

Érika Mara de Moraes Machado
Maria Luiza Bernardo Madeira

Matrícula: 202310058
Matrícula: 202310981

RELATÓRIO TÉCNICO - ALGORITMO MLP

Este trabalho tem como objetivo a análise comparativa de algoritmos de classificação supervisionada aplicados a dois datasets: **Iris** e **Wine**. Foram utilizadas duas abordagens principais: a rede neural artificial **MLP (Multi-Layer Perceptron)** e o algoritmo baseado em instâncias **KNN (K-Nearest Neighbors)**. Além disso, foram incluídas duas variações do KNN, uma implementada manualmente (KNN Hardcore) e outra com a biblioteca scikit-learn, mas estas tiveram caráter secundário e serviram apenas como suporte na comparação geral entre KNN e MLP.

Os datasets escolhidos foram o **Iris**, com 150 amostras de flores descritas por quatro atributos numéricos e divididas em três classes, e o **Wine**, com 178 amostras de vinhos italianos descritas por 13 atributos químicos e também divididas em três classes. Em ambos os casos, os dados foram separados em **80% para treino e 20% para teste**, utilizando `random_state=42` para garantir reprodutibilidade.

A primeira abordagem testada foi o **MLPClassifier**, que é uma rede neural artificial de múltiplas camadas totalmente conectadas. Antes do treinamento, foi realizado o pré-processamento das amostras com a técnica de **padronização (StandardScaler)**, garantindo que todas as features tivessem média zero e desvio padrão igual a um. O modelo foi configurado com duas camadas escondidas contendo, respectivamente, 64 e 32 neurônios, e com limite de 1000 iterações para o processo de aprendizado, o que assegurou a convergência em todas as execuções. A rede neural foi avaliada por meio da acurácia, da precisão, do recall e do relatório de classificação completo, além da geração de matrizes de confusão. Os resultados obtidos mostraram que o MLP apresentou **desempenho consistente e elevado**, atingindo acurácia próxima de 96% no dataset Iris e valores também superiores a 90% no dataset Wine. As matrizes de confusão mostraram que a rede neural conseguiu separar corretamente a maioria das classes, com poucos erros de classificação entre categorias semelhantes.

Já o **KNN** foi testado em suas duas versões implementadas anteriormente. Na implementação manual, foi realizada a programação do cálculo da distância Euclidiana e da seleção dos vizinhos mais próximos, enquanto na versão com scikit-learn utilizou-se diretamente a classe `KNeighborsClassifier`. Ambas apresentaram resultados satisfatórios, com acurácia variando entre 90% e 97% a depender do valor de `k`. A principal diferença observada foi que a versão manual, apesar de útil para fins didáticos, não possui otimizações e se torna menos eficiente em bases maiores, enquanto a versão do scikit-learn é mais prática, rápida e escalável.

Ao comparar os três métodos, pode-se afirmar que todos foram eficazes na tarefa de classificação dos datasets escolhidos. O **MLPClassifier** se destacou por apresentar resultados mais robustos e consistentes, com grande capacidade de generalização e baixo número de erros. O **KNN Hardcore** cumpriu bem seu papel didático, pois proporcionou uma compreensão detalhada do funcionamento interno do algoritmo KNN, ainda que tenha limitações em termos de eficiência computacional. Já o **KNN com scikit-learn** mostrou-se a versão mais prática para aplicações reais, uma vez que alia precisão, velocidade e flexibilidade na configuração de parâmetros.

Link para o vídeo de apresentação no YouTube: