



Implémentez un modèle de scoring

Projet 7 du parcours
« Data Scientist » d'OpenClassrooms

Mark Creasey

Sommaire

Implémentez un modèle de
scoring

- 01 La problématique
- 02 Les données
- 03 Modélisation
- 04 Dashboard
- 05 Conclusion

01 Présentation de la problématique

Mission – Implémenter un modele de scoring

La société financière « Prêt à dépenser » propose:

- de **crédits à la consommation**
- pour les personnes ayant **peu ou pas de tout d'historique** de prêt

Basée sur les données financières

- Données internes
- Données externes

Critères de succès

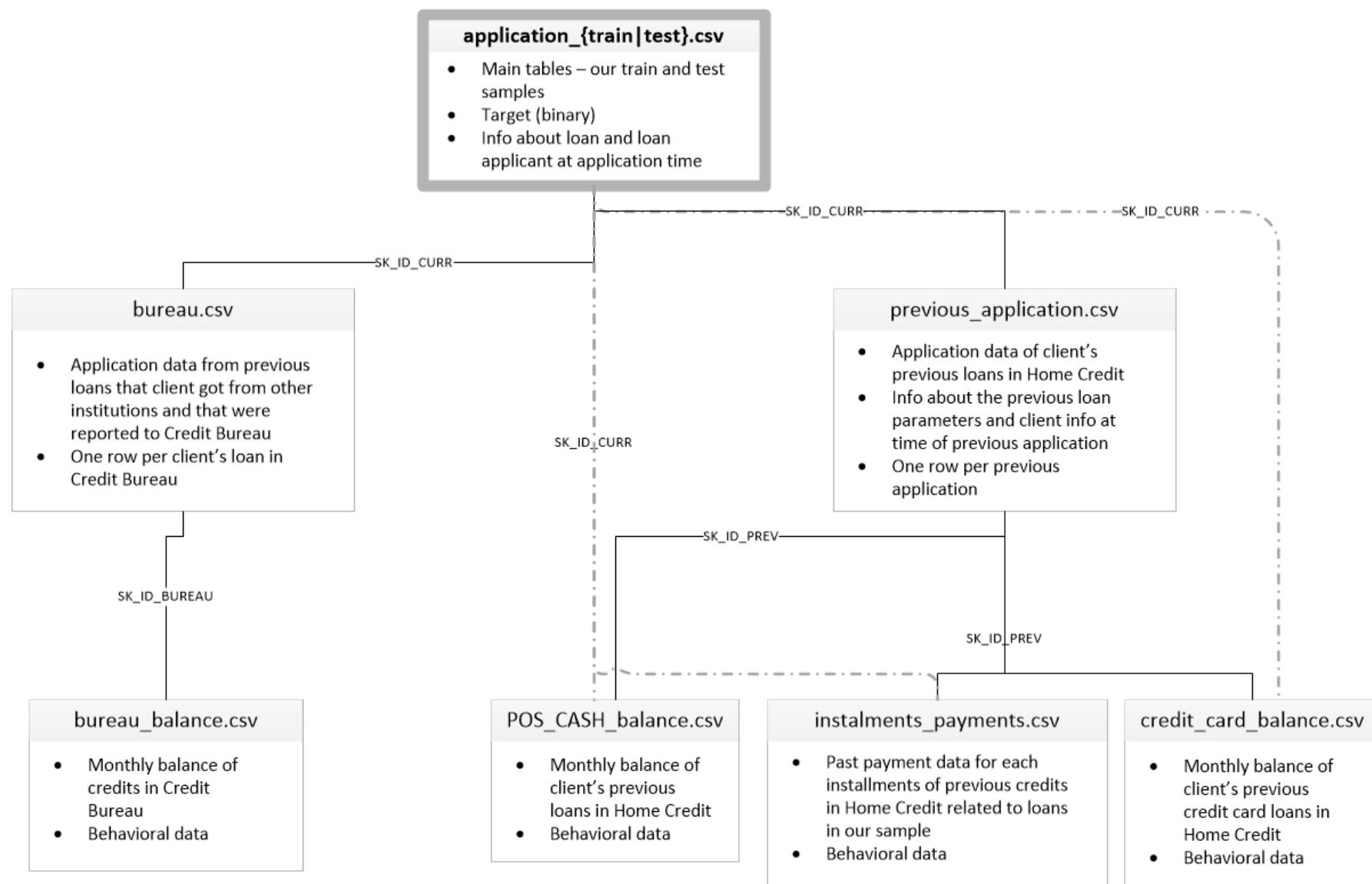
- Transparence de la décision sur l'octroi du crédit
- Deploiement d'un dashboard permettant de visualiser les informations clients pour interpréter la décision fait par la modèle



02 Les données

Nettoyage, exploration

Données financières sous 7 tables



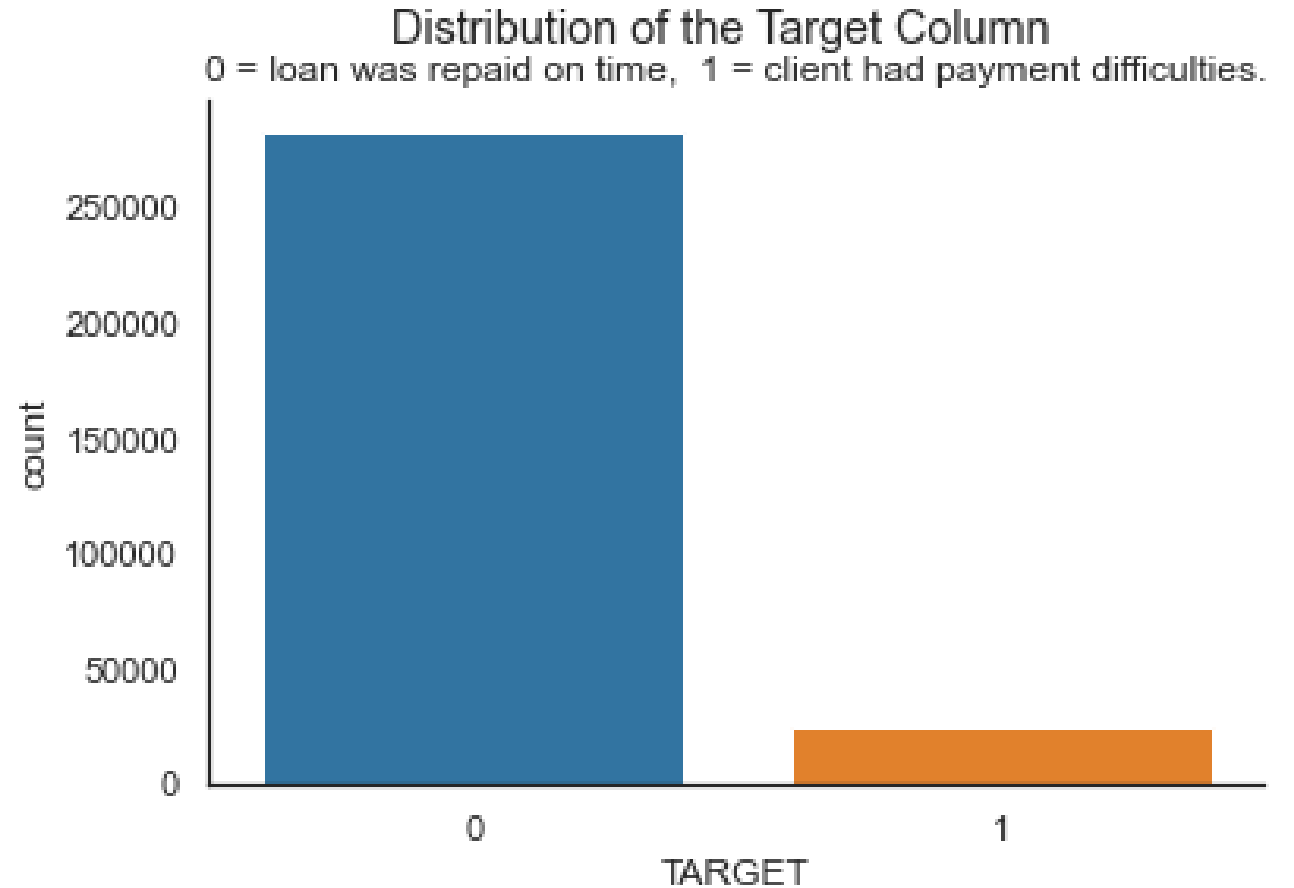
Distribution de cible très déséquilibré

Prédiction que tous les clients sont bons

→ précision de 93%

→ on aura identifié aucun client défaillant.

- 0 – clients sans problème de paiement
- 1 – clients défaillants



03. Modélisation

Sampling : Re-équilibrage des classes

- Aucun
 - conséquences
- Cost- sensitive
- Random under/over sample
- SMOTE
- SMOTE Tomek links

Métriques d'évaluation

- Precision
- Recall
- F1 Score
- F(beta) score
- ROC AUC

La fonction cout metier

Les modèles (classifiers)

Pipelines imblearn

- Preprocess
- Sampling
- Feature Selection
- Classification

Classifiers

- Dummy
- Ridge
- LogisticRegression
- RandomForest
- LightGBM

Evaluation des modèles

LGBM (weight-balanced) - Summary.

best params={'clf__class_weight': 'balanced', 'clf__max_depth': 6, 'clf__min_child_samples': 50}

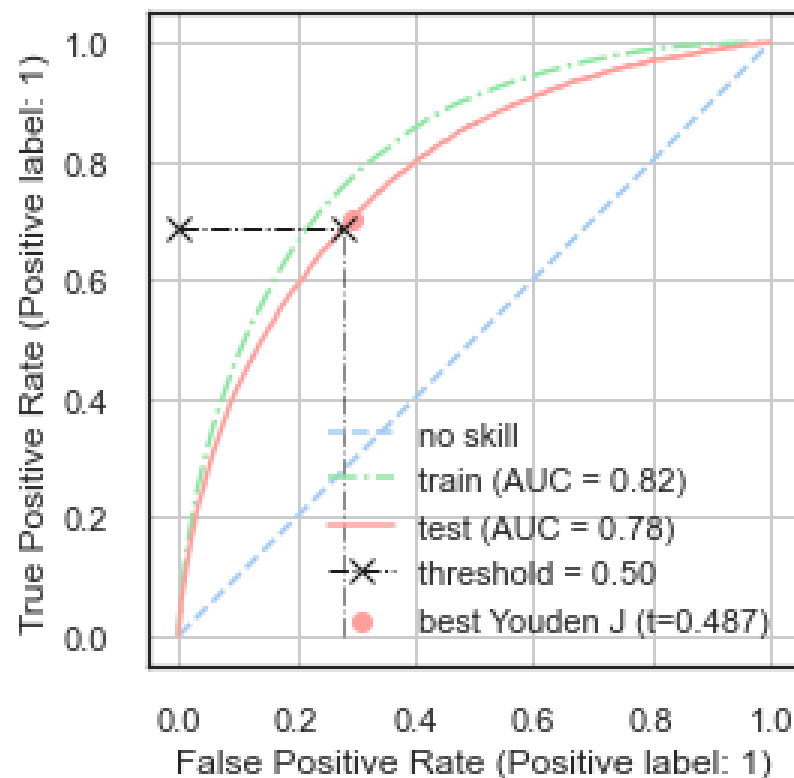
Confusion Matrix

True label	Predicted label	
	0	1
0	0.72	0.28
1	0.31	0.69

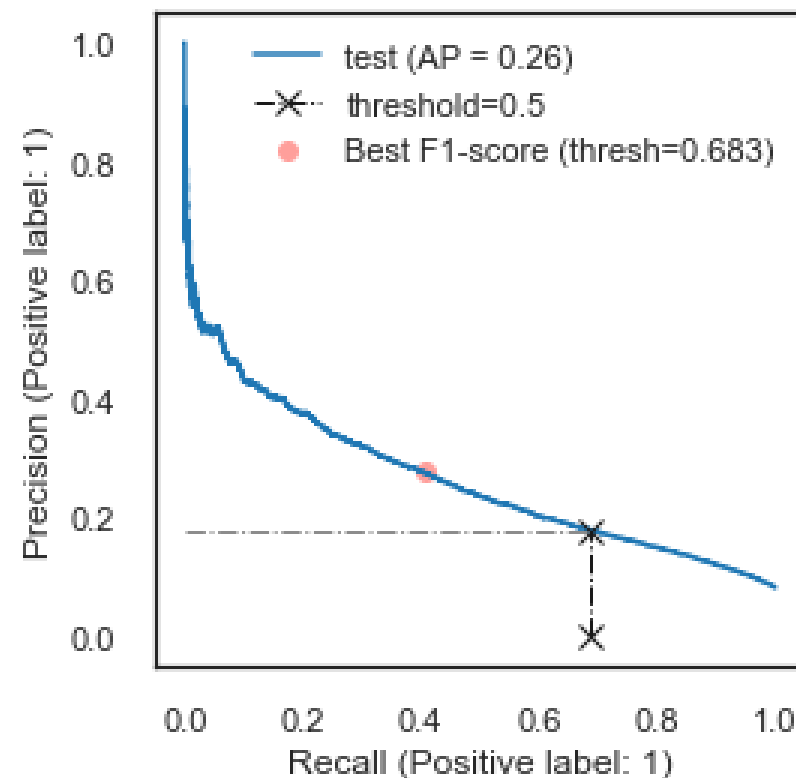
counts

True label	Predicted label	
	0	1
0	40891	15646
1	1561	3404

ROC Curve



Precision - Recall Curve



Optimisation du seuil (discrimination threshold)

Feature Importances

SHAP Interpretabilité

04 Dashboard

API du dashboard

- REST API
- Query parameters

Visualisation du dashboard

- Gauge
- lien vers dashboard deployment

05 Conclusion et améliorations à faire

Conclusions

Améliorations à faire

Questions

images: Mark Creasey

- mrcreasey@gmail.com
- Merci !