



MACHINE LEARNING

Curso de especialização em Ciência de Dados e Big Data

Prof. Hugo de Paula

ATIVIDADE EM LABORATÓRIO 4

BOOSTING

OBJETIVOS

- 1- A base de dados SONAR foi utilizada por Gorman e Sejnowski, em seu trabalho sobre redes neurais, para classificar sinais de sonar. A tarefa consistia em treinar uma rede neural para diferenciar entre um cilindro de metal – potencial mina terrestre – de uma rocha cilíndrica.
 - a. Cada registro da base possui 60 atributos numéricos entre 0 e 1.0, e uma classe.
 - b. São 11 padrões para minas e 97 padrões para rocha.
- 2- Este trabalho irá comparar o uso de redes neurais e de técnicas de *boosting* para a classificação de sinais de sonar.

EXERCÍCIO

- 1- Vamos construir três modelos:
 - a. Árvore de decisão.
 - b. AdaBoost (**AdaBoostClassifier**). Configure o parâmetro **n_estimators=50** do AdaBoost para 5. Esse parâmetro definirá o número máximo de modelos que serão gerados.
 - c. Redes neurais (**MLPClassifier**).
- 2- Explore as saídas dos modelos.
 - a. Descreva a topologia da rede que foi treinada. Compare o conjunto de regras da árvore de decisão com os modelos previstos pelo AdaBoost.
 - b. Verifique e compare a matriz de confusão e as medidas de performance: acurácia, precisão e revocação, de cada modelo.



- c. Analise a interpretabilidade dos modelos. Altere as opções de ***prunning*** e ***preprunning*** no AdaBoost para criar árvores mais simples, mas que tenham igual poder de previsão.

A atividade deve ser entregue na forma de um Notebook desenvolvido no Google Colab (<https://colab.research.google.com/>), contendo os nomes dos integrantes da dupla, e compartilhado com o e-mail hugodepaula@gmail.com, até 7 dias após a realização na aula.