



Aperçu du projet

L'Organisation Nationale de Lutte Contre le Faux Monnayage est une organisation publique qui a pour objectif de mettre en place des méthodes d'identification des contrefaçons de billets en euros. Elle nous confie aujourd'hui la mission d'analyser les différences exploitables entre un vrai et un faux billet pour mettre en place une méthode automatique qui pourra ensuite être exploitée par une machine.

Réponse

Une analyse exploratoire de la base de données montre notamment une **différence exploitable** sur la valeur de la **marge inférieure** d'un billet.

Les données manquantes sur cette variable ont été prédites par une régression linéaire (explicitement demandé) afin d'optimiser l'algorithme final.

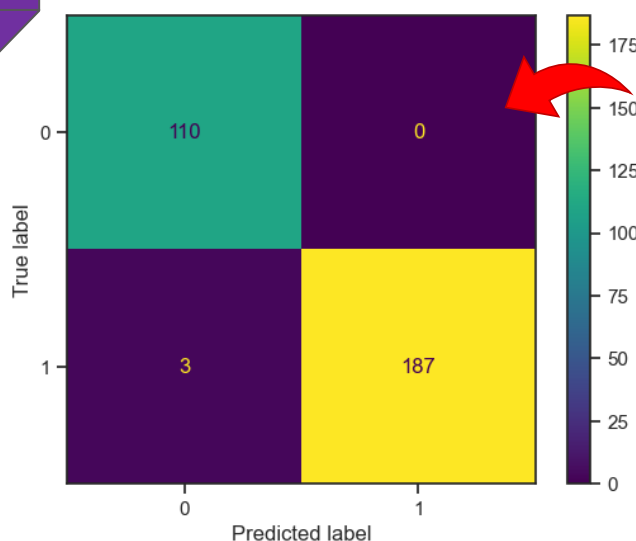
Le modèle a été entraîné **sans** l'information sur **la véracité d'un billet**, afin de le rendre plus robuste et dans des conditions **au plus proche de la situation réelle**.

Plusieurs modèles et plusieurs approches ont été abordées afin de comparer les modèles et de sélectionner le meilleur avant son **optimisation**. Le plus performant s'est révélé être la régression logistique.

Le modèle final a été réentraîné sur l'entièreté du jeu de données et est prêt pour des tests avant la mise en production.

Détails

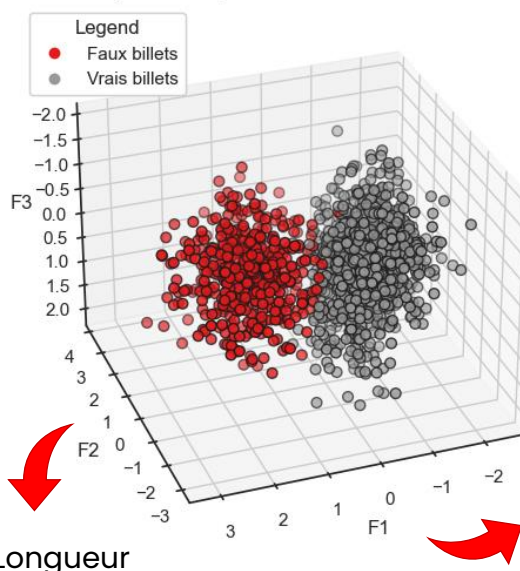
Matrice de confusion



L'optimisation du modèle lui a permis de ne laisser aucun faux billet (pas de faux positifs) sur l'échantillon de test

Différences clés

Projection des prédictions sur les 3 dimensions



Combinaison des marges (haut et bas) et des hauteurs (à gauche et à droite)