

knn-UFLA-IA

Este projeto é um trabalho realizado pelos alunos **Gustavo Rodrigues Sousa** e **Luiz Carlos Coelho Conde** para a disciplina de Inteligência Artificial ministrada pelo professor **Ahmed Ali Abdalla Esmin** pelo DCC da UFLA no primeiro semestre de 2020.

Requisitos

Para executar este programa, é necessário ter o Python instalado (versão 3 ou acima)

Executar

Para executar o programa basta estar localizado dentro da pasta do programa e rodar:

```
python main.py
```

ou

```
python3 main.py
```

Caso você tenha mais de uma versão do python instaladas.

Resultados

Numa primeira abordagem foi tentado realizar a divisão dos dados sendo os 80% dos primeiros dados para treino e o restante para predição, porém como os dados estão ordenados pelas suas classes, tanto no arquivo de iris quanto no arquivo de spam, o treinamento ficou viciado e ficou especialista em somente 1 tipo de classe mas não os dos outros, por isso foi realizado uma abordagem diferente dos dados, sendo dividido pelo índice mod 5.

Iris

Não houve uma alteração significativa conforme o aumento de k

Matrizes de confusão

para k = 1

class	Iris-setosa	Iris-versicolor	Iris-virginica
Iris-setosa	10	0	0
Iris-versicolor	0	9	0
Iris-virginica	0	1	10

para k = 3

class	Iris-setosa	Iris-versicolor	Iris-virginica
Iris-setosa	10	0	0
Iris-versicolor	0	9	0
Iris-virginica	0	1	10

para k = 5

class	Iris-setosa	Iris-versicolor	Iris-virginica
Iris-setosa	10	0	0
Iris-versicolor	0	9	0
Iris-virginica	0	1	10

para k = 7

class	Iris-setosa	Iris-versicolor	Iris-virginica
-------	-------------	-----------------	----------------

class	Iris-setosa	Iris-versicolor	Iris-virginica
Iris-setosa	10	0	0
Iris-versicolor	0	8	0
Iris-virginica	0	2	10

Spam

Conforme o aumento de k, o aumento de falsos positivos e de verdadeiros positivos é relevante, conforme é possível ver nas matrizes de confusão. isso indica que os dados utilizados para o treino são majoritariamente positivos, ou seja, spams.

Matrizes de confusão

para k = 1

class	1	0
1	294	67
0	69	491

para k = 3

class	1	0
1	337	167
0	26	391

para k = 5

class	1	0
1	348	225
0	15	333

para k = 7

class	1	0
1	352	292
0	11	266