

Cấu trúc của Mạng Nơ-ron truyền thẳng (Feed-forward Neural Networks)

- **Các Node (Nút):** Được biểu diễn bằng các hình tròn, là đơn vị xử lý trong mạng.
 - **Các Kết nối:** Được biểu diễn bằng các đường thẳng, nối giữa các node.
 - **Các Lớp (Layers):** Các node được tổ chức thành các lớp.
 - **Lớp đầu vào (Input Layer):** Nơi dữ liệu bắt đầu đi vào.
 - **Các lớp ẩn (Hidden Layers):** Nơi dữ liệu được xử lý qua các phép tính.
 - **Lớp đầu ra (Output Layer):** Nơi tạo ra kết quả cuối cùng.
 - **Truyền thẳng (Feed-forward):** Dữ liệu luôn di chuyển về phía trước qua các lớp, không có kết nối nào quay ngược lại. Nếu có, đó sẽ là mạng nơ-ron hồi quy (recurrent neural network).
-

Quá trình hoạt động và các thành phần chính

1. **Trọng số (Weights):** Khi dữ liệu đi qua một kết nối, một trọng số được áp dụng cho nó.
2. **Tổng có trọng số:** Mỗi node sẽ tính tổng các đầu vào đã được nhân với trọng số tương ứng.
3. **Hàm kích hoạt (Activation Function):** Một hàm phi tuyến được áp dụng cho tổng có trọng số. Các hàm kích hoạt phổ biến bao gồm:
 - Hàm Sigmoid (ví dụ: tan h, hàm logistic).
 - Rectified Linear Units (ReLU).

Lợi ích của độ sâu trong mạng nơ-ron

1. **Tính chất mô-đun & có thể mở rộng**
 - Các lớp là mô-đun → có thể thêm, bớt hoặc thay đổi cấu trúc tùy bài toán.
2. **Tính chất phổ quát**
 - Một lớp ẩn duy nhất có thể xấp xỉ bất kỳ hàm liên tục nào (universal approximation), *nếu* có đủ node.
 - Tuy nhiên, nhiều lớp ẩn (deep) thường dễ học hơn trong thực tế.
3. **Tạo đặc trưng thông qua hợp thành (composition)**
 - Mỗi lớp học đặc trưng từ lớp trước → xếp chồng tạo ra đặc trưng phức tạp hơn.
 - Ví dụ ảnh:
 - Lớp đầu → học cạnh, đường thẳng
 - Lớp giữa → kết hợp cạnh → hình dạng
 - Lớp sâu hơn → nhận diện đối tượng (cú, mèo,...)
4. **Tạo trừu tượng hóa (abstraction)**
 - Mỗi lớp biểu diễn cấp độ trừu tượng cao hơn của dữ liệu.

- Mạng có thể biểu diễn “hình ảnh chứa cú hay không” bằng một bit duy nhất.

5. Bottleneck Layer

- Thiết kế các lớp sau ngày càng nhỏ → ép thông tin thành các đặc trưng cốt lõi.
- Giảm nhiễu, giữ lại thông tin có ích cho dự đoán.