

Risque de défaut: Robustesse face à la dérive de modèles

1 – INTRODUCTION

L'estimation de la probabilité de défaut (PD) permet aux institutions financières d'évaluer le risque de non-remboursement d'un prêt. Cependant, les modèles prédictifs peuvent perdre leur précision au fil des années en raison de l'évolution des données et des conditions économiques.

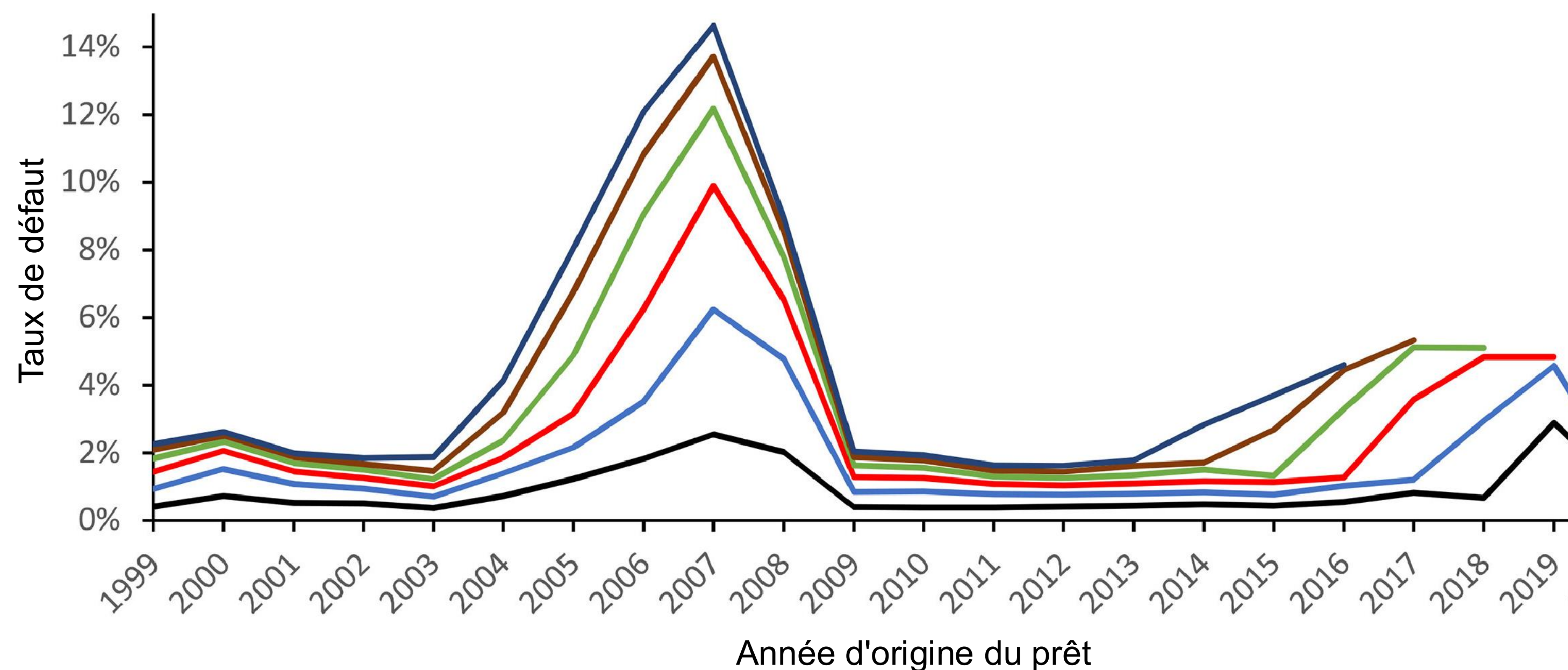


Figure 1: Incidents de paiement sur les crédits américains de 1995 à 2019 [1].

Buts:

- Comparer différents modèles (apprentissage automatique);
- Évaluer leur robustesse face à la dérive des modèles;
- Identifier ceux conservant une stabilité de leur performance face à la dérive.

2 – MÉTHODOLOGIES EXPÉRIMENTALES

Jeu de données

Le projet s'appuie sur un vaste jeu de données de l'entreprise américaine Freddie Mac contenant les informations sur plus de 50 millions de prêts hypothécaire sur une période allant de 1999 à 2021 [1].

Le jeu utilisé est un sous-ensemble du jeu complet évaluant l'état du prêt 12 mois après l'initialisation sous-divisé en 3 sous-espaces de temps:

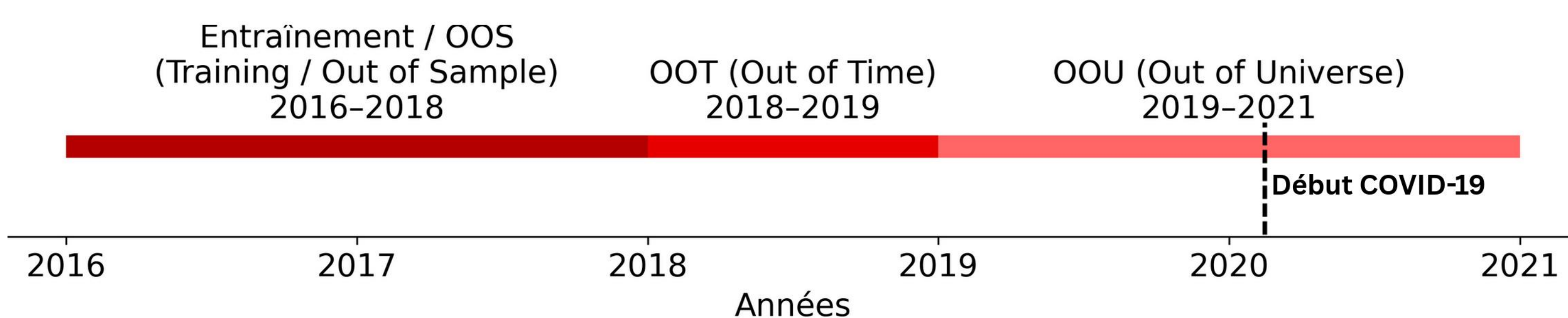


Figure 2: Ligne du temps illustrant la distribution des jeux de données dans le temps.

Exploration

Les défauts de paiements sont très faiblement représentés parmi l'ensemble des données.

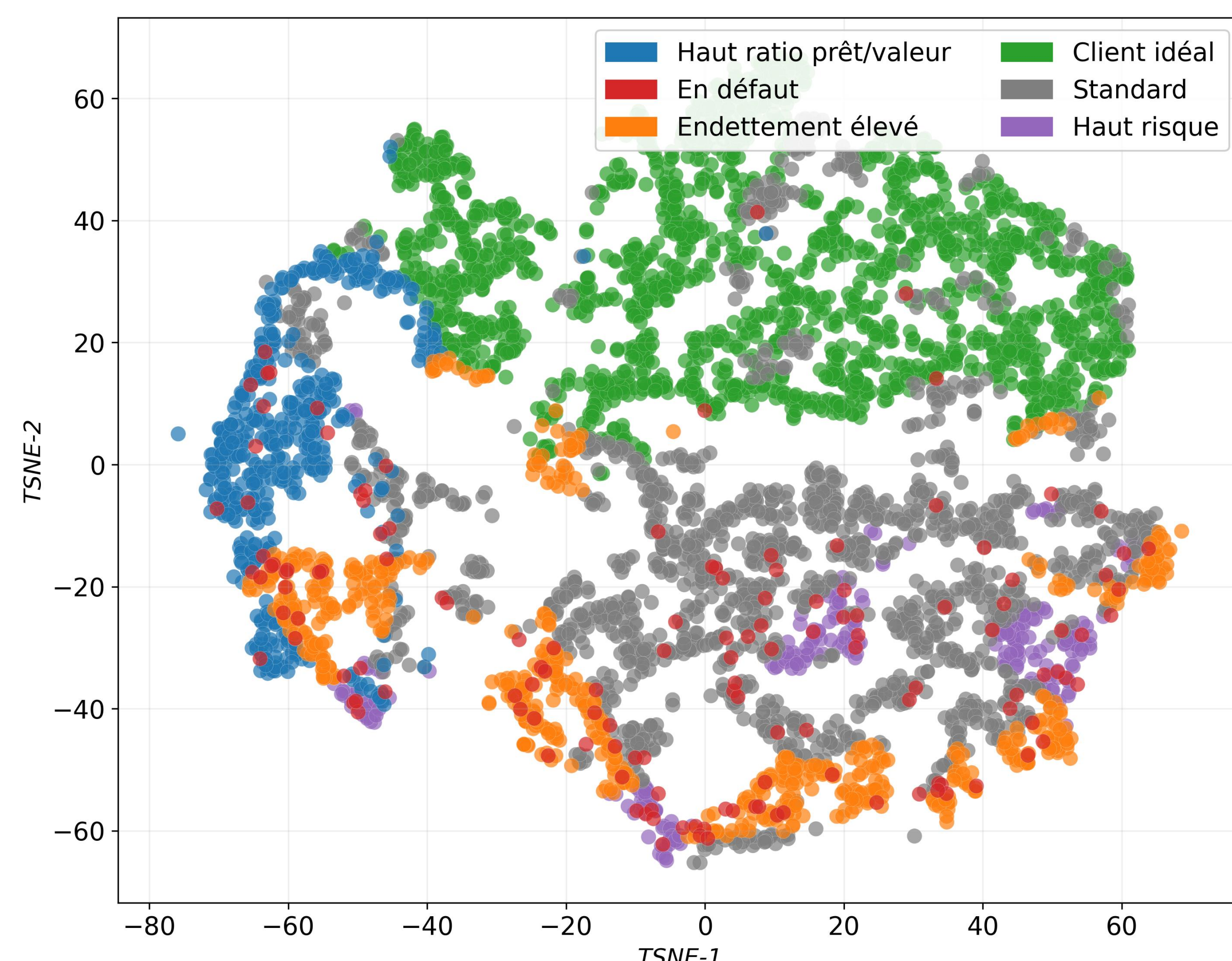


Figure 3: Représentation t-distributed Stochastic Neighbour Embedding (t-SNE) des données.

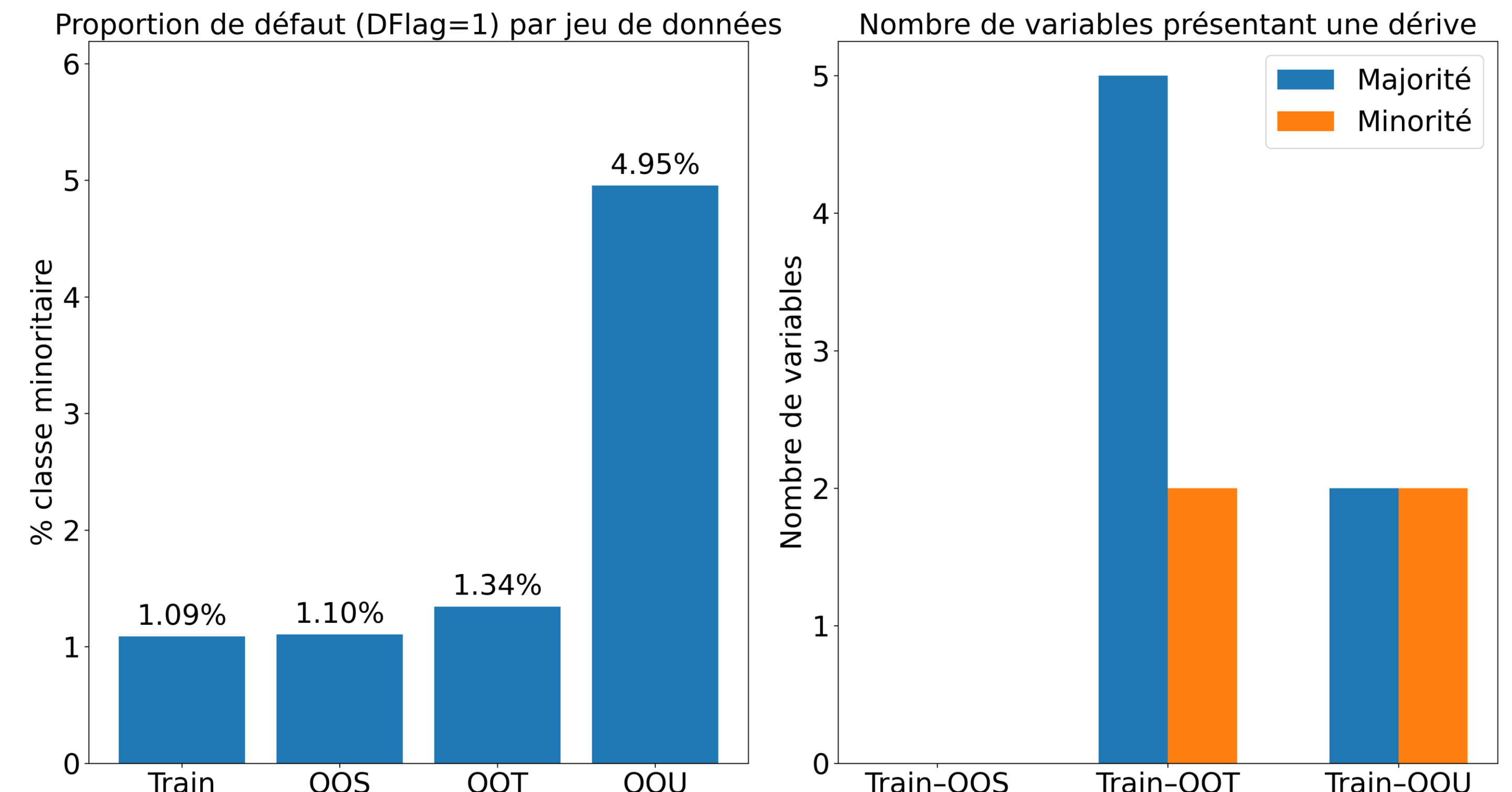


Figure 4: Évolution de PR-AUC et Gini au cours des différentes périodes temporelles sous-échantillonnées pour les modèles de Régression Logistique, XGBoost et Random Forest.

Prétraitement

Les données subissent un prétraitement de base, incluant une normalisation et une imputation des données incomplètes. Le sous-échantillonnage et le sur-échantillonnage se sont avérés inefficaces en prétraitement pour pallier le problème de déséquilibre des classes.

Modèles évalués

- Régression Logistique
- Random Forest
- XGBoost

Métriques

- Gini : standard en prédiction de défaut de crédit
- PR-AUC: combine précision et rappel

3 – ANALYSE DE RÉSULTATS EXPÉRIMENTAUX

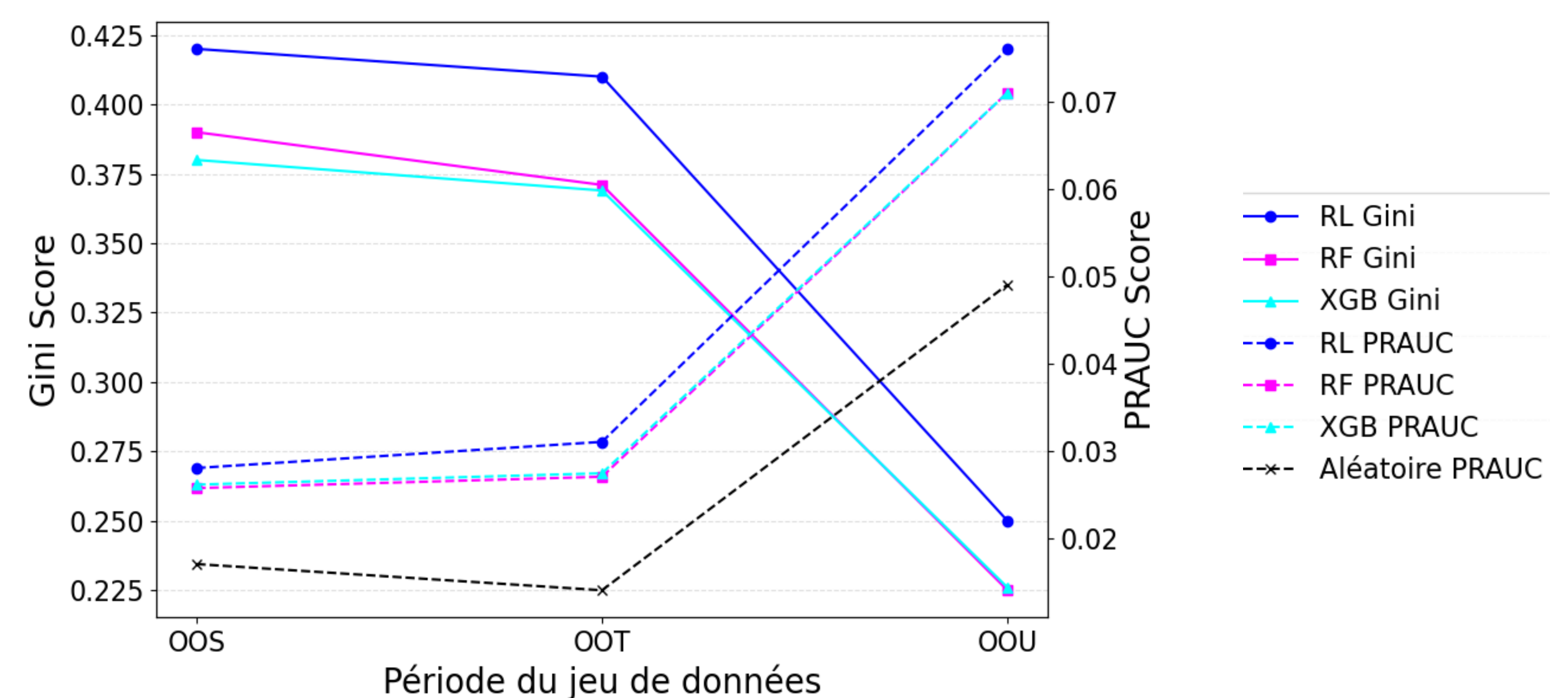


Figure 5: Distribution des défauts par jeu de données et nombre de variables avec dérive entre jeux, selon la classe.

Classement des modèles par robustesse:

#1 Régression Logistique

#2 XG Boost

#3 Random Forest

4 – DISCUSSION ET CONCLUSION

Est-ce que les algorithmes d'apprentissage automatique sont vraiment appropriés pour la prédiction du défaut de crédit?

- Initialement substantiellement meilleures que les méthodes standards;
- Renouvellements itératifs nécessaires pour préserver des performances stables;
- Aspect boîte noire augmente l'importance de la robustesse du modèle.

Travaux futurs

- Répéter sur différents jeux de données (périodes, type de prêt, emplacement géographique);
- Prédiction de la durée de vie utile d'un modèle avant obsolescence;
- Meilleur explicabilité des résultats.

5 – RÉFÉRENCE

[1] Mushava, Jonah (2023), "Credit scoring with class imbalance data: An out-of-sample and out-of-time perspective.", Mendeley Data, V3, doi: 10.17632/bzr2xttvz.3