

TD4 K- Nearest Neighbor

Questions de cours :

1. Critiquer l'algorithme KNN (ses limites).
2. Si on voulait utiliser l'algorithme Knn pour une régression, comment procéderait-on ?
3. Principe de K-NN : dis-moi, qui sont tes voisins, je te dirais qui tu es ! Est-il toujours vrai ?
4. Pour quel type d'applications Knn donne de bon resultats (rechercher sur le net) ?

Exercice 1 : Finir la colonne distance du 1^{er} tableau. Vérifier que nos conclusions sur la classe étaient correctes.

Exercice 2 : Faire un tableau et trouver la classe du dernier exemple (avec valeur manquante).

Exercice 3 : En utilisant l'algorithme de Knn (k = 1, puis K = 3), la distance de Manhattan, puis la distance d'Euclide, classer les instances :

1. O 55 F 100
2. N ? M 85
3. O 49 ? 110

Etude	Age	Sexe	Prêt	Rembourse ?
O	25	M	40	O
N	35	M	60	O
N	45	F	?	O
?	20	F	20	N
N	35	M	120	N
O	52	F	18	N
O	23	M	95	O
O	?	M	62	O
N	60	F	100	N
O	48	F	220	O
O	33	?	150	O

- **Distance euclidienne :** $d(x,y)=\sqrt{\sum_{i=1}^n (x_i - y_i)^2}$
- **Distance de Manhattan :** $d(x,y)=\sum_{i=1}^n |x_i - y_i|$

Faites les TDs, et l'examen sera un jeu d'enfants pour vous, sinon ça sera ... !!!!