

Mini teste-2

A apresentar na aula da semana a seguir à data descrito em cada teste. As respostas não devem exceder mais do que duas páginas A4.

Teste-2

1. Execute um 10-XVal para o dataset *vote* com o algoritmo J48 e com o AdaBoostM1 sobre J48. Apresente os resultados obtidos em termos de erro. Sugira uma explicação para a comparação dos erros obtidos e também para os valores de AUC.
2. Que tipo de benefícios esperaria da aplicação de Bagging sobre Naive Bayes num dataset específico, sabendo que o resultado do modelo individual Naive Bayes nesse dataset é erro = 0.085. Justifique.

3. Para o dataset “diabetes” apresente um estudo sobre custo de erros. Usando por exemplo J48 e NaiveBayes, apresente resultados para diferentes matrizes de custo. Faça as avaliações orientadas aos custos e articule conclusões sobre os resultados obtidos usando modelos sensíveis ao custo.
4. Apresente as principais diferenças entre algoritmos de *clustering* baseadas em partições e algoritmos hierárquicos. Use exemplos para ilustrar a sua resposta.