

Inteligência Artificial TP- 555

Matrícula: 809

Aluno: Antonio Manuel Ngunza Luciano

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: Acontece que possivelmente o polinômio utilizado não tem a flexibilidade, ou seja, a complexidade suficiente para aproximar os dados, o grande problema pode estar na complexidade de dados que fornece as curvas que parecem menos úteis para fazer uma ótima previsão. No caso de se utilizar o polinômio de ordem 1 (subajuste) ele não terá flexibilidade suficiente para aproximar os dados, ele não consegue capturar a curvatura dos pontos treinamentos, ele erra muito para os pontos de treinamento quanto para os pontos vistos durante treinamento. Uma forma de resolver esse problema é necessário aumentar a complexidade do modelo no caso da regressão polinomial aumentar sua ordem, para o modelo com polinômio de grau 30 (sobreajuste) tem capacidade de erro de treinamento elevado e erro de predição muito baixo, e no caso de um polinômio de ordem 2 produz melhor aproximação dos dados, a aproximação será melhor quanto maior o conjunto de treinamento e menos o ruído, modelos ótimos em termos de balanço e entre os erros de capacidade e de generalização.

As técnicas de Regularização visa obter um modelo bem comportado (ou seja, que não sobreajuste) através da incorporação de informações adicionais ao processo de treinamento do modelo, na forma de restrições de suavidade junto ao mapeamento ou de penalizações proporcionais a norma do vetor de pesos.