

Note technique pour le projet de scoring (projet 7)

Introduction :

Ce projet de scoring est un projet de machine learning qui vise à développer un modèle de prédiction de score pour évaluer la solvabilité d'un client potentiel. Le but est de fournir une estimation précise de la capacité de remboursement d'un client et de minimiser le risque de non-remboursement.

Méthodologie :

Le modèle de scoring est basé sur un algorithme de machine learning supervisé. Nous avons utilisé un ensemble de données historiques de clients pour entraîner notre modèle. Les caractéristiques des clients, telles que leur âge, leur revenu, leur historique de crédit, leur niveau d'endettement, ... sont utilisées pour créer des variables explicatives. Le score de crédit est utilisé comme variable cible pour la formation du modèle.

Les données historiques ont été divisées en deux ensembles, un ensemble d'entraînement et un ensemble de validation. L'ensemble d'entraînement a été utilisé pour entraîner le modèle, tandis que l'ensemble de validation a été utilisé pour évaluer la performance du modèle.

Nous avons utilisé certains algorithmes de classification à savoir l'algorithme de régression logistique, le random Forrest classifier et le light gradient boosting machine classifier (LGBM) comme modèles. Nous avons effectué également une analyse des variables importantes pour comprendre l'impact de chaque variable explicative sur le score de crédit.

Une fois que le modèle a été entraîné, nous l'avons appliqué à un ensemble de données de test pour évaluer sa performance en termes de précision et de rappel. Nous avons ajusté ensuite le modèle en fonction des résultats de l'analyse des performances.

Conclusion :

Ce projet de scoring est un outil précieux pour évaluer la solvabilité d'un client potentiel. Nous avons utilisé l'algorithme de LGBM classifier pour créer un modèle de prédiction de score en utilisant des données historiques de clients. L'analyse des variables importantes nous a permis de comprendre l'impact de chaque variable sur le score de crédit. L'application du modèle à un ensemble de données de test nous permettra d'évaluer la performance du modèle et de l'ajuster en conséquence.