

Demi Module : Data Mining

Rapport de projet sous thème :

Recommandation de Produits sur un Site E-commerce

Réalisé par :

EL HANAFY Amine

Encadré par :

Professeur : Mohamed SABIRI

Année universitaire : 2023/2024

1. Introduction :

Le mini-projet vise à mettre en place un système de recommandation au sein d'une boutique en ligne, avec pour objectif principal d'anticiper les besoins des clients. L'implémentation de cet outil de recommandation personnalisée repose sur l'utilisation de techniques de data mining, notamment l'algorithme FP-Growth, pour analyser l'historique des transactions.

2. Contexte et Objectif :

La mise en place d'un système de recommandation dans le cadre du commerce électronique présente des avantages tant pour les clients que pour l'entreprise. Pour les clients, cela facilite la découverte de produits pertinents, améliorant ainsi leur expérience utilisateur. Pour l'entreprise, la satisfaction des clients conduit à une fidélisation accrue, une diminution du taux de désabonnement, et une augmentation des revenus grâce à la proposition de nouveaux produits via des stratégies de vente croisée.

3. Description des Données E-commerce :

Les données utilisées dans le projet proviennent d'un ensemble public disponible sur l'UCI Machine Learning Repository. Ces données comprennent des informations telles que le numéro de facture (InvoiceNo), le code du produit (StockCode), la description du produit (Description), la quantité achetée (Quantity), la date de la facture (InvoiceDate), le prix unitaire (UnitPrice), l'identifiant du client (CustomerID), et le pays de commande (Country).

Le prétraitement des données comprend la suppression des produits offerts en cadeau, ne conservant que les produits effectivement ajoutés au panier par les clients, et la consolidation des produits achetés par chaque client.

4. Modélisation avec FP-Growth :

L'algorithme FP-Growth est utilisé pour modéliser les règles d'association basées sur les transactions du site e-commerce. Cette technique de data mining permet d'identifier les règles fréquentes sans générer explicitement des ensembles de candidats, ce qui améliore l'efficacité du processus. Les règles extraites indiquent les produits associés, accompagnés de leurs probabilités d'achat.

5. Paramètres d'Exécution :

Deux hyperparamètres sont utilisés pour ajuster la modélisation FP-Growth :

- **minSupRatio** : Le support minimum pour qu'un ensemble d'items soit identifié comme fréquent.
- **minConf** : La confiance minimale pour générer une règle d'association à partir d'un ensemble d'items fréquent.

6. Anticipation des Besoins Clients :

Le système de recommandation développé propose un produit à chaque client, basé sur son panier actuel, maximisant ainsi la probabilité d'achat. Le résultat final est une table composée des identifiants clients, des produits recommandés, et des probabilités d'achat associées.

7. Visualisations et Analyses :

Le code intègre plusieurs visualisations pour enrichir l'analyse des résultats :

- **Boîte à moustaches des probabilités** : Une représentation graphique des probabilités extraites des règles d'association.
- **Statistiques descriptives des probabilités** : Une analyse statistique des probabilités associées aux règles d'association.

- **Histogrammes des moyennes de probabilités par facture** : Visualisation des tendances des probabilités moyennes par facture.
- **Graphique en boîte des moyennes de probabilités par facture** : Un résumé graphique des moyennes de probabilités par facture.
- **Nuage de points des prix estimés par probabilité** : Exploration de la relation entre les prix estimés et les probabilités d'achat.

8. Résultats et Analyse :

Le système a démontré une précision de 35% dans la prédiction du prochain achat parmi un catalogue de plus de 3000 articles. L'analyse des résultats met en avant la simplicité et l'explicabilité du modèle FP-Growth, offrant une compréhension claire des règles spécifiques au commerce.

9. Conclusion :

Le projet souligne l'efficacité des règles d'association, mises en œuvre via l'algorithme FP-Growth, pour améliorer l'expérience client, renforcer la fidélité, et stimuler les revenus. La transparence et la facilité d'implémentation du modèle en font une solution attractive, offrant une alternative compréhensible aux modèles d'intelligence artificielle plus complexes.

10. Perspectives Futures :

Les résultats encourageants du projet ouvrent la voie à des améliorations futures. L'exploration de paramètres supplémentaires, l'ajout de variables pertinentes, et l'application de techniques avancées pourraient encore optimiser le système de recommandation.

11. Recommandations :

Il est recommandé d'intégrer ce système de recommandation dans le processus opérationnel de la boutique en ligne, permettant ainsi de capturer les avantages en termes de fidélisation client et d'augmentation des ventes.

12. Conclusion Finale :

Le mini-projet démontre le potentiel significatif d'un système de recommandation basé sur les règles d'association dans le contexte du commerce électronique. L'équilibre entre précision, simplicité et explicabilité en fait une solution pertinente pour améliorer l'expérience client et stimuler la croissance de l'entreprise.