

Projet Fouille de donnés

Le machine learning, ou apprentissage automatique, représente une branche de l'intelligence artificielle. Il consiste à entraîner des algorithmes pour qu'ils apprennent à partir de données et effectuent des prédictions ou des classifications sans être explicitement programmés pour chaque tâche. Aujourd'hui, l'essor des paiements numériques a généré d'immenses volumes de données, dont celles issues des transactions UPI (Unified Payments Interface) en Inde, un système qui facilite les transferts bancaires instantanés via mobile.

Dataset: <https://www.kaggle.com/datasets/skullagos5246/upi-transactions-2024-dataset>

Vous devez faire :

- Une analyse du jeu de données dans le but de déterminer les raisons d'un échec de transaction par rapport aux différentes caractéristiques. Une transaction est un échec lorsque son statut a la valeur "failed".
- Ensuite vous allez faire la prédiction de l'échec de la transaction en utilisant 3 modèles d'apprentissage supervisés. Déterminer laquelle de ces modèles est la meilleur et faites la prédiction de l'échec.
- Par la suite sur la base de la prédiction de l'échec, proposer 3 modèles de clustering permettant de déterminer 3 catégories différentes que vous avez vous-même définies.