

Relatório da Atividade Sobre KNN da Disciplina Sistemas Inteligentes

João Marcello Mendes Moreira

E-mail: joaomarcello.mm@gmail.com

Professor: João Dallyson Sousa de Almeida

Resumo da Atividade. Implementar um programa que carregue um dataset e execute o KNN de modo que o melhor valor para o parâmetro K seja definido automaticamente.

Implementação

Utilizou-se a linguagem Python 3.7 para realização da atividade. O programa carrega o Iris dataset, composto de 150 amostras com 4 características e possui 3 classes. O dataset é dividido da proporção de 77% para treino e 33% pra validação.

A implementação do KNN utiliza a distância euclidiana para mensurar os vizinhos mais próximos de uma amostra. Testa-se ao todo os seguintes valores para K (3, 5, 7, 9, 11, 15, 19). Ao final, o programa mostra o valor de K que obteve a melhor acurácia no conjunto de validação.

Figura 1 - Exemplo de parte da saída do programa. Na figura, vemos o resultado do KNN para K=19.

```
===== Resultado para k = 19 =====
      precision    recall  f1-score   support

    0         1.000      1.000      1.000        18
    1         0.833      0.882      0.857        17
    2         0.857      0.800      0.828        15

 accuracy          0.900          50
 macro avg         0.897      0.894      0.895          50
weighted avg         0.900      0.900      0.900          50

--> O melhor valor para k foi 5 levando em conta a acuracia ( 0.94 ).
```