

Model

A Função de Custo, também conhecida como Função de Perda ou Loss Function, é um componente absolutamente crucial no treinamento de modelos de machine learning, especialmente redes neurais. Ela quantifica o "quão errado" o seu modelo está.

O objetivo do treinamento de um modelo é **minimizar essa função de custo**. É por meio da otimização da função de custo que o modelo ajusta seus pesos e vieses para fazer previsões mais precisas.

Função de Custo - MSE (Mean Squared Error)

A **Mean Squared Error (Erro Quadrático Médio)** é uma das funções de custo mais amplamente utilizadas, especialmente em problemas de regressão, onde o objetivo é prever um valor contínuo.

A Idéia por Trás do MSE:

O MSE mede a média dos quadrados das diferenças entre os valores previstos pelo modelo e os valores reais (os rótulos verdadeiros). Ele penaliza erros maiores de forma mais significativa devido ao termo quadrático.

A Fórmula da Função MSE:

Para um conjunto de N amostras, onde y_i é o valor real (verdadeiro) e \hat{y}_i é o valor previsto pelo modelo para a i -ésima amostra, o MSE é calculado como:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

Onde:

- N é o número total de amostras (ou o tamanho do lote, se estivermos falando de treinamento em mini-lotes).
- y_i é o valor real (ground truth) para a i -ésima amostra.
- \hat{y}_i é o valor previsto pelo modelo para a i -ésima amostra.
- $(y_i - \hat{y}_i)$ é o erro para a i -ésima amostra.
- $(y_i - \hat{y}_i)^2$ é o erro quadrático para a i -ésima amostra.
- $\sum_{i=1}^N$ indica a soma de todos os erros quadráticos.
- $\frac{1}{N}$ indica que estamos calculando a média desses erros quadráticos.

Características Principais do MSE:

- **Diferenciável:** É uma função contínua e diferenciável, o que é essencial para o uso de algoritmos de otimização baseados em gradientes (como o Gradiente Descendente).
- **Penaliza Erros Grandes:** Devido ao termo quadrático, erros maiores são penalizados de forma mais severa do que erros menores. Por exemplo, um erro de 2 contribui com $2^2 = 4$ para o custo, enquanto um erro de 4 contribui com $4^2 = 16$.
- **Sensível a Outliers:** A penalidade quadrática torna o MSE bastante sensível a outliers (valores atípicos), pois um único erro grande pode aumentar significativamente o custo total.
- **Unidades:** A unidade do MSE é o quadrado da unidade da variável de saída (por exemplo, se você está prevendo preços em dólares, o MSE estará em dólares quadrados). Por isso, às vezes é preferível usar o RMSE (Root Mean Squared Error), que é a raiz quadrada do MSE e tem a mesma unidade da variável de saída, tornando-o mais interpretável.

Exemplo de Cálculo:

Suponha que seu modelo fez as seguintes previsões para três amostras, e os valores reais são:

Amostra (i)	Valor Real (y_i)	Valor Previsto (\hat{y}_i)	Erro ($y_i - \hat{y}_i$)	Erro Quadrático ($(y_i - \hat{y}_i)^2$)
1	10	11	$10 - 11 = -1$	$(-1)^2 = 1$
2	20	18	$20 - 18 = 2$	$(2)^2 = 4$
3	5	7	$5 - 7 = -2$	$(-2)^2 = 4$

Agora, vamos calcular o MSE:

1. Soma dos Erros Quadráticos:

$$\sum_{i=1}^3 (y_i - \hat{y}_i)^2 = 1 + 4 + 4 = 9$$

2. Divide pelo número de amostras ($N = 3$):

$$MSE = \frac{1}{3} \times 9 = 3$$

Portanto, o MSE para essas previsões é 3.

Aqui está uma imagem que ilustra o conceito do MSE:

