

TÌM HIỂU VỀ BACKPROPAGATION

1. Giới thiệu

Thuật toán Backpropagation là hiểu đơn giản là giúp tính toán gradient của hàm mất mát đối với các trọng số của mạng.

Ta cần phải hiểu cách tính đạo hàm riêng của hàm mất mát theo từng trọng số w và bias

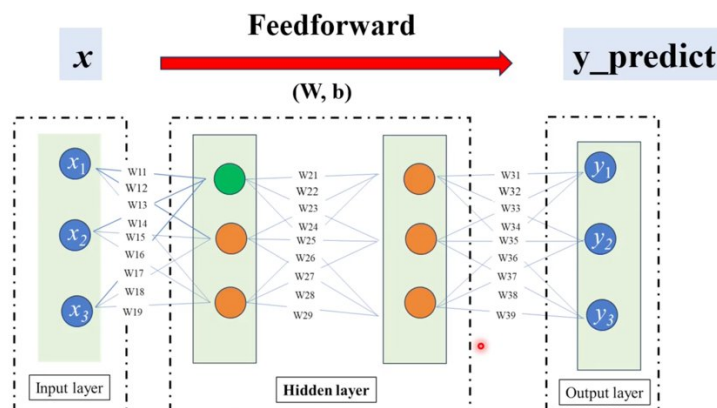
Trong mạng. Từ đây ta cập nhật trọng số để tăng độ chính xác của mô hình.

Ta cần phải biết công thức đạo hàm riêng.

$$\begin{aligned}\text{VD: } f(x) &= 2x+1 & g(f(x)) &= 2(f(x))+5 \\ g'(f(x)) &= 2 \cdot (2x+1)' = 2 \cdot 2 = 4\end{aligned}$$

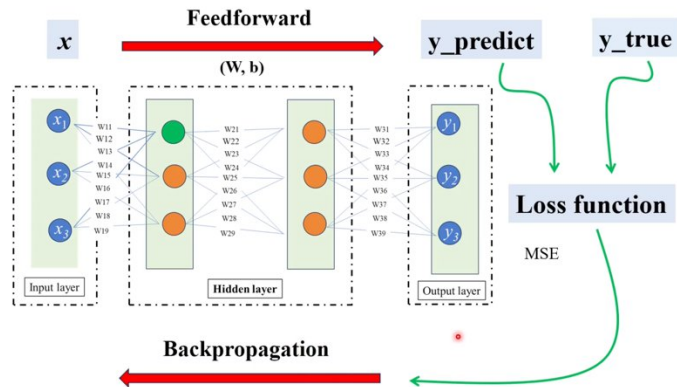
2. Nguyên lý hoạt động

a, Feedforward



Tư tưởng của nó chính là ta đi y_{predict} dựa vào input đầu vào và đối với các trọng số lúc này ta lấy giá trị ngẫu nhiên

b, Backpropagation



-Lúc này ta cần hàm Loss function để đánh giá kết quả của mô hình, ta có thể dùng MSE. Ta cần tối ưu hàm mất mát hay nói cách khác là tìm các giá trị W để mô hình tối ưu.

-Ta cần biết về thuật toán gradient decent để tối loss function cho mỗi một node cụ thể.

-Cơ bản để tối ưu hàm mất mát ta cần đi ngược hướng của gradient để đi về điểm cực tiểu (có thể điểm đó không phải tối ưu nhất nhưng nó cũng có thể là cực tiểu địa phương).

- Ta có thể nhận giá trị tối ưu khi $fn(x) - fn-1(x) < threshold$. Với giá trị threshold là giá trị ta tự cho.

-Ta sẽ tính đạo hàm cho từng node cụ thể cho bài toán áp dụng đạo hàm hợp và áp dụng gradient decent để tối ưu.