

## Inteligência Artificial

### Verificação Especial – 2019 – 3º Ano

Elabore um relatório sobre um estudo comparativo de aplicação dos algoritmos de classificação (classificadores) indicados abaixo em um conjunto de dados à escolha dentre os disponíveis no *UCI Machine Learning Repository*. Utilize a ferramenta *scikit-learn* na realização do estudo.

- K-NN (k Vizinhos Mais Próximos)
- Naïve CBayes (Classificador Bayesiano Ingênuo)
- C4.5 (ou outro classificador baseado em árvore de decisão)
- MLP (Multi-layer Perceptron com Back-Propagation)
- SVM (*Support Vector Machine*)

O relatório deverá conter, no mínimo, as seguintes informações:

- A indicação do conjunto de dados utilizado.
- Uma explicação sobre a semântica dos atributos do conjunto.
- Uma visão estatística do conjunto que reporte a distribuição dos dados (global e por classe), evidenciando características relevantes (ex. número de dados ausentes, médias e desvios).
- As funções de pré-processamento aplicadas aos dados, devidamente explicadas e justificadas.
- Os parâmetros utilizados para cada classificador, assim como o método utilizado em sua escolha. Sugestão: Particione o conjunto de dados em treino, validação e teste, e utilize o conjunto de validação para escolher a melhor configuração de parâmetros. Indique a proporção de tuplas adotada no particionamento.
- Apresentação dos desempenhos dos classificadores no conjunto de teste.
- Uma análise do desempenho de cada classificador, incluindo possíveis justificativas que tenham levado ao referido desempenho.
- Uma análise comparativa dos desempenhos apresentados pelos classificadores, indicando qual (ou quais) classificador(es) você selecionaria para classificar novos casos que possam surgir no contexto da aplicação do conjunto de dados investigado.
- Indicação de sugestões de possíveis ações (não executadas por decurso de prazo) que, ao serem desenvolvidas futuramente, possam vir a melhorar o desempenho de cada classificador, quando necessário. Justifique cada sugestão.

Itens a serem entregues:

- Relatório (pdf)
- Código desenvolvido (python)

Os seguintes aspectos serão considerados na avaliação do trabalho:

- Correta aplicação de conceitos e recursos
- Clareza e organização na apresentação do conteúdo
- Completude dos conteúdos expostos
- Coerência e profundidade das explicações das escolhas realizadas
- Coerência e profundidade das análises apresentadas

Importante:

- O trabalho é individual.
- Como o trabalho é individual, o conjunto de dados escolhido pelo aluno não poderá ter sido escolhido por outro aluno da turma.
- O conjunto de dados *Iris Plant* não poderá ser escolhido.

Entrega: 18/11, via EaD.