Détection de la fraude basée sur l'intelligence artificielle et les bases de données orientées graphe

Tabary, Edouard, edouard.tabary@bnpparibas-pf.com, BNP Paribas Personal Finance, Bordeaux (Orateur)

Cavarroc, Julie, julie.cavarroc@bnpparibas-pf.com, BNP Paribas Personal Finance, Bordeaux (Orateur)

Thématique : Analyse des données : Prévisions, statistiques, corrélations, Machine Learning, …

Résumé

BNP Paribas Personal Finance (BNPP PF) propose des produits de paiement fractionné à un certain nombre de partenaires. L'un de ces produits, un paiement en 3 ou 4 fois sur le e-commerce 100% dématérialisé, engendre un montant significatif de fraude. Les personnes ayant l'intention de frauder interchangent certaines informations personnelles afin de contourner les règles et les listes noires mises en place.

Pour contrer ce phénomène, BNPP PF a construit un dispositif innovant autour d'une **base de données orientée graphes** combinée à un **score en Machine Learning**.

Mots clés : Modèle de score – Fraude – Machine Learning – Base de données graphes – Innovation – Temps réel

1. Introduction

BNPP PF propose des produits de paiement fractionné à un certain nombre de partenaires. Cette solution peut par exemple prendre la forme d'un paiement en 3 ou 4 fois sur des sites de e-commerce en France pour des montants d'achat qui peuvent varier entre 90€ et 3000€. Sur ce périmètre, aucun justificatif n'est demandé, les clients remplissent uniquement un formulaire en ligne. Plusieurs milliers de demandes de financement sont reçues chaque jour, le processus est 100% dématérialisé et les clients reçoivent une réponse quasiment instantanée.

Un montant significatif de fraude est observé annuellement pour ce produit. En effet, étant donné le niveau limité de contrôle inhérent à ce processus sur l'identité des clients, ceux qui ont l'intention de frauder interchangent certaines informations personnelles afin de contourner les règles et les listes noires mises en place. BNPP PF a construit un dispositif autour d'une base de données orientée graphes combiné à un score en Machine Learning pour contrer ce phénomène.

2. Méthodologie

La base de données graphe est constituée :

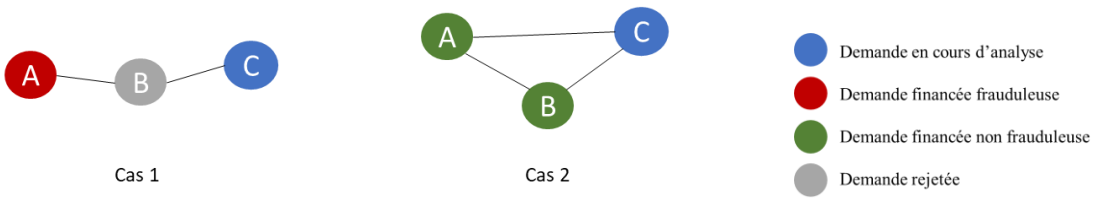
- De sommets qui représentent les demandes de financement reçues et qui contiennent les informations contenues dans les dossiers clients (quelques dizaines de champs)
- D'arêtes, qui permettent de relier les sommets entre eux dès lors que l'on considère que les demandes proviennent d'un même client (ou du même foyer)

Les arêtes sont déterminées à partir de règles qui peuvent utiliser l'ensemble des informations disponibles sur la demande reçue (par exemple, deux dossiers ayant un même numéro de téléphone seront connectés par une arête).

En temps réel, lorsqu'une nouvelle demande est reçue, celle-ci est intégrée dans la base de données graphe (un nouveau sommet est créé ainsi que de nouvelles arrêtes pour la connecter aux sommets préexistants). Selon la manière dont cette nouvelle demande se connecte aux dossiers antérieurs elle pourra être refusée ou acceptée.

Ce dispositif est très contraint en temps de calcul car l'analyse du dossier et la réponse doivent être faites de manière quasi-instantanée sur le site du partenaire e-commerce. Les algorithmes développés doivent donc retourner une prédiction en quelques millisecondes pour ne pas dégrader l'expérience d'achat du client.

Illustration :



Le dossier « C » est une demande de financement en cours d'analyse.

Dans le **cas 1**, on observe un chemin entre « C » et un dossier frauduleux « A ». De plus, « A » et « C » ne sont pas connectés et ne partagent par conséquent aucun élément d'identité commun. Il est donc probable que le dossier « C » soit une demande frauduleuse provenant d'un client qui cherche à dissimuler son identité.

Dans le **cas 2**, « C » est connecté de manière directe à deux dossiers non-frauduleux. De plus, les trois dossiers sont interconnectés deux à deux. Il est donc probable que la demande provienne d'un client honnête qui a fait plusieurs achats en ligne. On voit à travers ces exemples simplifiés que le graphe nous renseigne sur la qualité des dossiers reçus à la fois par sa topologie et par la nature des nœuds qui le composent.

Face au volume de données utilisées, les variables extraites du graphe combinées aux variables collectées lors de l'octroi sont analysées avec un modèle Machine Learning basé sur une forêt aléatoire.

Le modèle de score a été développé sur une plateforme interne à BNPP.

3. Originalité / perspective

L'utilisation du modèle ainsi obtenu permet de détecter environ 20 % des dossiers frauduleux en rejetant seulement 3% des demandes (statistiques établies sur un échantillon hors-temps). Ce projet a été mis en production fin mai 2022.

Ce projet est innovant grâce à l'utilisation des bases de données graphes dans le domaine bancaire, combinée à un score machine learning.

Le modèle actuel se base sur le comptage des dossiers qui se connectent à la nouvelle demande de financement. La base de données n'est exploitée que partiellement et localement autour de la nouvelle demande.

Pour mieux tirer profit de la structure de la base de données, une première piste de recherche porte sur la détection de motifs. Il s'agit de trouver si tout ou partie du réseau formé autour de la nouvelle demande existe ailleurs dans la base de données. Selon la similarité avec le ou les réseaux trouvés, il sera possible de calculer un risque que la nouvelle demande soit frauduleuse. Plus formellement, il s'agit du problème de recherche de motif dans un graphe.

Une autre piste est d'envisager les approches récentes en intelligence artificielle telle que les réseaux de neurones de graphes conçus pour manipuler des données au format graphe. Ils peuvent alors aider assez naturellement à exploiter globalement et rapidement la base de données pour notamment calculer un indice de similarités entre des réseaux.

Enfin, BNPP PF aimerait étendre ce dispositif à d'autres géographies.