

PROJET 19 : Détection de Fraude Carte de Crédit

OBJECTIF DU PROJET

Protéger les actifs financiers. Le défi est l'Aiguille dans une Botte de Foin (Données Déséquilibrées).

FICHIER DE DONNÉES

- fraude_carte_credit.csv
- 800 lignes
- Inclut des valeurs manquantes, valeurs aberrantes et déséquilibres pour un apprentissage réaliste

TÂCHES À RÉALISER

■ TÂCHE PRINCIPALE

Classer Classe (Fraude/Légitime) avec un Rappel Élevé

■ TÂCHES SECONDAIRES (OPTIONNELLES)

1. Gérer le Déséquilibre de Classe (SMOTE/Sous-échantillonnage)
2. Visualiser la distribution Fraude vs Légitime (t-SNE)
3. Définir un seuil "Coût-Bénéfice" (Coût du Faux Positif vs Faux Négatif)
4. Détecter les "Nouveaux Types" de fraude (Détection d'Anomalie Non Supervisée)

LIVRABLES ATTENDUS

1. ****Exploration des Données (EDA)****
 - Analyse statistique descriptive
 - Visualisations des distributions
 - Analyse des corrélations
 - Identification des valeurs manquantes et aberrantes
2. ****Ingénierie des Fonctionnalités****
 - Création de nouvelles variables pertinentes

- Transformation des variables existantes
- Encodage des variables catégorielles
- Normalisation/Standardisation si nécessaire

3. **Modélisation**

- Choix et justification du/des modèle(s)
- Entraînement et validation
- Optimisation des hyperparamètres
- Évaluation avec métriques appropriées

4. **Résultats et Interprétation**

- Performance du modèle principal
- Résultats des tâches secondaires (si réalisées)
- Insights et recommandations
- Visualisations des résultats

CONSEILS

- Commencez par une exploration approfondie des données
- L'ingénierie des fonctionnalités est CRUCIALE pour ce projet
- Documentez votre démarche et vos choix
- Comparez plusieurs approches si possible
- Pensez à l'aspect pratique et utilisable de vos résultats

Bon travail ! ■