

Relatório de Implementação do algoritmo KNN

Introdução

O trabalho consiste na implementação do algoritmo KNN de duas maneiras, com o objetivo fixar os conhecimentos adquiridos em sala de aula. A primeira implementação é manual, sem o uso de bibliotecas que já possuem o KNN implementado, e a segunda é usando bibliotecas como o *Sklearn*. Toda a implementação foi feita sobre a base de dados [Iris](#).

Metodologia

- Base de dados: Iris dataset.
- Processamento dos dados: Normalização Min-Max aplicada apenas aos dados de treino, e replicada sobre os dados de teste.
- Algoritmo KNN: usando distância euclidiana e votação uniforme, e random state = 13, para que as duas implementações analisem os mesmos dados. 50% dos dados foram usados para treino, e o resto para teste.
- Implementação manual: Usando majoritariamente as bibliotecas *numpy* e *pandas* para facilitar os cálculos, com todas as funções do KNN e de cálculo de métricas implementadas a partir delas.
- Implementação com Sklearn: Usando a função *KNeighborsClassifier* e outras relacionadas

Resultados

Valor de K	Classificador	Acurácia	Precisão(média)	Revocação(média)
1	Manual	0.96	0,9608	0,9594
	Sklearn	0.96	0,9594	0,9608
3	Manual	0.9733	0,9731	0,9731
	Sklearn	0.9733	0,9731	0,9731
5	Manual	0.9867	0,9876	0,9861
	Sklearn	0.9867	0,9861	0,9876
7	Manual	0.9733	0,9753	0,9733
	Sklearn	0.9733	0,9733	0,9753

Conclusão

É possível observar que as duas implementações tem resultados bastante parecidos, principalmente pelo fato de que a *acurácia* para os valores de K testados foram iguais, o que valida a implementação manual. Também é notável que para os testes com K's iguais a 3, 5 e 7, as outras métricas ficaram praticamente iguais.

As ligeiras diferenças podem ser explicadas possivelmente pelas implementações diferentes feitas, onde os cálculos de arredondamento podem ser diferentes, e as decisões do algoritmo em caso de empates também (por exemplo, vizinhos com mesma distância do ponto a se classificar).