

Mini Projet : Prédiction de la rentabilité des opérations journalières d'une banque X

Med Lemine **SIDIBBA**

Doctorant en Informatique

A. Objectif :

Nous procédons à une étude pratique sur un Dataset des transactions journalières d'une banque X, pour prédire la rentabilité en utilisant des techniques de Machine Learning.

B. Etapes du Projets :

IMPORTING LIBRARIES & DATASET

```
# import requirement libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
# import sedad dataset
data = pd.read_csv('chemin d'accès de dataset.csv')
df = pd.DataFrame(data)
df
```

CLEANING DATASET

Changement de type de quelques colonnes
Vérifier les valeurs doubles et celles manquantes

Supprimer les colonnes qui ne rentrent pas dans l'opération de prédiction de la rentabilité

EXPLORATORY DATA ANALYSIS (EDA)

CORRELATION INSIGHTS

DATA PREPROCESSING

CONSTRUCTION DE MODELES MACHINE LEARNING

Resultas

Modeles	accuracy	precision	recall	F1-score
LogisticRegression	0.9999301	0.9999625	0.9994780	0.999720
DecisionTreeClassifier	1.0	1.0	1.0	1.0
SVC	0.934586	0.7734354	0.5303765	0.5406748
GradientBoostingClassifier	1.0	1.0	1.0	1.0
KNeighborsClassifier	0.9886085	0.9477479	0.9628894	0.9551703
MLPClassifier	0.9931511	0.9963566	0.9488517	0.9712189

Conclusion

- Le modèle DecisionTreeClassifier et GradientBoostingClassifier ont obtenu les meilleurs résultats avec une précision de 100 %. Cela suggère que ces modèles ont pu bien généraliser à partir des données de formation et prédire avec précision les résultats sur les données de test.
- Les modèles LogisticRegression et MLPClassifier ont également montré des scores de précision élevés de 99,99625% et 99,63566% respectivement, indiquant qu'eux aussi étaient efficaces dans leurs prédictions.
- Le modèle de KNeighborsClassifier avait une précision décente de 94,77479%, mais il était surpassé par les autres modèles, à l'exception de SVC.
- Le modèle SVC avait la précision la plus faible, soit 77,34354%, ce qui suggère qu'il n'a peut-être pas été aussi efficace dans ses prédictions que les autres modèles.