Prédiction du Churn des Clients Télécom

Analyse et modélisation prédictive

Contexte

- Les opérateurs télécoms font face à un défi : la perte de clients (churn)
- Conserver un client coûte 5x moins cher qu'en acquérir un nouveau
- Besoin : anticiper les départs pour mettre en place des stratégies de fidélisation

Objectif du projet

- Construire un modèle capable de prédire si un client va résilier son contrat
- Identifier les facteurs clés du churn
- Proposer des actions business concrètes pour réduire le churn

Données utilisées

- Source : Dataset bigml_59.csv (3 333 clients, 21 variables)
- Variable cible : churn (résiliation oui/non)
- Variables principales : plans, utilisation (minutes, appels, charges), service client

Méthodologie

- Préparation des données (nettoyage, encodage, standardisation)
- Séparation en train/test (70/30)
- Entraînement de 3 modèles : Régression
 Logistique, Arbre de Décision, Random Forest
- Comparaison des performances

Résultats

- Régression Logistique : modèle simple, bonne interprétation
- Arbre de Décision : capture les non-linéarités mais surapprentissage possible
- Random Forest : meilleure performance (Accuracy & AUC élevés)
- Facteurs clés : customer service calls, total day minutes, international plan

Visualisation

- Matrices de confusion
- Courbes ROC comparatives
- Importance des variables
- (Graphiques insérés depuis le notebook)

Recommandations Business

- Cibler les clients avec un plan international (plus enclins à résilier)
- ► Améliorer le service client → moins de plaintes, meilleure satisfaction
- Promotions personnalisées pour clients à usage élevé mais insatisfaits

Conclusion

- Le churn est prédictible avec les données disponibles
- Le modèle Random Forest est retenu
- L'entreprise peut mettre en place des actions préventives ciblées
- Prochaines étapes : tester d'autres modèles (XGBoost), déployer en production