



**Data Science
Academy**

www.datascienceacademy.com.br

Machine Learning

Backpropagation, Gradiente Descendente e
Chain Rule



Por Que Precisamos de Backpropagation (Retropropagação)?

Ao projetar uma rede neural, primeiro, precisamos treinar um modelo e atribuir pesos específicos a cada uma das entradas. Esse peso decide o quão vital é esse recurso para nossa previsão. Quanto maior o peso, maior a importância. No entanto, inicialmente, não sabemos o peso específico exigido pelas entradas. Então, o que fazemos é atribuir um peso aleatório às nossas entradas e nosso modelo calcula o erro na previsão. Depois disso, atualizamos nossos valores de peso e executamos novamente o código (retropropagação). Após várias iterações, podemos obter valores de erro mais baixos e maior precisão.

Para que a retropropagação funcione usamos um algoritmo que verifica quanto o valor de cada peso afeta o erro do modelo, calculando as derivadas parciais. Esse algoritmo é chamado de Gradiente Descendente e é aplicado através da Chain Rule (Regra da Cadeia).

Em outras palavras, a retropropagação visa minimizar a função de custo ajustando os pesos e vieses (bias) da rede. O nível de ajuste é determinado pelos gradientes da função de custo em relação a esses parâmetros.

Uma pergunta pode surgir: por que calcular gradientes?

O gradiente de uma função $C(x_1, x_2, \dots, x_m)$ no ponto x é um vetor das derivadas parciais de C em x .

A derivada de uma função C mede a sensibilidade à alteração do valor da função (valor de saída) em relação a uma alteração no argumento x (valor de entrada). Em outras palavras, a derivada nos diz a direção que C está seguindo.

O gradiente mostra quanto o parâmetro x precisa mudar (na direção positiva ou negativa) para minimizar C . O cálculo desses gradientes acontece usando uma técnica chamada Regra da Cadeia.

É o que veremos nas próximas aulas.

Referências:

Algoritmo Backpropagation Parte 1 – Grafos Computacionais e Chain Rule

<http://deeplearningbook.com.br/algoritmo-backpropagation-parte1-grafos-computacionais-e-chain-rule/>