

Destaco como aprendizado neste projeto o desafio de tratar, analisar e aprender a partir de muitas variáveis.

A partir disso destaco o aprendizado em utilizar a regressão linear para validação e seleção dos dados de tal forma que ao treinar o modelo de predição seja alcançada um melhor resultado sem que o mesmo seja treinado em informações que agregam pouco ou nada. O uso da regressão se fez necessário dado o grande número de variáveis, mais de 80, o que dificulta a análise e escolhas das mesmas.

Fora isso, nas aulas também foram ditas importantes dicas em relação aos dados estatísticos, pois nem sempre os dados estarão perfeitos, ou seja, de forma gaussiana/homocedástico/VIF e nesses momentos temos de tentar aperfeiçoar os dados ao máximo, mas nem sempre teremos um mundo teórico estatístico perfeito nos dados e devemos aceitar e lidar com isso.