

Módulo 2 | Aula #4

Treinamento de
redes neurais

Tempo estimado de leitura:
5 min

Overfitting? O que é isso?

Overfitting é um problema que pode ocorrer durante o treinamento de um modelo. Como você sabe, o modelo aprende com os dados de treinamento e, com isso, analisa futuras entradas para prever as respostas corretas.

No contexto de avaliação de imóveis, por exemplo, o modelo aprende com os exemplos e ele nos diz o preço ideal de venda de novos imóveis, pois ele aprendeu a realizar esta tarefa.

Quando o modelo aprende a realizar as previsões apenas para os dados de treinamento e não de maneira generalizada, ocorre o overfitting. Basicamente, isso é devido ao ajuste acentuado que fazemos na função e ela acaba "**decorando**" a **saída** para cada entrada ao invés de aprender a realizar a tarefa.

Como podemos evitar o overfitting?

Para evitar o *overfitting*, podemos considerar os seguintes pontos:

1. Utilizar uma **função de regularização** - quando temos muitas características úteis a serem aprendidas
2. **Diminuir a quantidade de características** - selecionamos as que são menos relevantes para o problema em questão

Como podemos identificar o overfitting?

Para identificar o *overfitting*, devemos analisar se a performance do modelo usando os dados de treinamento ficou muito boa, enquanto que para os dados de teste ficou muito ruim. Isso significa que o modelo sabe as saídas corretas para os dados que foram usados durante o treino, mas quando apresentamos novos dados ele não sabe retornar boas saídas.