



# EXAMEN FOUILLE DE DONNÉE SUJET 1

## EXAMEN DATA MINING

### EXERCICE 1 : CLASSIFICATION PAR ARBRE DE DECISION 08PTS

Une banque dispose des informations suivantes sur un ensemble de ses clients :

Client	M	A	R	E	I
1	moyen	moyen	village	oui	oui
2	élevée	moyen	bourg	non	non
3	faible	âgé	bourg	non	non
4	faible	moyen	bourg	oui	oui
5	moyen	jeune	ville	oui	oui
6	élevée	âgé	ville	oui	non
7	moyen	âgé	ville	oui	non
8	faible	moyen	village	non	non

#### QUESTIONS :

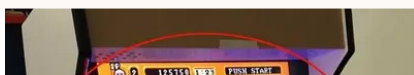
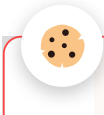
- 1 Quelle est l'entropie de la population ?
- 2 Pour la construction de l'arbre de décision, utilisez-vous l'attribut numéro de client ? Pourquoi ?
- 3 Lors de la construction de l'arbre de décision, quel est l'attribut à tester à la racine de l'arbre ?
- 4 Construisez l'arbre de décision complet.
- 5 Quel est le taux d'erreur de cet arbre estimé sur l'ensemble des clients 1 à 8 ?
- 6 Donnez un intervalle de valeurs pour l'erreur réelle en utilisant une confiance de 90%.

#### ON SE DONNE LES 4 CLIENTS SUIVANTS :

Client	M	A	R	I	E
9	moyen	âgé	village	oui	non
10	élevée	jeune	ville	non	non
11	faible	âgé	village	non	non
12	moyen	moyen	bourg	oui	oui

7) Comment chacun de ces clients est-il classé avec l'arbre de décision que vous avez proposé dans la question 4?

Pour ces 4 clients, on apprend par ailleurs que les clients 9 et 10 gèrent leur compte par Internet, et que les clients 11 et 12 ne le font pas.



AliExpress



클릭하기

8) Quel est le taux d'erreur estimé sur les clients 9, 10, 11 et 12? Combien y a-t-il de faux positifs et de faux négatifs?

## EXERCICE 2 : APPRENTISSAGE DES RÈGLES D'ASSOCIATION 05PTS

Soit le tableau suivant d'un magasin de vente de fruits frais :

Transaction ID	Items
T1	Peanuts, Lemon, Grape
T2	Peanuts, Lemon
T3	Peanuts, Orange, Apple
T4	Apple, Orange
T5	Apple, Grape
T6	Peanuts, Orange, Apple

- 1 Montrer que l'union de deux itemsets fréquents n'est pas toujours fréquents.
- 2 Tracer les résultats de l'utilisation de l'algorithme Apriori sur ce tableau avec un seuil de support  $s=33.34\%$  et un seuil de confiance  $c=60\%$ .
- 3 Montrer les ensembles candidats et fréquents pour chaque balayage de la base de données.
- 4 Énumérer tous les ensembles fréquents finaux. Indiquez également les règles d'association générées et mettez en évidence celles qui sont fortes, en les triant par degré de confiance.

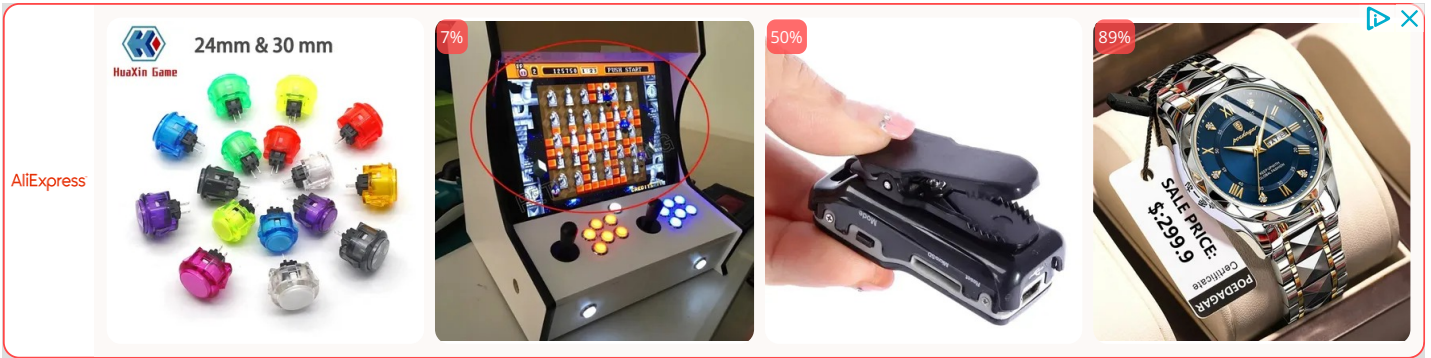
## EXERCICE 3 : MACHINE A VECTEURS DE SUPPORT ET RESEAUX DE NEURONES 07PTS

A) Répondre pa vrai ou Faux tout en justifiant votre reponse

- 1 On peut s'attendre à ce que les vecteurs de support restent généralement les mêmes lorsque l'on passe d'un noyau linéaire à des noyaux polynomiaux d'ordre supérieur.
- 2 Les limites de décision à marge maximale construites par les machines à vecteurs de support présentent l'erreur de généralisation la plus faible de tous les classifieurs linéaires.
- 3 Toute limite de décision obtenue à partir d'un modèle génératif avec des distributions gaussiennes conditionnelles de classe peut en principe être reproduite avec un SVM et un noyau polynomial de degré inférieur ou égal à trois.
- 4 Les valeurs des marges obtenues par deux noyaux différents  $K_1(x,x_0)$  et  $K_2(x,x_0)$  sur le même ensemble d'apprentissage ne nous indiquent pas quel classifieur sera le plus performant sur l'ensemble de test.

B) Les réseaux de neurones, également connus sous le nom de réseaux de neurones artificiels (ANN) constituent un sous-ensemble de l'apprentissage automatique et sont au cœur des algorithmes de l'apprentissage en profondeur. Soit ? et ? deux variables booléennes.

- 1 Concevoir un réseau de neurones à deux entrées permettant d'implémenter la fonction booléenne  $A \wedge B$ .
- 2 Concevoir un réseau de neurones à deux couches implémentant la fonction booléenne  $A \text{ XOR } B$ .
- 3 En déduire un réseau de neurones implémentant la fonction booléenne :  $\neg(A \text{ XOR } B)$  sachant que :  $\neg(A \text{ XOR } B) = \neg((A \wedge B) \vee (\neg A \wedge B))$   
 $= \neg(A \wedge B) \wedge \neg(\neg A \wedge B)$



## CORRIGE EXAMEN FOUILLE DE DONNEE :

BIENTOT

Si vous avez trouvé les examens corrigés en fouille de donnée de Mr Joël\_Yk intéressants et utiles, pourquoi ne pas les partager avec d'autres personnes qui pourraient également en bénéficier ? Partagez ce lien sur les réseaux sociaux ou envoyez-le à vos amis et collègues. Vous pourriez aider quelqu'un à améliorer ses compétences en programmation ou à trouver des solutions à des problèmes complexes. N'oubliez pas que la connaissance doit être partagée pour grandir. Merci pour votre soutien et votre partage !

Contact WhatsApp : +237 652027193 | Réaliser Par [Joël\\_Yk](#)



☆☆☆☆☆ Aucune note. Soyez le premier à attribuer une note !

### AJOUTER UN COMMENTAIRE

Nom \*

E-mail

Site Internet

**B** *I* U **Tt** A " ☰ 🔗 🗑️ 🖼️ 📄 ⏪ ☺️ ✓ Aperçu