

Kiến Thức Nền Tảng Machine Learning

1. Machine Learning (Học máy) là gì?

Machine Learning là một nhánh của Trí tuệ nhân tạo (AI), nơi máy tính tự học hỏi quy luật từ dữ liệu cũ để đưa ra các dự đoán hoặc quyết định cho dữ liệu mới mà không cần lập trình chi tiết từng bước.

2. Phân loại Machine Learning

Loại hình	Đặc điểm	Ví dụ sử dụng
Học có giám sát (Supervised)	Dữ liệu có nhãn/đáp án sẵn.	Dự báo giá nhà, nhận diện khuôn mặt.
Học không giám sát (Unsupervised)	Dữ liệu không nhãn, máy tự tìm nhóm.	Phân loại khách hàng, tìm điểm bất thường.
Học tăng cường (Reinforcement)	Học qua thử và sai để nhận thưởng.	Xe tự lái, chơi game (AlphaGo).

3. Phân biệt các loại Hồi quy (Regression)

- Hồi quy Tuyến tính Đa biến:** Tìm một **đường thẳng** kết nối nhiều yếu tố đầu vào để dự báo một con số liên tục. (Ví dụ: *Giá nhà dựa trên diện tích và vị trí*).
- Hồi quy Phi tuyến tính:** Dùng khi mối quan hệ là một **đường cong** phức tạp, không theo tỉ lệ thẳng. (Ví dụ: *Tốc độ tăng trưởng vi khuẩn*).
- Hồi quy Logistic:** Thực chất dùng để **phân loại** (xác suất 0 hoặc 1). Kết quả trả về là lựa chọn. (Ví dụ: *Email là rác hay không rác*).

4. Cây quyết định (Decision Tree)

Là mô hình phân loại dựa trên các câu hỏi logic dạng "Nếu - Thì" (If-Then). Cấu trúc gồm:

- Nút gốc (Root):** Câu hỏi quan trọng nhất bắt đầu cây.
- Nút lá (Leaf):** Kết quả dự đoán cuối cùng.

Sử dụng khi cần sự minh bạch, dễ giải thích tại sao máy lại đưa ra quyết định đó.