

Aprendizagem Automática

Perguntas de preparação para 2ª frequência

1. Indique de que modo um classificador linear binário poderá ser usado para classificar um dataset com 3 classes.
2. Indique as principais diferenças entre os algoritmos de classificação de comité bagging e boosting e stacking.
3. Descreva o modo de decisão num algoritmo de comité usando um sistema de *voting* (votação)
4. O algoritmo *Random Forest* introduz aleatoriedade (daí o nome) no processo de construção das árvores que constituem o comité. Indique como é introduzida essa aleatoriedade?
8. Considere a seguinte matriz de confusão de um problema de classificação com 3 classes.

real\classificador	Iris-setosa	Iris-versicolor	Iris-virginica
Iris-setosa	47	2	1
Iris-versicolor	1	42	7
Iris-virginica	4	9	37

- a. Calcule o valor da exatidão.
 - b. Calcule o valor da precisão de cada classe sabendo que a mesma é calculada pela fórmula $VP/(VP+FP)$.
 - c. Calcule o valor da precisão do classificador usando macro-média
 - d. Calcule o valor da precisão do classificador usando média ponderada
13. Uma rede neuronal é treinada através da atualização dos seus pesos ao longo de várias iterações (épocas). A **função de loss** mede a distância entre a previsão e o valor objetivo (o que se pretende que seja a saída da rede) indicando, assim, quão bem a rede funcionou para esse exemplo específico. A figura seguinte mostra o valor da função de *loss* ao longo do treino da rede (épocas/*Epochs*) para os conjuntos de treino e validação.
- a. Qual o objetivo da utilização de um conjunto de validação? Justifique a resposta.
 - b. Em que época considera que a rede terá os parâmetros mais adequados ao problema? Justifique a resposta.

