

Teoria T03

Preencha o círculo dos itens verdadeiros e deixe em branco o dos falsos.

- ☐ O uso de instâncias redundantes em um conjunto de treinamento pode ser bem quisto considerando métodos para seleção de características.
- ☐ *Outliers* podem levar a um super-ajuste do modelo.
- ☐ Com a técnica de encestamento no pré-processamento dos dados, há alteração dos valores originais dos dados a serem submetidos para treinamento.
- ☐ Vale à pena investir num classificador binário com 90% de acurácia onde 8% dos exemplos pertencem a uma das classes.
- ☐ Para representação numérica de *features* categóricos ordinais, o mapeamento para valores inteiros é uma saída. A codificação *one-hot* usada para *features* categóricos nominais também seria opção viável.
- ☐ Se comparada com L2, a geometria da regularização L1 favorece a preservação das ponderações sobre *features* ao cabo do treinamento do modelo.
- ☐ É possível utilizar L2 para eliminar *features* menos relevantes.
- ☐ No SBS a própria performance do modelo é utilizada para selecionar o subconjunto de *features* desejáveis.
- ☐ A amostragem de *features* geradas aleatoriamente com substituição é importante em *random forests* para que tenhamos árvores menos correlacionadas.
- ☐ A principal razão das *random forests* utilizarem amostragem de *features* ao invés do conjunto inteiro original no treinamento de cada árvore individual é que o custo de processamento total seria enorme.