

Neural Network

- Mạng nơ-ron nhân tạo (Artificial Neural Networks - ANN) ban đầu được phát triển để mô phỏng cách bộ não con người hoạt động. Mặc dù ANN lấy cảm hứng từ nơ-ron sinh học, nhưng chúng thực tế chỉ mô phỏng đơn giản lại các tín hiệu và kết nối thần kinh bằng toán học. Trong khi đó, hiểu biết về bộ não con người vẫn còn rất hạn chế, và những nỗ lực mô phỏng hoàn toàn nó hiện không khả thi.
- Một lý do lớn giúp deep learning bùng nổ là nhờ sự gia tăng dữ liệu số (digital data) cùng với sự phát triển của phần cứng mạnh như GPU. Các thuật toán truyền thống như hồi quy tuyến tính không thể tận dụng hết lượng dữ liệu lớn, trong khi các mạng nơ-ron lớn hơn lại có thể nâng cao hiệu suất đáng kể khi có đủ dữ liệu.

Neural là gì?

Một đơn vị hồi quy logistic có thể được coi như một **nơ-ron đơn**, đầu vào là x , đầu ra là a (activation).

Neural Network

Một mạng neural là một tập hợp các neural có thể chia thành 3 lớp:

- **Input layer:** vector đầu vào x .
- **Hidden layer:** lớp trung gian, trích xuất các đặc trưng nội tại.
- **Output layer:** dự đoán xác suất.

Mạng nơ-ron có thể tự học đặc trưng phù hợp từ dữ liệu mà không cần xác định trước.

- Mỗi lớp gồm nhiều **nơ-ron** (neuron), mỗi nơ-ron thực hiện phép tính tương tự **hồi quy logistic**.
- Một mạng có thể có nhiều lớp, từ vài lớp đến hàng trăm lớp.

Forward propagation Algorithm

Forward propagation là quá trình tính toán để mạng nơ-ron đưa ra dự đoán.

Giả sử có một cấu trúc mạng nơ-ron:

- **Input layer:** 64 đầu vào
- **Hidden layer 1:** 25 nơ-ron
- **Hidden layer 2:** 15 nơ-ron
- **Output layer:** 1 nơ-ron đầu ra (xác suất)

Các bước tính toán:

1. **Từ đầu vào $x \rightarrow a_1$:** Sử dụng công thức tuyến tính và hàm kích hoạt, với 25 tham số trọng số và bias.
2. **Từ $a_1 \rightarrow a_2$:** Tính toán tương tự, đầu vào là a_1 , đầu ra là 15 giá trị a_2 .
3. **Từ $a_2 \rightarrow a_3$:** Đầu ra cuối cùng là a_3 , một giá trị xác suất (scalar).
4. **Phân loại nhị phân:** So sánh a_3 với ngưỡng

Lưu ý:

- $a_0 = x$ theo quy ước.
- $f(x)$ biểu thị hàm mà mạng nơ-ron tính được.
- Forward propagation là quá trình lan truyền thông tin từ trái sang phải (từ input đến output).
- Khác với **backward propagation** (lan truyền ngược), dùng để huấn luyện mô hình.