


1. Quel est le problème ?
2. Pourquoi ce problème est-il important ?
3. Objectif du modèle
4. Compréhension des données
5. Nettoyage des données : Gestion des valeurs manquantes et outliers
6. Feature Engineering : Création de variables significatives
7. Analyse exploratoire : Identification des patterns de fraude
8. Entraînement de modèles : Classification binaire
9. Évaluation : Métriques de performance sur la détection de fraude



Personne 1 – Introduction & cadrage du problème

1. Quel est le problème ?

 Détection automatique des transactions frauduleuses

- La fraude est rare mais coûteuse
- Problème de **classification binaire**
 - 0 : transaction normale
 - 1 : transaction frauduleuse

2. Pourquoi ce problème est-il important ?

- Pertes financières élevées
- Risque pour les banques / clients
- Impossible de tout vérifier manuellement
- Erreurs possibles :
 - Faux négatifs  (fraude non détectée)
 - Faux positifs  (transaction normale bloquée)

Personne 2 – Objectif & compréhension des données

3. Objectif du modèle

- Prédire si une transaction est frauduleuse ou non
- Maximiser la détection des fraudes
- Réduire les fausses alertes

- Trouver un **bon compromis précision / rappel**

4. Compréhension des données

- Description du dataset
- Variable cible : fraude / non fraude
- Variables explicatives : montants, comportements, etc.
- **Fort déséquilibre des classes**

Personne 3 – Prétraitement des données

5. Nettoyage des données

- Gestion des valeurs manquantes
- Traitement des outliers
- Normalisation / standardisation si nécessaire

6. Feature Engineering

- Création de variables plus informatives
- Transformation des variables existantes
- Amélioration de la capacité prédictive du modèle

Personne 4 – Analyse exploratoire & modélisation

7. Analyse exploratoire (EDA)

- Comparaison fraude vs non fraude
- Identification de patterns :
 - montants atypiques
 - comportements suspects
- Visualisations

8. Entraînement des modèles

- Problème de classification binaire
- Test de plusieurs algorithmes
- Entraînement sur les données d'apprentissage

Personne 5 – Évaluation & conclusion

9. Évaluation des modèles

- Accuracy (avec prudence)
- Precision
- Recall
- F1-score
- Matrice de confusion
- Analyse du seuil de décision

Mini conclusion

- Objectif atteint
- Modèle performant
- Importance du choix des métriques
- Importance du seuil de décision