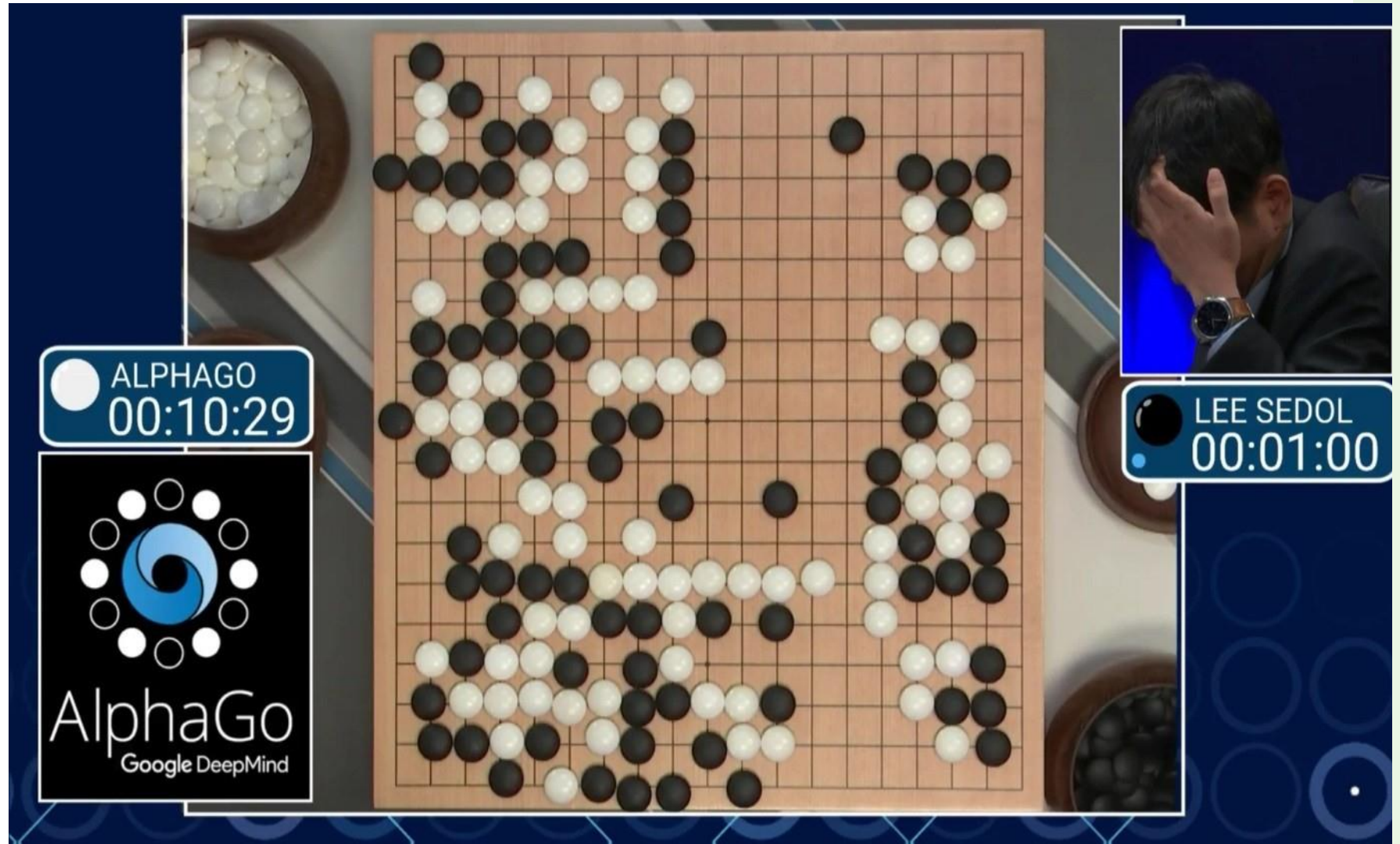
HỌC TĂNG CƯỜNG

■ Ví dụ:

https://vi.wikipedia.org/wiki/AlphaGo_o_%C4%91%E1%BA%A5u_v%E1%BB%9Bi_Lee_Sedol

Cờ vây có độ phức tạp cực kỳ cao

<https://www.youtube.com/watch?v=54uqCtM3fOQ&t=70s>



HỌC TĂNG CƯỜNG

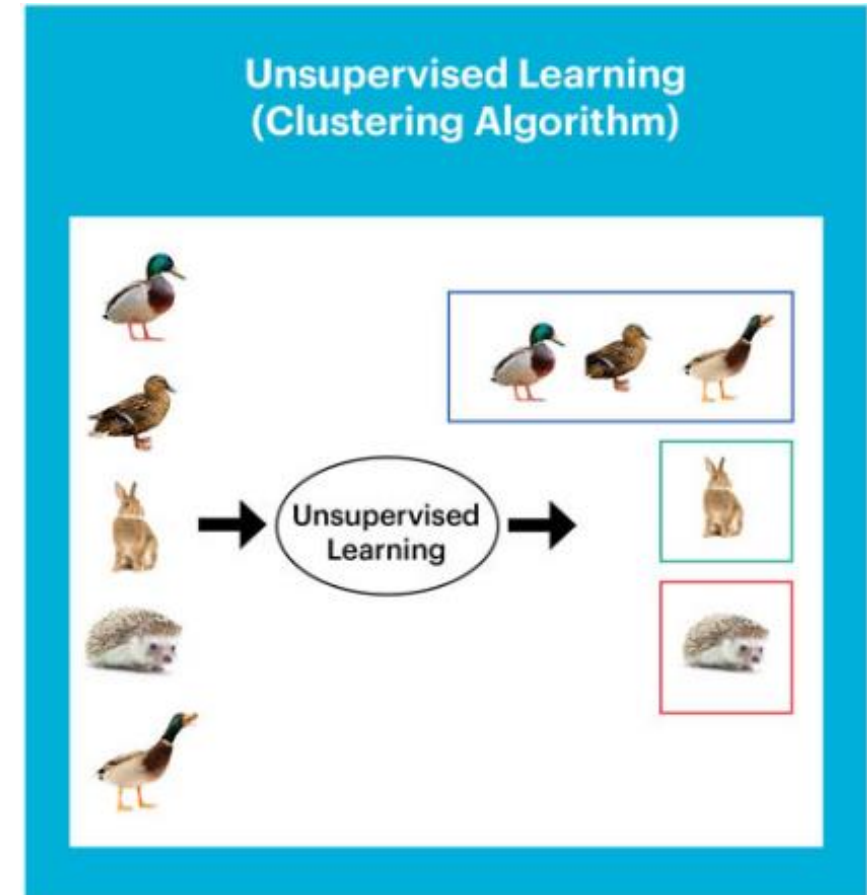
- Cờ vây có độ phức tạp cực kỳ cao
- Tổng số nước đi : 10^{761} cờ vua (10^{120})
- Không thể áp dụng **IBM DeepBlue** (thuật toán thắng người trong môn cờ vua 20 năm trước đây)
- AlphaGo:
 - **Supervised**: dữ liệu từ ván cờ do con người chơi với nhau được đưa vào huấn luyện
 - **Reinforcement**: tự chơi với chính nó để tìm ra nước đi mới

3. HỌC KHÔNG GIÁM SÁT

- Học không giám sát (**Unsupervised learning**) là nơi bạn chỉ có dữ liệu đầu vào (X) và không có biến đầu ra tương ứng.
- Mục tiêu của việc học không giám sát là để mô hình hóa cấu trúc nền tảng hoặc sự phân bố trong dữ liệu để hiểu rõ hơn về nó.
- Đây được gọi là học không giám sát vì không giống như việc học có giám sát, không có câu trả lời đúng và không có vị “giáo viên” nào cả.
- Các thuật toán được tạo ra chỉ để khám phá và thể hiện các cấu trúc hữu ích bên trong dữ liệu.

3. HỌC KHÔNG GIÁM SÁT

- Các vấn đề học tập không giám sát có thể được phân ra thành hai việc chia nhóm và kết hợp.
- **Chia nhóm:** Vấn đề về chia nhóm là nơi muốn khám phá các nhóm vốn có bên trong dữ liệu, chẳng hạn như phân nhóm khách hàng theo hành vi mua hàng.
- **Kết hợp:** Vấn đề về học tập quy tắc kết hợp là nơi muốn khám phá các quy tắc mô tả dữ liệu của bạn, chẳng hạn như những người mua X cũng có khuynh hướng mua Y.



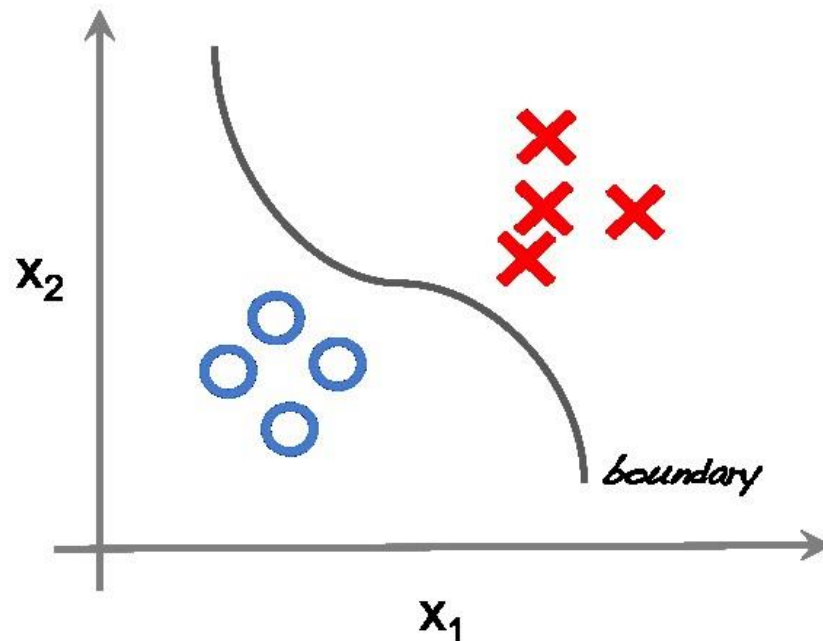
3. HỌC KHÔNG GIÁM SÁT

- Có rất nhiều bài toán trong mô hình học không giám sát, nổi bật nhất là hai bài toán: phân cụm và xấp xỉ phân phối xác suất. Bài toán phân cụm có nhiệm vụ tìm kiếm các nhóm có tương quan trong dữ liệu.
- ví dụ về thuật toán phân cụm là *k-Means*, trong đó *k* đại diện cho số cụm cần tìm trong dữ liệu. Một ví dụ điển hình của các thuật toán xấp xỉ phân phối là thuật toán xấp xỉ mật độ nhân (*Kernel Density Estimation*) trong đó sử dụng một nhóm nhỏ mẫu dữ liệu có liên quan chặt chẽ với nhau để tìm cách xấp xỉ phân phối cho các điểm mới trong không gian dữ liệu của bài toán. Các mô hình này có thể được sử dụng để học ra các mẫu dạng trong dữ liệu mà không cần gán nhãn trước.

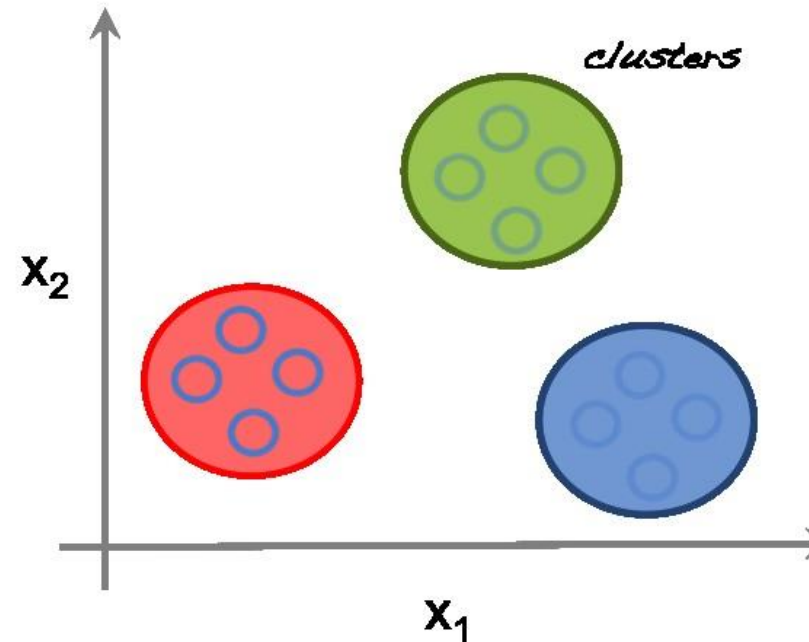
3. HỌC KHÔNG GIÁM SÁT

- Chỉ cung cấp đầu vào, không cung cấp đầu ra
- Mạng học để phân cụm dữ liệu

Supervised learning



Unsupervised learning

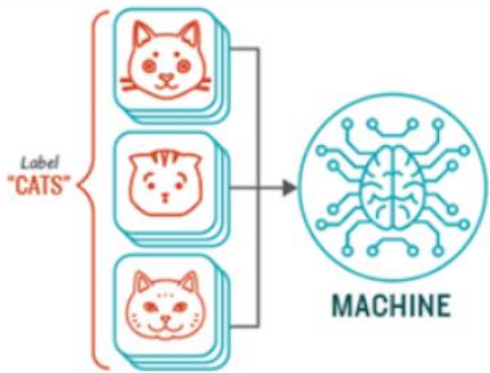


Supervised or unsupervised ?

Supervised

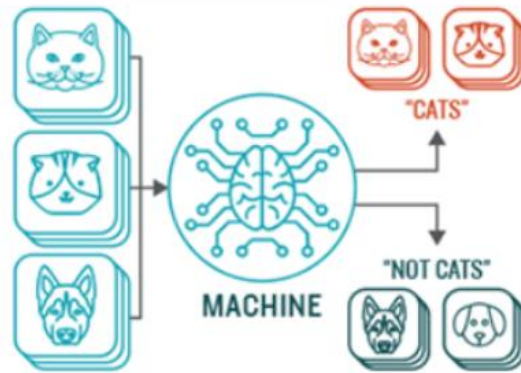
STEP 1

Nhận dữ liệu gán nhãn, tìm cách dự đoán nhãn



STEP 2

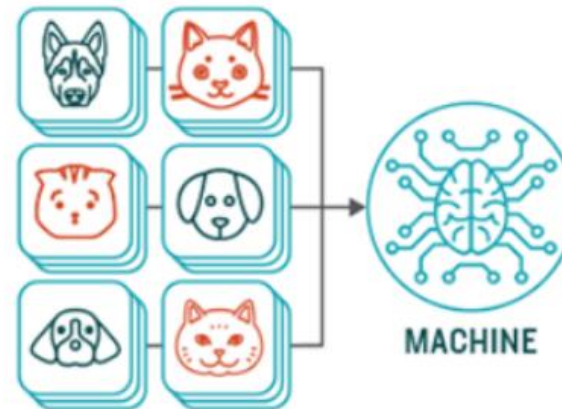
Dự đoán nhãn sao cho sai số dự đoán ít nhất



unsupervised

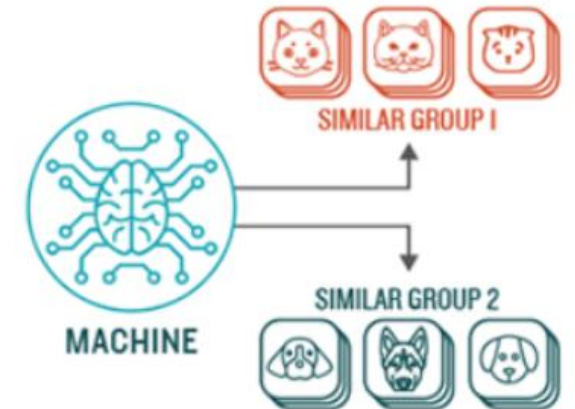
STEP 1

Nhận dữ liệu **không gán nhãn**, tìm **quan hệ** giữa các dữ liệu



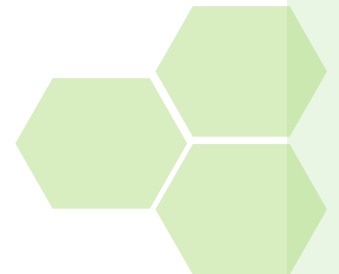
STEP 2

Tìm quan hệ ẩn trong dữ liệu



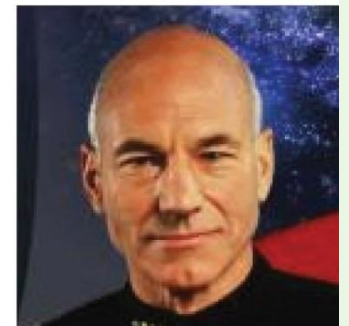
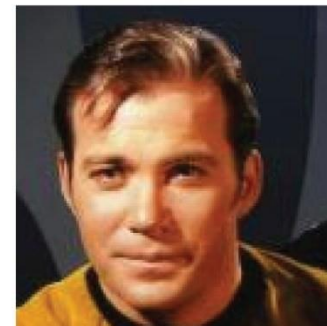
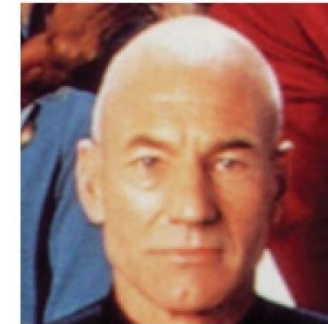
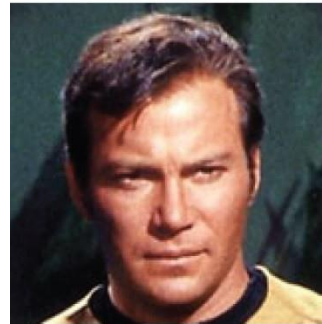
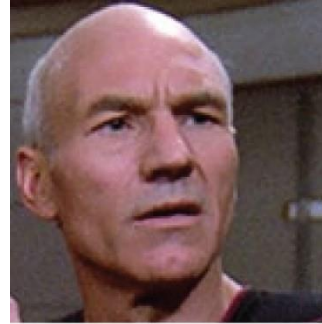
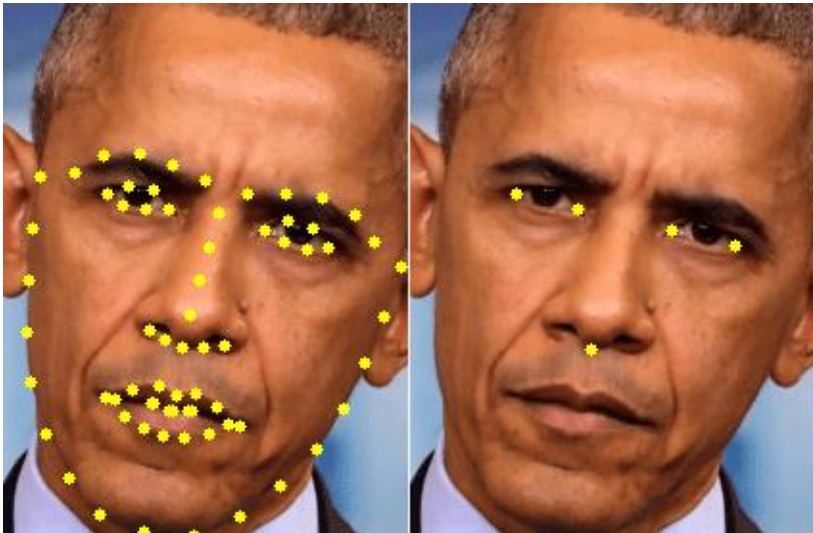
Supervised or unsupervised ?

- Đối với học có giám sát chúng ta biết trước tên của các nhãn là "mèo" hoặc "không phải mèo" trong khi trong học không giám sát, tên của các nhãn không có trong bộ dữ liệu, chúng ta chỉ tìm được quan hệ rằng: có một số ảnh giống nhau như thế nào.
- Ta không biết hai tập này là gì, và chỉ biết rằng chúng "giống nhau". Sau khi tìm được sự giống nhau này rồi, ta vẫn có thể gán nhãn cho dữ liệu tương tự như bài toán phân lớp. Tuy nhiên, trong nhiều trường hợp, điều này đòi hỏi khá nhiều kiến thức của chuyên gia.



Supervised or unsupervised ?

■ Face recognition





Supervised or unsupervised ?

Học có giám sát là nhiệm vụ của Machine Learning là học một hàm ánh xạ đầu vào thành đầu ra dựa trên các cặp đầu vào-đầu ra mẫu.

Học không giám sát là nhiệm vụ của Machine Learning trong việc suy ra một chức năng để mô tả cấu trúc ẩn từ dữ liệu không được gán nhãn.

Chức năng chính

Trong học tập có giám sát, mô hình dự đoán kết quả dựa trên dữ liệu đầu vào được gán nhãn.

Trong học tập không giám sát, mô hình dự đoán kết quả mà không có dữ liệu được gán nhãn bằng cách tự xác định các mẫu.

Độ chính xác của kết quả

Các kết quả được tạo ra từ các phương pháp học có giám sát là chính xác và đáng tin cậy hơn.

Kết quả được tạo ra từ các phương pháp học tập không giám sát không chính xác và đáng tin cậy.

Các thuật toán chính

Có các thuật toán để hồi quy và phân loại trong học tập có giám sát.

Có các thuật toán để phân cụm trong học tập không giám sát.

Thank you !

