



Akademi Education

Prédiction du Churn des Clients Bancaires

Analyse et Modélisation du Comportement Client



Réalisé par: Riché FLEURINORD et Micka LOUIS
Sous la direction des membres du jury d'Akademi

24 octobre 2025

Présentation de l'équipe

02



Riché FLEURINORD

Data Scientist
Economist-Statistician
Financial Analyst

Parcours:
Akademi
CTPEA

University of pennsylvania (Wharthon Online)
INAGHEI(UEH)



Micka LOUIS

Data Scientist
Economist-Statistician
Accountant

Parcours:
Akademi
CTPEA
INAGHEI(UEH)



Vue d'ensemble du projet

1) Contexte du projet

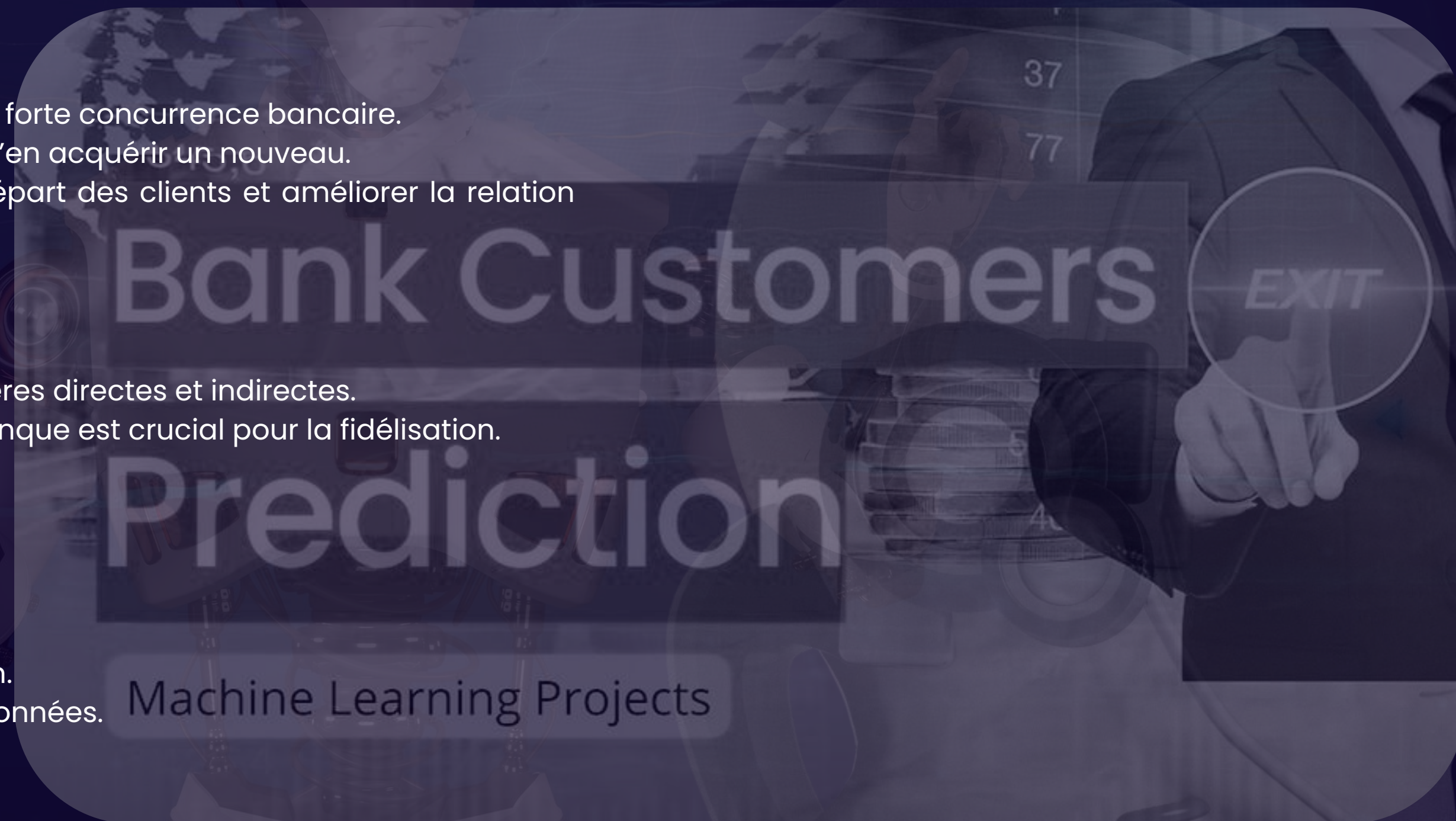
- La fidélisation client est essentielle face à la forte concurrence bancaire.
- Conserver un client coûte moins cher que d'en acquérir un nouveau.
- Projet de Data Science pour anticiper le départ des clients et améliorer la relation client.

2) Problématique du churn

- Chaque départ entraîne des pertes financières directes et indirectes.
- Comprendre pourquoi un client quitte la banque est crucial pour la fidélisation.

3) Objectifs principaux

- Prédire le risque de départ des clients.
- Identifier les facteurs clés du churn.
- Proposer des actions concrètes de rétention.
- Faciliter la prise de décision basée sur les données.





Akademi

Compréhension du Métier

1) Objectifs métier :

- Prédire le départ des clients avant qu'il ne se produise.
- Identifier les clients à risque pour des actions de fidélisation ciblées.
- Optimiser la rentabilité et la satisfaction client.

2) Questions clés :

- Quels clients présentent un risque élevé de churn ?
- Quels facteurs expliquent le départ des clients ?
- Quelles actions peuvent réduire le churn efficacement ?

3) Impact attendu :

- Meilleure allocation des ressources marketing.
- Réduction des pertes financières liées aux départs.
- Renforcement de la relation client et de la fidélité.

04



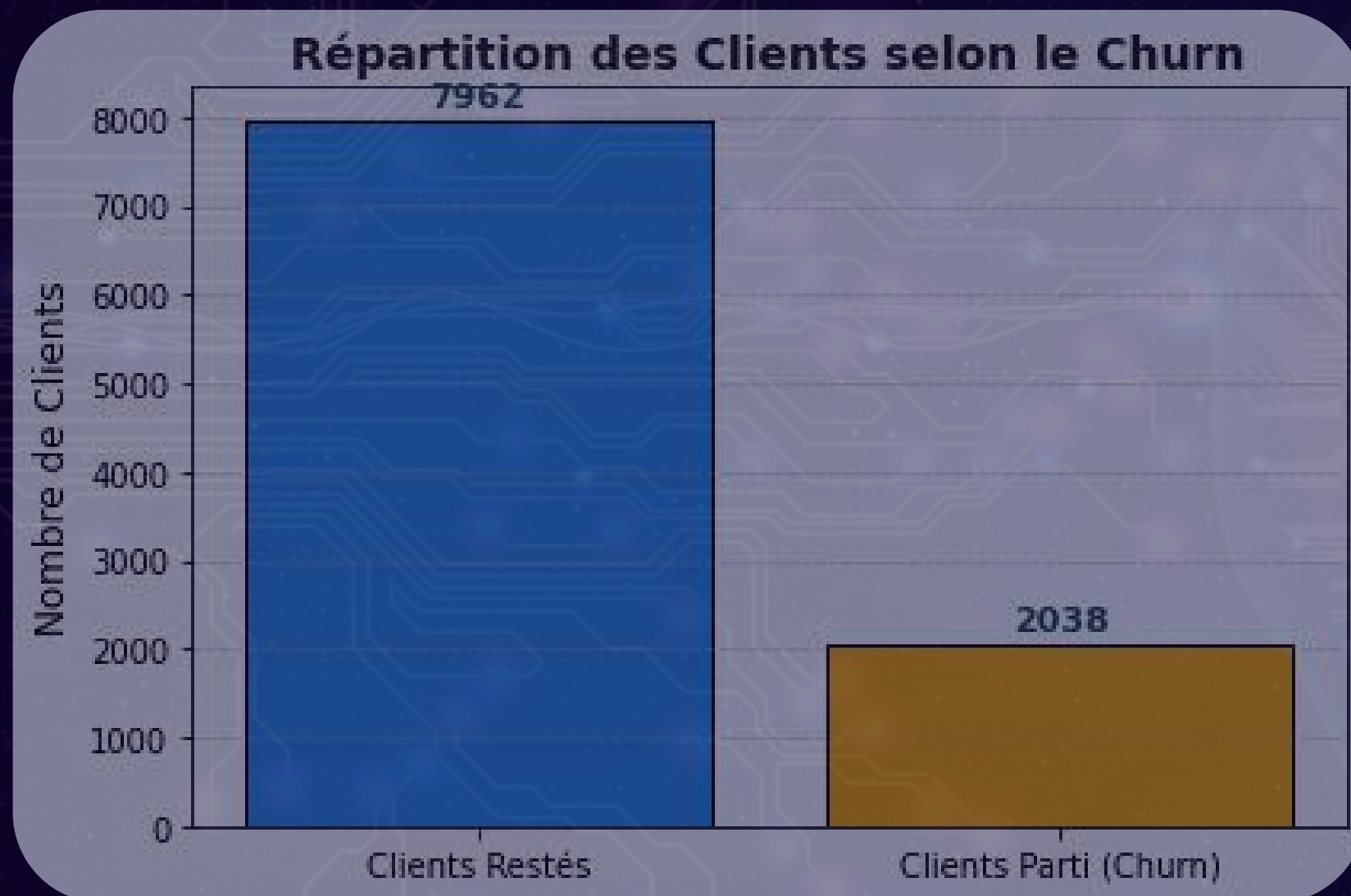
CRISP-DM (Étape 1)





Compréhension des Données

Ojectif: Explorer la structure, la cohérence et la qualité des données



1) Dataset :

- 10 000 clients, 14 variables
- Variables catégorielles : Surname, Geography, Gender, Card Type
- Variables numériques: Age, CreditScore, Balance, EstimatedSalary, NumOfProducts, Exited, etc.

2) Observations clés :

- Dataset propre : 0 valeur manquante, pas de doublons
- Exited (churn) : 20,4 %
- Géographie et genre équilibrés
- Variables numériques variées, certaines nécessitent attention (Balance, Age)

3) Conclusion :

- Dataset fiable et cohérent, prêt pour la préparation, l'EDA et la modélisation.



Compréhension de la Variable cible: Exited

1)**Définition:** représente le statut de départ d'un client.

- **Exited = 1:** le client a volontairement quitté la banque (procédure de clôture effectuée).
- **Exited = 0:** le client est toujours actif, sans démarche de départ.

2)**interprétation pratique:** la variable mesure un comportement décisionnel, lié à la fidélité ou à la rupture volontaire de la relation bancaire.

3)**Objectif du modèle:** prédire les clients susceptibles de quitter la banque, afin de prévenir le churn et renforcer la fidélisation.

Préparation des Données

Ojectif: Préparer un jeu de données propre, cohérent et exploitable pour la modélisation prédictive.



Étapes principales :

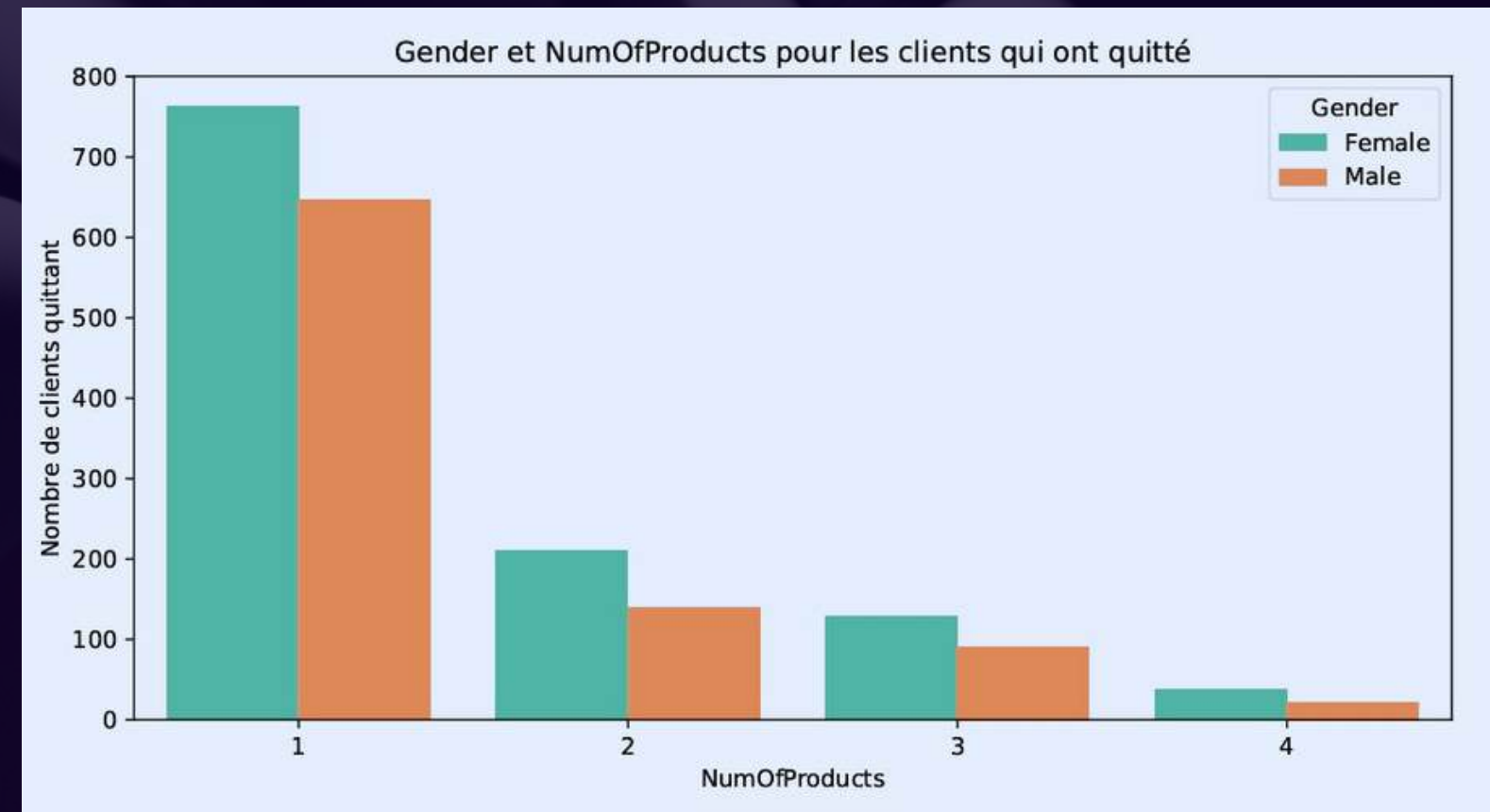
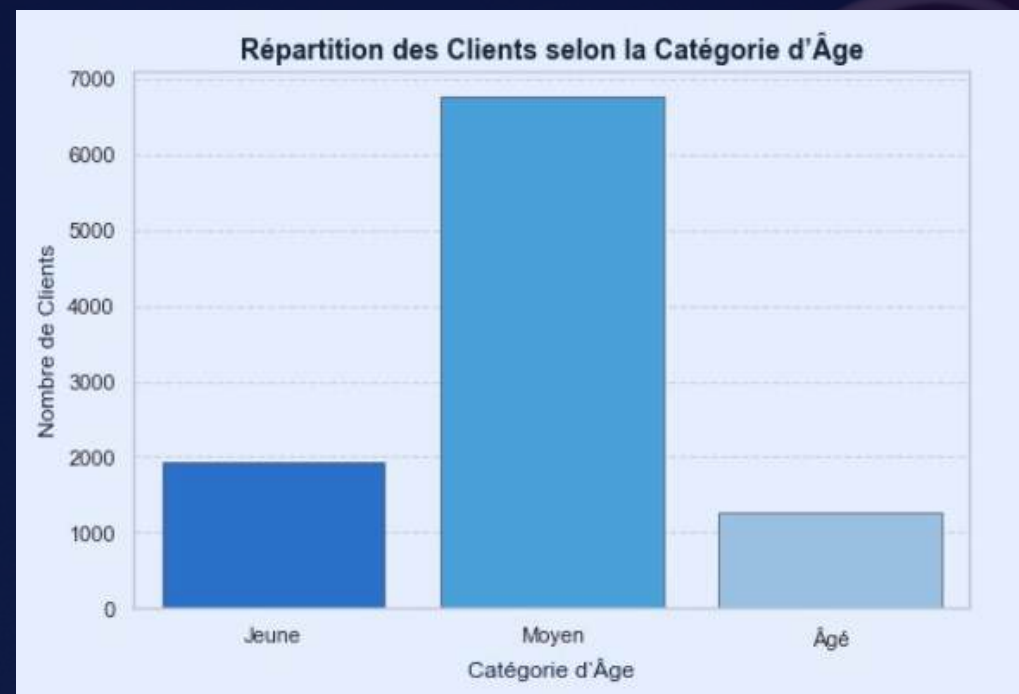
- Chargement & Vérifications.
- Nettoyage
- Feature Engineering
- Encodage
- Split des données
- Sauvegarde & Reproductibilité

Résultat final: Un dataset fiable, standardisé et prêt pour la modélisation.

Ce pipeline assure la qualité des données et la compatibilité du modèle lors du déploiement Streamlit.

Analyse Exploratoire des Données (EDA)

Analyse Statistique: Analyse univariée et multivariée



- La majorité des clients appartiennent à la catégorie d'âge "Moyen", supposant une forte concentration autour de la population active.
- Le churn est plus élevé chez les clients (hommes et femmes) détenant un seul produit, tandis qu'il diminue à mesure que le nombre de produits augmente.

Modélisation

Ojectif: Identifier le meilleur algorithme pour prédire le churn client à partir des données préparées.

Méthodologie

- Division des données
- Pipelines complets
- GridSearchCV + Validation croisée pour l'optimisation

Modèles testés

- Régression Logistique
- Random Forest
- XGBoost
- Neural Network (MLP)

Modèle Final : XGBoost

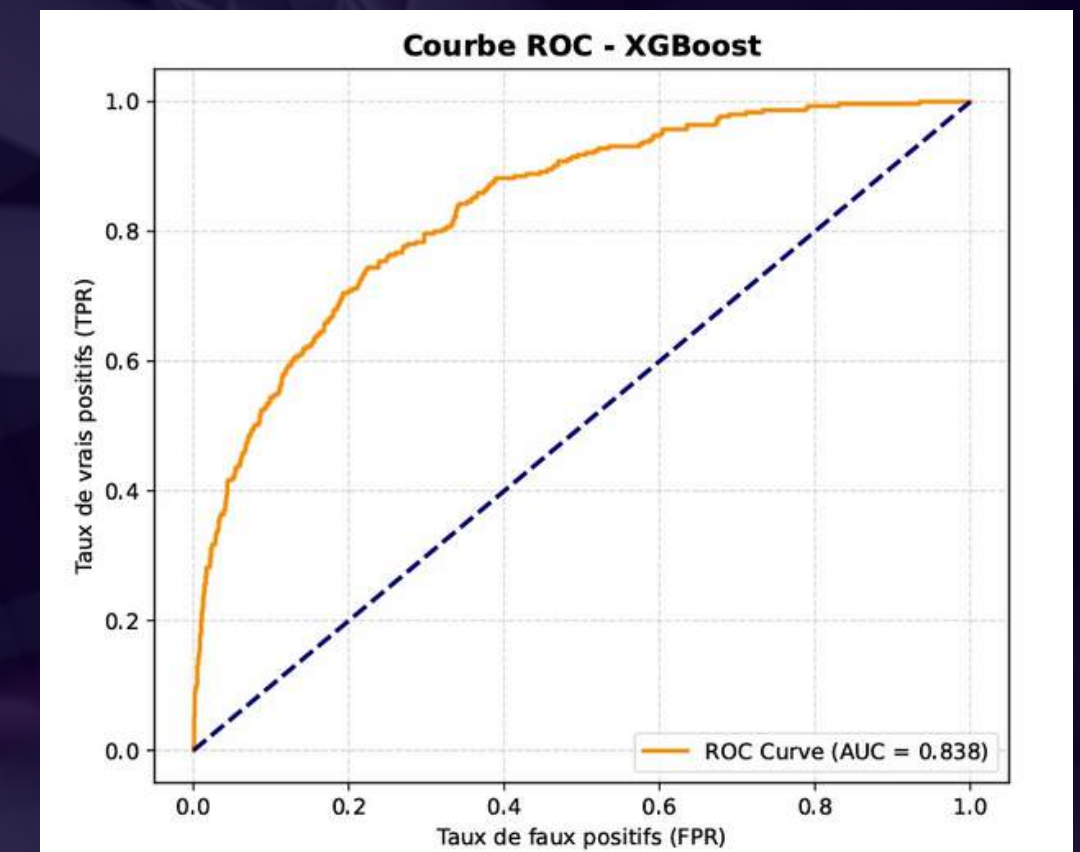
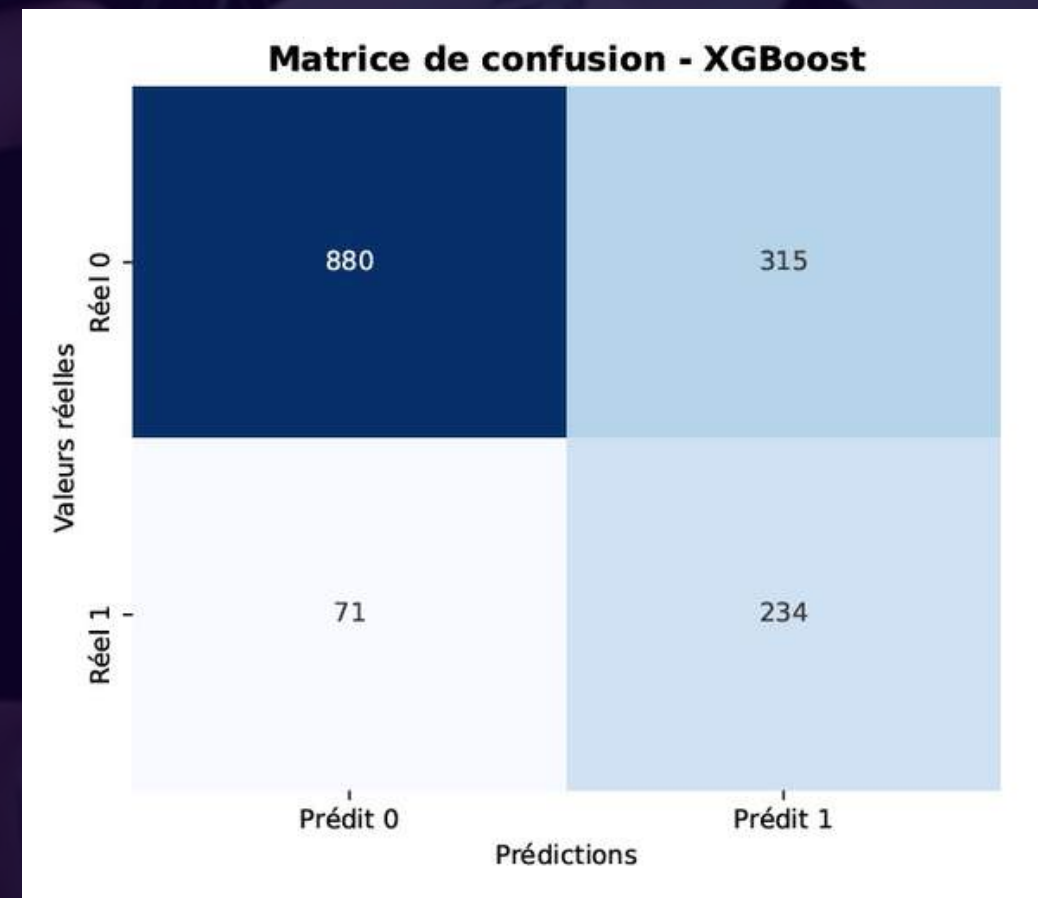
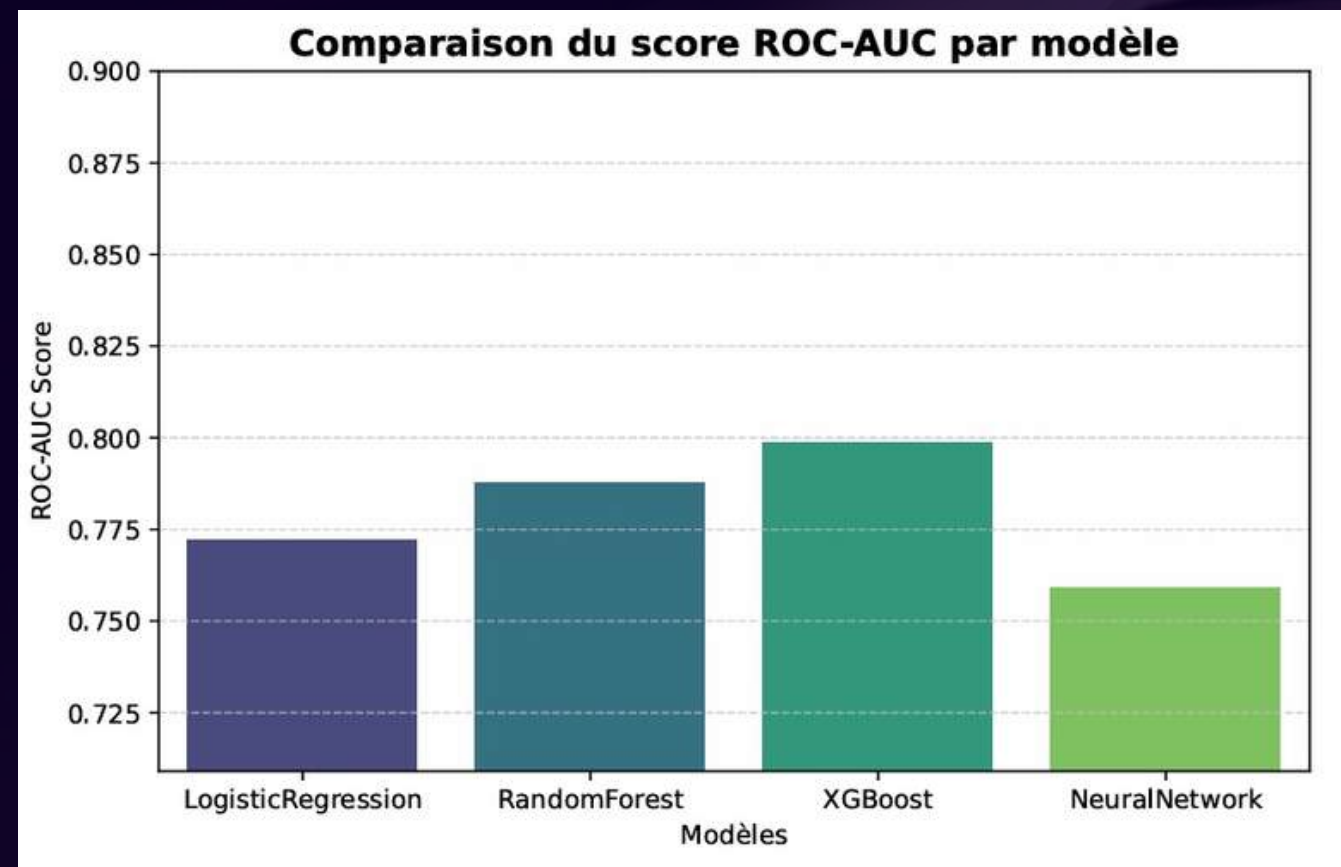
- Excellent compromis: précision/robustesse/ généralisation
- Meilleur score ROC-AUC = 0.7996
- Sélectionné pour le déploiement Streamlit



Évaluation du Modèle

Objectif de l'étape:

- 1) Vérifier la performance réelle du modèle sur des données non vues.
- 2) S'assurer qu'il est fiable, stable et utile pour la prise de décision.
- 3) Identifier le modèle offrant le meilleur compromis entre précision, rappel et robustesse.





Déploiement du Modèle XGBoost

Le déploiement marque la dernière étape du projet :

- Transformation du modèle XGBoost en une application web interactive avec Streamlit
- Utilisable par les managers pour prédire le départ des clients en temps réel



Client #1 : Jeune et stable financièrement, il présente un risque modéré de départ lié à une satisfaction moyenne

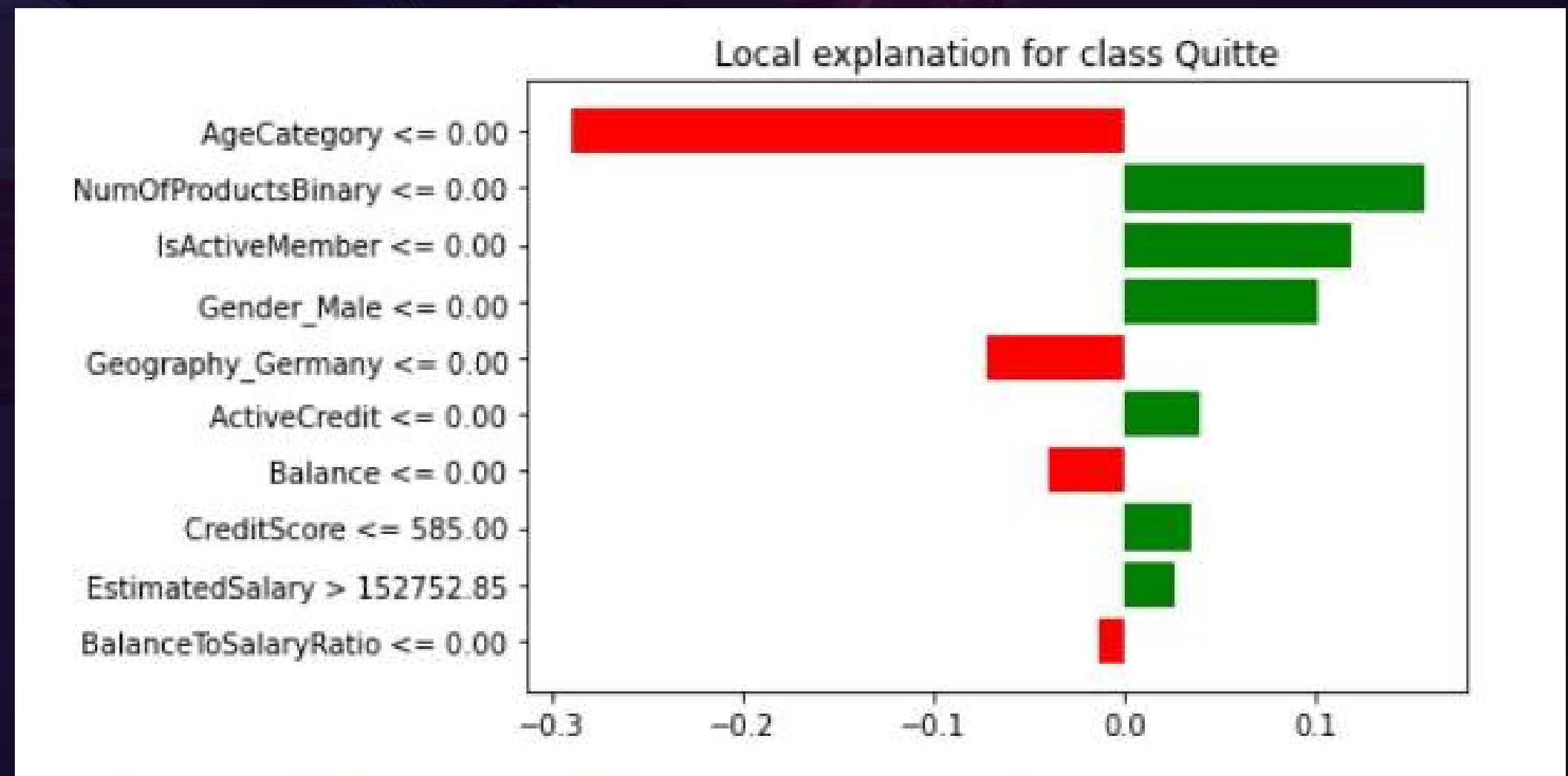
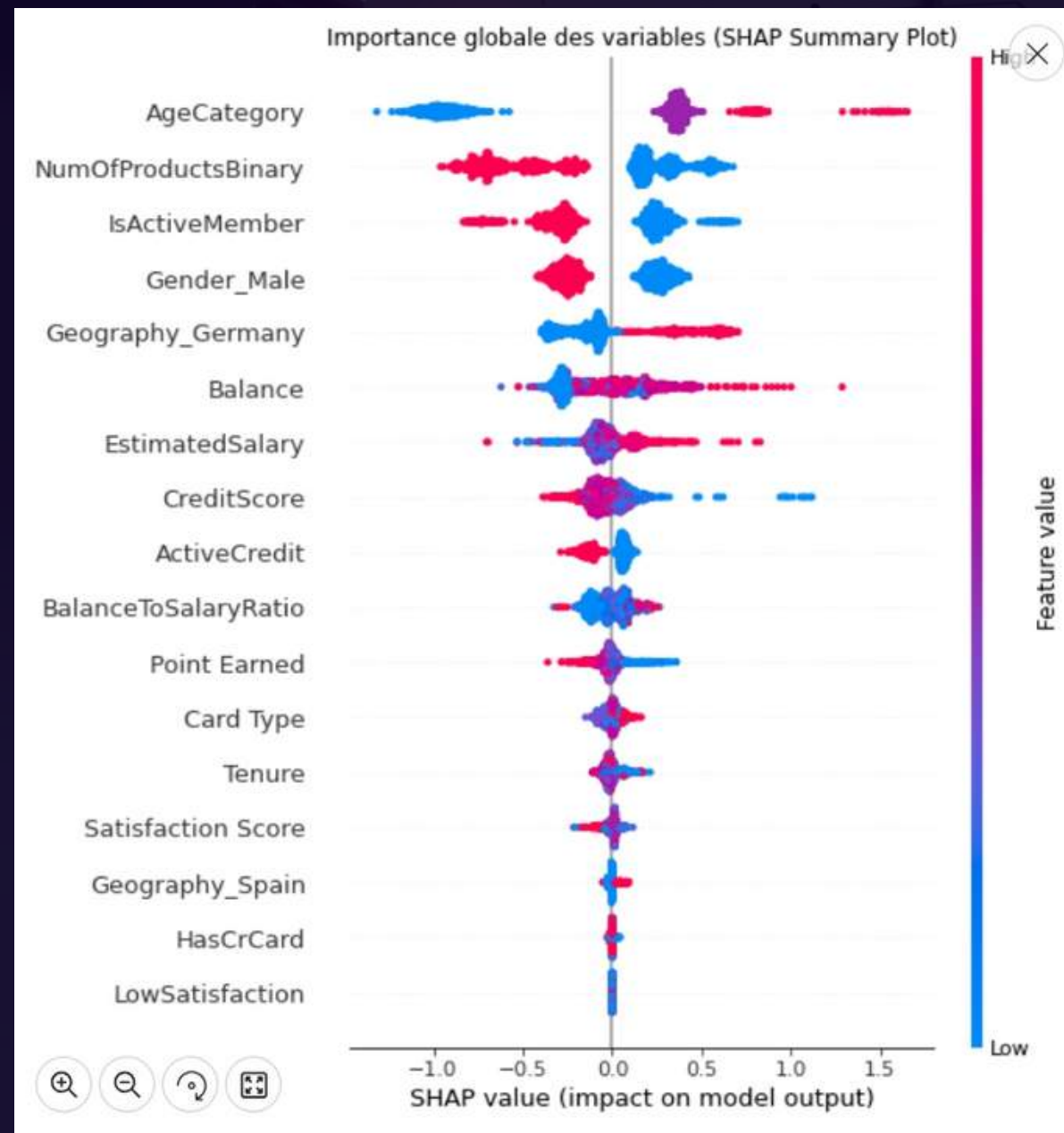
Client #2 : Âgé, fidèle et très satisfait, il affiche un risque très faible de churn.

Client #3 : Financièrement fragile et insatisfaite, elle montre un risque élevé de quitter la banque.



Explicabilité et Interprétabilité du Modèle

Objectif: Rendre le modèle XGBoost transparent et compréhensible pour les décideurs, en identifiant les facteurs expliquant la fidélité ou le départ des clients.





Akademi

Recommandations métiers

- 1- Fidélisation proactive des clients à risque de churn**
- 2- Segmentation ciblée et stratégies de cross-selling**
- 3- Optimisation des campagnes marketing et allocation des ressources**





Limites du Projet

- 
- 1– Absence de dimension temporelle**
 - 2– Taille restreinte du jeu de données**
 - 3– Manque de variables comportementales récentes**





Akademi

Étapes suivantes

1. **Suivi en production et monitoring du modèle**
2. **Actualisation et enrichissement des données**
3. **Optimisation et tests de modèles avancés**



Akademi

Merci!

Et si nous pouvions anticiper quels clients risquent de partir... pour mieux les garder ?



+509 44138336



+509 46153943



richefleurinord20@gmail.com



mickalouis25@gmail.com

https://github.com/richefleuriord/Bank_Customer_Churn_Prediction.git