



Implémentez un modèle de scoring

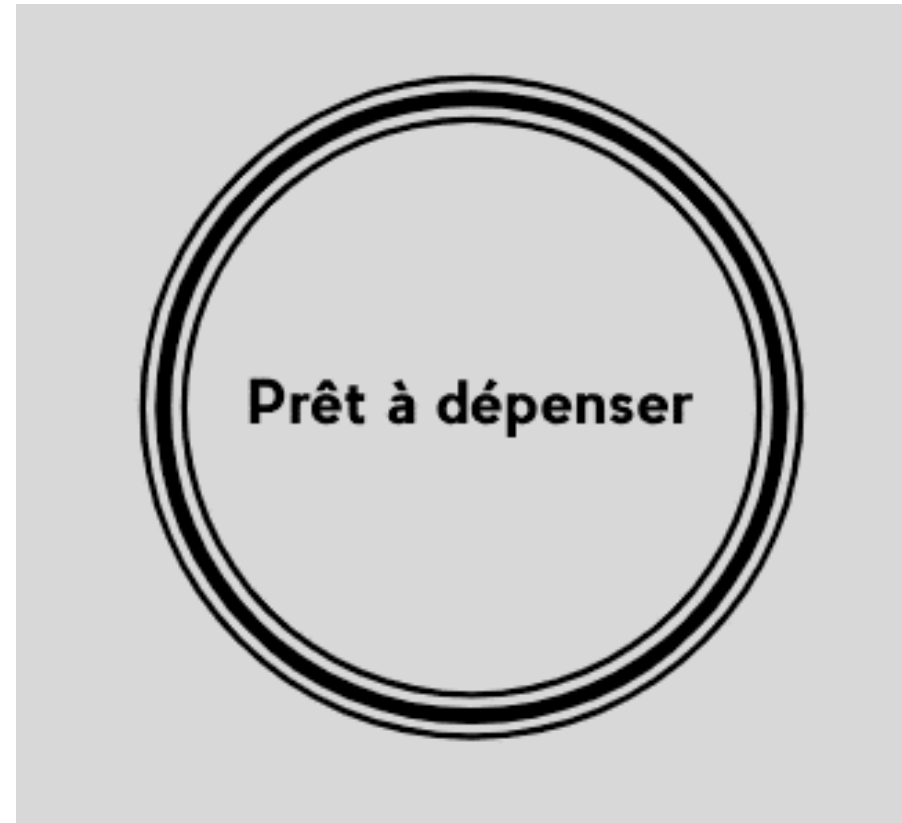
Octobre 2021

Sommaire

- Rappel de la problématique
- Exploration et analyse des données
- Explication de l'approche de modélisation
- Démonstration Dashboard
- Conclusion

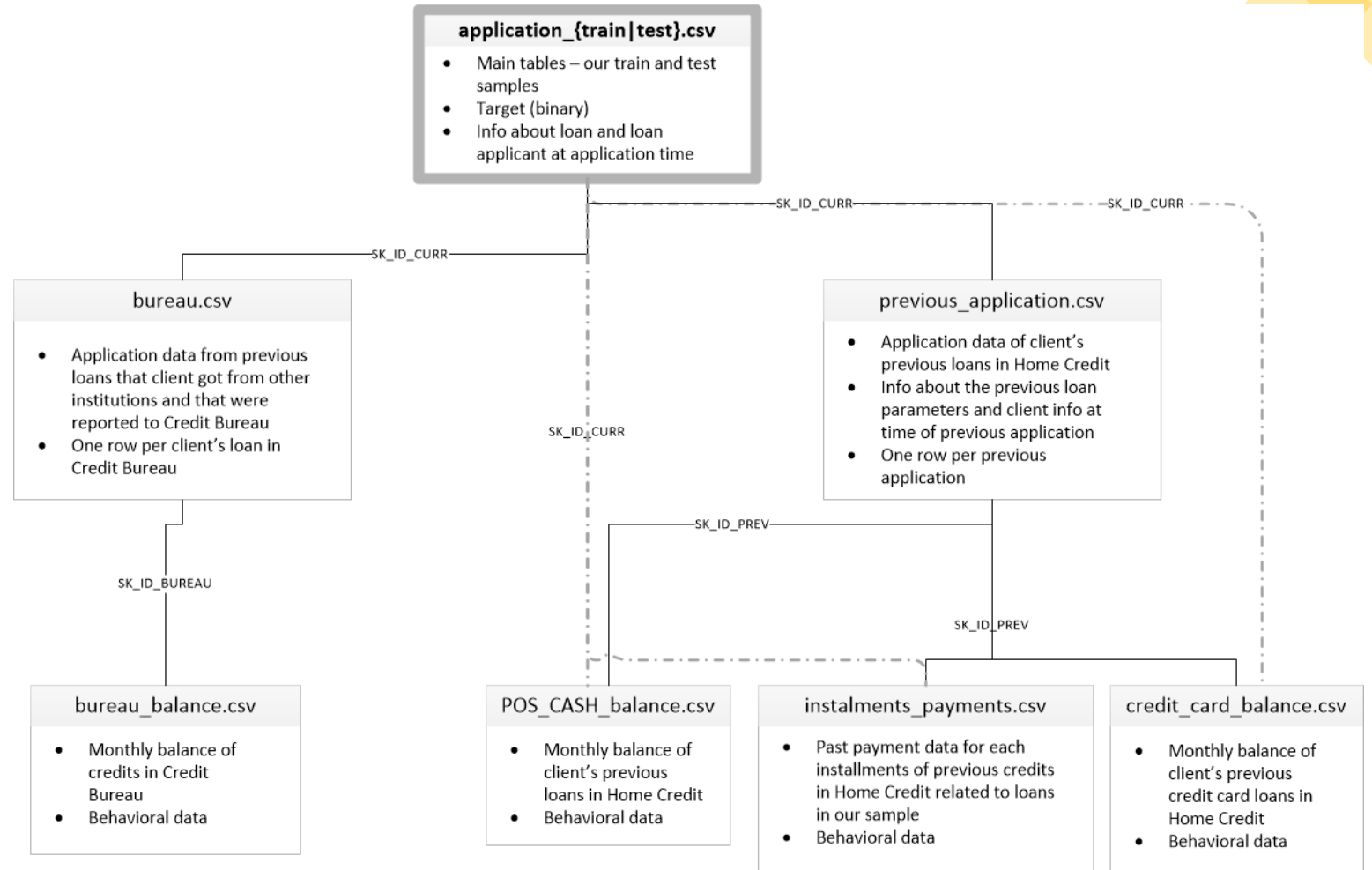
Rappel de la problématique

- Développer un modèle de scoring de la probabilité de défaut de paiement du client
- Développer un dashboard interactif
 - Visualisation du score
 - Visualiser les informations du client
 - Comparer les informations du client avec les autres applicants



Exploration et analyse des données

- 9 Datasets
- Crédit de biens de consommation (POS)
- Pas d'information de devise



Variable cible

		True Labels	
		Repaid	Defaulted
Model Prediction	Will Repay	True Repaid	False Defaulted
	Will Default	False repaid	True Defaulted

- Variable cible déséquilibré
- 1 = Clients à défaut, 0 = Clients Réguliers
- **Les erreurs du modèle ont des conséquences inégaux**

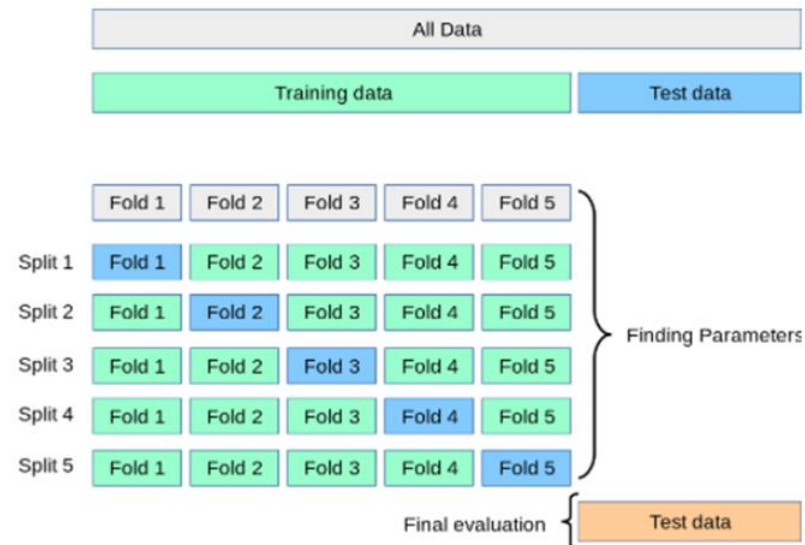


Explication de l'approche de modélisation

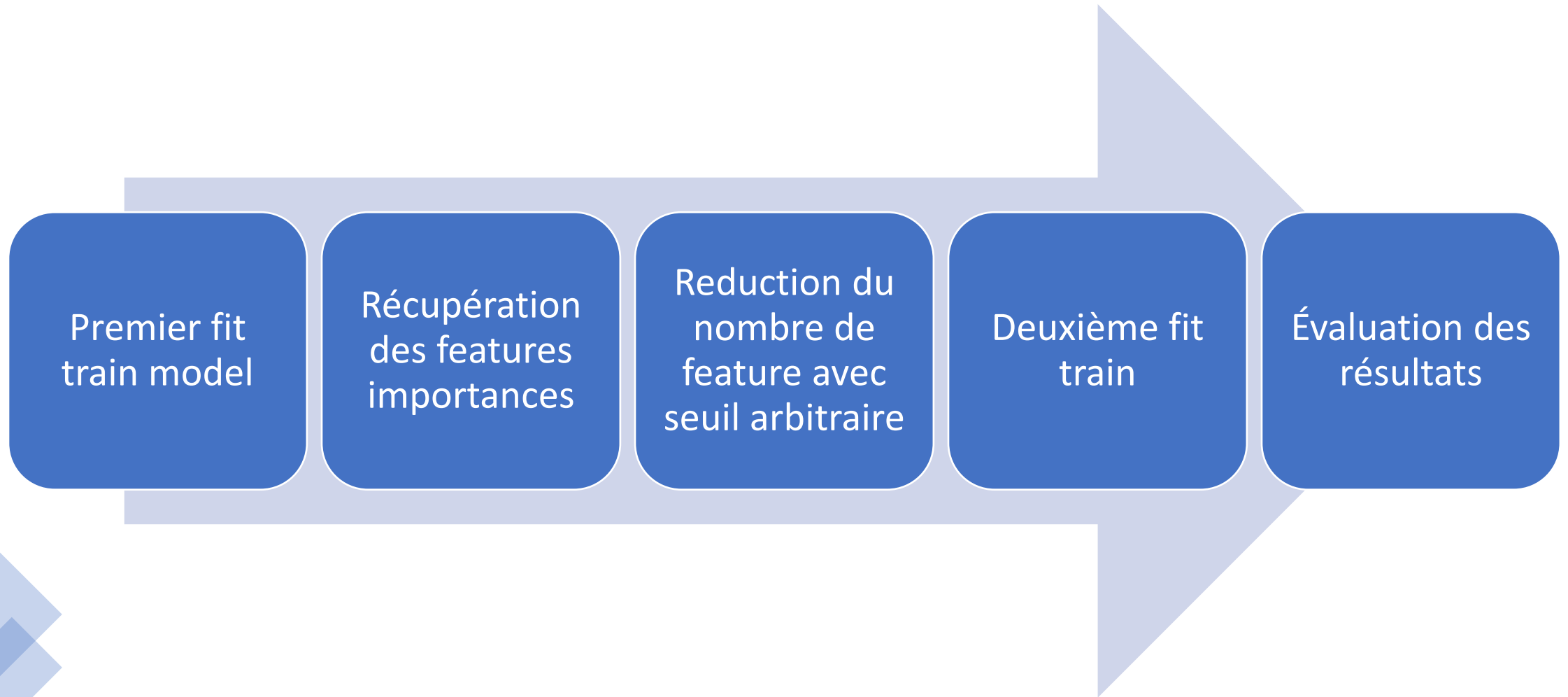


Notebook kaggle

- Feature engineering avec tous les datasets
 - Ratio, min, max, sum, size
- Model → LGBM classifier + kfolds
 - Eval_metric = 'auc'
- Modification effectué
 - Reduction du jeu de donné en appliquant des « bins » en fonction du crédit demandé.
 - Reduction du nombre de features.



Pipeline du model





Démonstration Dashboard

Conclusion

- Model rudimentaire, score proche des champions du concours
- Emphase mit sur les éléments évalués du projet
 - Démonstration d'une fonction pour maximiser les profits
 - Interprétabilité et fonctionnalité du Dashboard
- Considération pour mise en production, amélioration possible