

Méthodologie de projet d'analyse prédictive en 7 étapes

Appliquez cette méthodologie en sept étapes à n'importe quel cas pour répondre à une question simple : quel sera le comportement futur ? Dans le cas présent, nous l'appliquons à la détection des fraudes.



Étape 1 : Définir le problème :

Cette étape est cruciale. Veillez à penser aux données qui corroborent le problème, et pas seulement aux données dont vous disposez. Expliquez clairement l'analyse cible entre ceux qui posent la question et ceux qui y répondent.

- **Cas de fraude** : comment déterminer si un paiement par carte de crédit peut être frauduleux ou non ?



Étape 2 : Collecte des données :

Une autre étape cruciale consiste à collecter les données qui corroborent l'étape 1. Les données sont les racines et la sève d'un projet d'analyse prédictive. C'est ici que toutes les données nécessaires sont collectées.

- **Cas de fraude** : collectez toutes les données relatives aux transactions, qu'elles aient été reconnues comme frauduleuses ou non



Étape 3 : Préparer les données :

Nettoyer, enregistrer, modéliser, transposer et analyser : ce ne sont là que quelques-unes des actions qui peuvent être effectuées sur les données pour les préparer à l'analyse prédictive.

- **Cas de fraude** : les données d'une carte de crédit sont nettoyées pour être analysées



Étape 4 : Effectuer une analyse statistique et une ingénierie des caractéristiques :

Les données sont analysées afin de tester directement les hypothèses développées précédemment ou d'extrapoler des informations en visualisant les différentes mesures.

- **Cas de fraude** : étude des valeurs aberrantes, analyse des réponses et équilibrage des données



Étape 5 : Créer des modèles prédictifs :

Après avoir soigneusement préparé les données, les différents modèles prédictifs peuvent être testés et les expériences nécessaires réalisées pour obtenir un modèle aux capacités prédictives satisfaisantes. Le test et la validation des modèles peuvent être effectués à l'aide de différentes techniques (p. ex. en divisant les ensembles de données en ensembles de formation et ensembles de test).

- **Cas de fraude** : trouver le meilleur modèle pour prédire la fraude



Étape 6 : Évaluer les modèles :

La phase d'évaluation des performances des modèles, qui dépend des caractéristiques spécifiques de la variable à prédire, est également très importante.

- **Cas de fraude** : affiner et évaluer les performances des modèles



Étape 7 : Mise en production :

Une fois que tous les tests nécessaires ont été effectués, que la qualité des modèles a été évaluée et que les données de sortie ont été validées, le modèle d'analyse prédictive peut enfin être mis en production et utilisé.

- **Cas de fraude** : une méthode validée pour détecter les transactions frauduleuses

Vous avez des données. Nous avons les solutions. Imaginez les possibilités.

Discutez dès maintenant avec
un membre de l'équipe Minitab !