EBAC
CIÊNCIA DE DADOS
M23E02 - RANDOM FOREST

BOOTSTRAP

A partir de um dataframe é gerado outros com o mesmo número de linhas com repetição

2

FEATURE SELECTION

São geradas colunas aleatoriamente em nº:

- classificação: m = √p
- regressão: m = p/3

RANDOM FOREST

> Baseado na técnica Bagging

3

BASE LEARNERS

É o nome para cada um desses novos dataframes gerados

4

MODELO DE CLASSIFICAÇÃO

Para cada um dos base learners é aplicado uma árvore de decisão

5

AGGREGATING

A agregação pode ser feita pela frequência ou pela média