

Mini-projet : Mise en œuvre des techniques de data mining

Objectifs du mini-projet

1) Vous familiariser avec le langage Python, 2) Vous faire traiter des jeux de données (datasets) en utilisant Python, 3) Vous faire apprendre les différentes techniques de data mining.

Détails du mini-projet

1 à 2 étudiant-e-s peuvent choisir un des mini-projets listés ci-dessous. Le lien donné vous amène à la page du mini-projet où vous pouvez avoir : 1) Une explication du mini-projet, 2) Les packages nécessaires à installer, 3) Le jeu de données utilisé (dataset) et le lien de son téléchargement, 4) Le code source en Python

Travail à rendre

Les étudiants doivent préparer et envoyer à l'adresse : tahar.mehenni AT univ-msila.dz :

- a- Le code source du mini-projet (en pdf) (Attention : le code source qui existe sur le lien est très détaillé et peut contenir des instructions et des modèles qui ne sont pas demandés)
- b- Un rapport détaillé (en pdf) des étapes réalisées du mini-projet, depuis la préparation des données jusqu'à l'évaluation des résultats.

Evaluation du mini-projet

- Le mini-projet sera noté sur 8 points : code source (2 pts) + rapport (2 pts) + démonstration (4 pts)
- Une séance de démonstration des mini-projets sera programmée après les examens.

Liste des mini-projets

- 1- Détection des cartes de crédit frauduleuses par les arbres de décision:
<https://www.kaggle.com/code/denidiana/fraud-detection-credit-card-with-decision-tree/>
- 2- Détection des cartes de crédit frauduleuses par Naïve Bayes classification :
<https://www.kaggle.com/code/yuridias/credit-card-fraud-detection-knn-naive-bayes>
- 3- Prédiction des ventes sur Walmart par les arbres de décision :
<https://www.kaggle.com/code/cmcoutosilva/zedelivery-python>
- 4- Analyse des achats avec Apriori :
<https://www.kaggle.com/code/priy998/groceries-analysis-using-apriori-algorithm>
- 5- Analyse des achats avec Apriori :
<https://www.kaggle.com/code/mariekaram/apriori-association-rule>
- 6- Boston House Pricing with decision trees :
<https://www.kaggle.com/code/arvindshaw/boston-house-pricing-regression-machine-learning>
- 7- Boston House Pricing with decision trees :
<https://www.kaggle.com/code/nacerkroudir/boston-house-pricing>
- 8- Clustering des étudiants avec KMeans:
<https://www.kaggle.com/code/rodriguekalash/unsupervised-machine-learning-clustering-technique>
- 9- Clustering des étudiants avec KMeans:
<https://www.kaggle.com/code/varunnagpalspyz/student-performance-in-exams-with-kmeans>
- 10- Prédiction des maladies du cœur par Arbre de décision
<https://www.kaggle.com/code/chemistahmedkamel/predict-heart-disease-status>