

Aprendizagem Automática 2022/23

Ficha prática 5

Exercício#5.1 - Algoritmo KNN - classificação

Neste exercício terá um contacto com o Algoritmo K vizinhos mais próximos aplicado num conjunto de dados de pequena dimensão.

Considere o seguinte subconjunto de dados do dataset IRIS, onde cada exemplo é caracterizado por 4 atributos numéricos (cm) e cada instância pertence a uma de 3 classes.

Att1	Att2	Att3	Att4	classe
5.1	3.5	1.4	0.2	Iris-setosa
4.9	3.0	1.4	0.2	Iris-setosa
4.7	3.2	1.3	0.2	Iris-setosa
7.0	3.2	4.7	1.4	Iris-versicolor
6.4	3.2	4.5	1.5	Iris-versicolor
6.9	3.1	4.9	1.5	Iris-versicolor
6.3	3.3	6.0	2.5	Iris-virginica
5.8	2.7	5.1	1.9	Iris-virginica
7.1	3.0	5.9	2.1	Iris-virginica
6.3	2.9	5.6	1.8	Iris-virginica

- 1) Calcule (à mão, ou usando uma folha de cálculo) a que classe pertence o exemplo {5.0, 3.0, 4.0, 2.0} se o conjunto anterior for fornecido ao algoritmo KNN, $K=3$.

Exercício#5.2 - Algoritmo KNN - regressão

Neste exercício terá um contacto com o Algoritmo K vizinhos mais próximos aplicado num conjunto de dados de pequena dimensão para regressão.

Considere o seguinte subconjunto de dados do dataset IRIS, semelhante ao exercício anterior usando 3 atributos para prever o 4º atributo.

Att1	Att2	Att3	Att4 (valor a prever)
5.1	3.5	1.4	0.2
4.9	3.0	1.4	0.2
4.7	3.2	1.3	0.2
7.0	3.2	4.7	1.4
6.4	3.2	4.5	1.5
6.9	3.1	4.9	1.5
6.3	3.3	6.0	2.5
5.8	2.7	5.1	1.9
7.1	3.0	5.9	2.1
6.3	2.9	5.6	1.8

- 1) Calcule (à mão, ou usando uma folha de cálculo) a previsão do atributo 4 em função de Att1, Att2, e Att3 {5.0, 3.0, 4.0} se o conjunto anterior for fornecido ao algoritmo KNN - regressão com K=3 (média com igual peso aos K vizinhos mais próximos)
- 2) Refaça o cálculo anterior considerando como a saída uma média ponderada por

$$W = 1 / (d+0.01) \quad \text{em que } d \text{ é a distância}$$

Exercício#5.3 - KNN

Veja quais as funções para fazer uma classificação em python/scikit-learn usando KNN usando o conjunto completo do dataset IRIS.

- 1) determine o desempenho do classificador resultante (score))
- 2) classifique a instância {5.0, 3.0, 4.0, 2.0}