

Inteligência Artificial e Machine Learning

Lista 3

José Elias cangombe

Matrícula - 799

1-Suponha que você esteja usando regressão polinomial. Você plota as curvas de aprendizado e percebe que há uma grande diferença entre o erro de treinamento e o erro de validação. O que está acontecendo? Quais são as três maneiras de resolver isso?
OBS.: Curvas de aprendizado: são gráficos mostrando o desempenho do modelo no conjunto de treinamento e no conjunto de validação em função do tamanho do conjunto de treinamento (ou da iteração do treinamento).

R: Sempre que na regressão polinomial apresenta uma grande diferença entre o erro de treinamento e o erro de validação é porque nesta ordem não há flexibilidade ou complexidade para aproximar bem os dados.

E para resolver este problema recorreremos as seguintes ordens:

Uma vez que um polinômio de ordem 2 produz melhor a aproximação dos dados. E esta aproximação será melhor quando maior o conjunto de treinamento e menor ruído. Ainda este modelo com polinômio de grau 2 parece ter aprendido bem as regularidades dos dados sem aprender o ruído no conjunto de treinamento. Nesse caso, os erros no conjunto de treinamento (**erro de capacidade**) e teste (**erro de generalização**) são parecidos e satisfatórios.

Já o modelo **com polinômio de grau 30** sofre de **sobre-ajustamento**, isto é, ele tem uma capacidade tão alta que além de aprender as regularidades, aprendeu também o ruído presente nos exemplos de treinamento. Nesse caso, o erro no conjunto de treinamento é baixo (**erro de capacidade**) e o erro no conjunto de teste é alto (**erro de generalização**).