
PROJET BANQUE

Contexte

Alors que vous travaillez pour une grande banque Française (CAC 40, etc...) votre supérieur hiérarchique compte mettre en place une offre de prêt automobile. Et pour garantir le succès de cette offre, il compte contacter les 10 000 personnes les plus appétentes à cette potentielle offre. Il vous demande donc de vous sortir la liste de ces 10 000 personnes.

Il vous faudra donc mettre en place un score d'appétence au prêt automobile. Pour cela, vous disposez d'une base comprenant près de 44 000 individus, exprimées sur 138 variables. La variable cible sera ici : `P_P_auto`. Un lexique est à disposition pour préciser la nature de chaque variable.

Traitement de la base

Il faudra naturellement préparer la base de données pour ce score. Voici quelques idées de traitement :

- Analyser les corrélations avec la variable cible pour enlever les variables qui seraient logiquement corrélées. Aidez vous du lexique pour cela pour déterminer si les corrélations sont induites ou non.
- Réfléchissez bien en termes de banque/assurance, quels sont les personnes qui ne seraient pas susceptibles de prendre un prêt automobile, ou à qui la banque n'a pas envie d'accorder un prêt automobile (ce genre de prêt se fait généralement sur 10-15 ans, on peut par exemple imaginer des contraintes d'âge, etc... n'hésitez pas à vous documenter sur les conditions d'acceptation d'un crédit auto pour préciser vos critères)
- N'hésitez pas à équilibrer votre jeu d'apprentissage pour optimiser les performances de votre modèle !

Après avoir entraîné votre modèle sur une partie de la base, il faudra le relancer sur l'ensemble des individus qui ont P_P_auto à 0 et sélectionner les 10 000 personnes ayant la proba d'appétence la plus élevée !