

Implémentation d'un modèle de Scoring

Description projet

CONTEXTE DU BESOIN

Une banque commerciale ou tout type d'institution financière qui fournit des services tels que l'acceptation de dépôts, l'octroi de prêts aux entreprises et l'offre des produits d'investissement de base doivent prendre des risques en accordant des prêts et des cartes de crédits aux clients. Ceci est délicat car leur survie est liée à la prise de risques appropriés. Ainsi, lorsque les banques sont confrontées à un risque, une gestion appropriée avec de bonnes stratégies doivent être mise en place. L'évaluation du risque de crédit est vitale pour les banques ; ils doivent s'assurer que les emprunteurs sont en mesure de payer leurs échéances avant de leur octroyer un prêt. En effet, lorsque le client fait défaut, la banque subit des pertes du principal et des intérêts, l'interruption des flux de trésorerie et une augmentation des coûts de collecte.

La gestion du risque de crédit est l'un des processus fondamentaux des banques visant à évaluer, mesurer et contrôler les risques découlant de la Probabilité de Défaut en remboursements de prêts. Ce processus affecte leur stratégie et les approches opérationnelles concernant la diversification du portefeuille de prêts et la tarification, l'allocation la plus efficace du capital, l'évaluation du capital et les exigences de liquidité, les conditions commerciales politiques d'octroi de crédits aux clients, etc. D'où le besoin de développer un système d'évaluation interne qui analyse le risque d'un emprunteur.

Pour pallier ce besoin rencontré par notre client, la volonté est de développer une solution de gestion de risque performante visant à prédire la probabilité de faillite d'un client afin d'aider dans la prise de décision.

OBJECTIF

L'objectif est de mettre en œuvre un outil de "scoring crédit" pour calculer la probabilité qu'un client rembourse son crédit, puis classifie la demande en crédit accordé ou refusé. La solution consiste à développer un algorithme de classification en s'appuyant sur des sources de données variées (données comportementales, données provenant d'autres institutions financières, etc.).

TACHES A EFFECTUER

- Prétraitement des données (nettoyage, doublons, valeurs manquantes)
- Feature engineering (gestion des variables numériques et catégoriques, variables pertinentes, data leakage)
- Etude de l'équilibre des classes de la variable Target (SMOTE par exemple)
- Entraîner un modèle de classification (plusieurs modèles dont un simple de référence)
- Evaluer les performances (métriques adaptées, score métier bien défini (seuil=0.5), optimiser les hyperparamètres,)
- Visualisation (score, les informations descriptives des clients, etc.)