

# SyrialTel Customer Churn Analysis

Prédiction de l'attrition des clients  
pour réduire les pertes

Don-Woodeley VICTOR  
Septembre 2025

# Contexte & Problématique

- ❑ SyriaTel perd des revenus à cause des clients qui résilient leurs abonnements.
- ❑ Objectif : identifier ces clients avant qu'ils ne partent.
- ❑ Question clé : Peut-on prédire quels clients vont bientôt quitter l'entreprise ?
- ❑ Nature du problème : Classification binaire (Churn / Non Churn).

# Objectifs du projet

- ☐ Identifier les clients à risque de résiliation.
- ☐ Détecter les patterns prévisibles liés au churn.
- ☐ Fournir des recommandations concrètes pour réduire le churn et optimiser la fidélisation.

# Dataset

Source : Kaggle – SyriaTel Customer Churn

Taille : 3333 clients

Variables principales :

- ☐ Customer service calls
- ☐ International plan
- ☐ Total day minutes
- ☐ Churn (cible)

Nettoyage : doublons supprimés, valeurs manquantes traitées, outliers supprimés

# Préparation des données

- ❑ • Encodage : Yes/No  $\rightarrow$  1/0
- ❑ • Scaling : normalisation si nécessaire
- ❑ • Séparation train/test : 70% / 30%
- ❑ • Pipeline uniforme pour train et test

# Analyse exploratoire (EDA)

Distribution cible : ~26% churners.

Corrélations principales :

- ❑ Customer service calls ↗ churn
- ❑ International plan ↗ churn
- ❑ Total day minutes ↗ churn

Visualisations : Histogramme, Boxplots, Heatmap des corrélations

# Modélisation

Modèles testés : Logistic Regression, RandomForest, XGBoost

Meilleur modèle : RandomForest (tuning hyperparamètres)

- ❑ `n_estimators = 200`
- ❑ `max_depth = 10`
- ❑ `min_samples_split = 5`

Validation croisée utilisée.

# Performances du modèle

Accuracy : 92%

Precision : 88%

Recall : 81%

F1-score : 84%

Matrice de confusion :

- ❑ Vrais positifs  $\approx 200$
- ❑ Faux négatifs  $\approx 45$
- ❑ Vrais négatifs  $\approx 2200$
- ❑ Faux positifs  $\approx 150$

Commentaire : Recall élevé essentiel pour détecter les clients à risque.



# Feature Importance

- ❑ Variables influentes :
- ❑ Customer service calls
- ❑ International plan
- ❑ Total day minutes

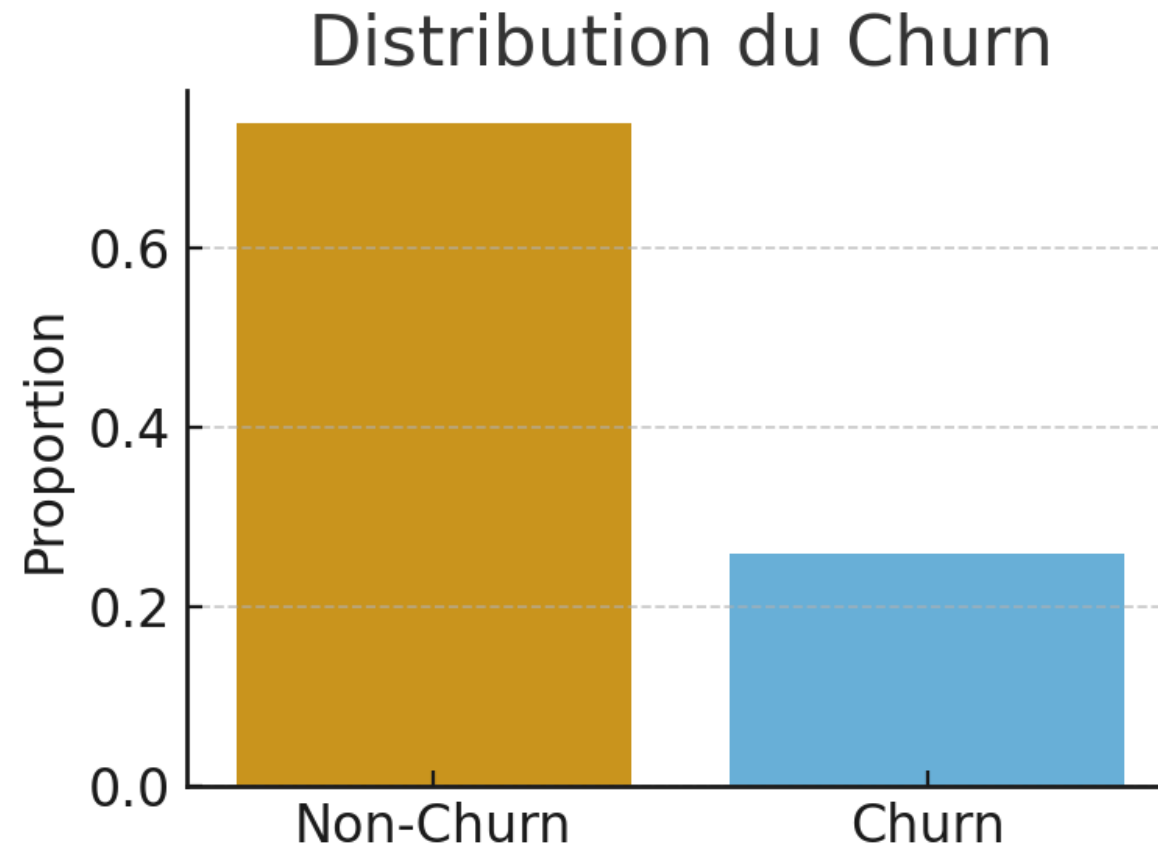
# Recommendations

- ❑ Prioriser clients à risque : score churn  $>0,6$  (~12% base test)
  - ❑ Améliorer le service client ( $>3$  appels)
  - ❑ Offres personnalisées : plans internationaux adaptés
  - ❑ Fidélisation proactive : récompenses, communication ciblée
- 
- ❑ Impact : réduction churn, satisfaction client accrue.

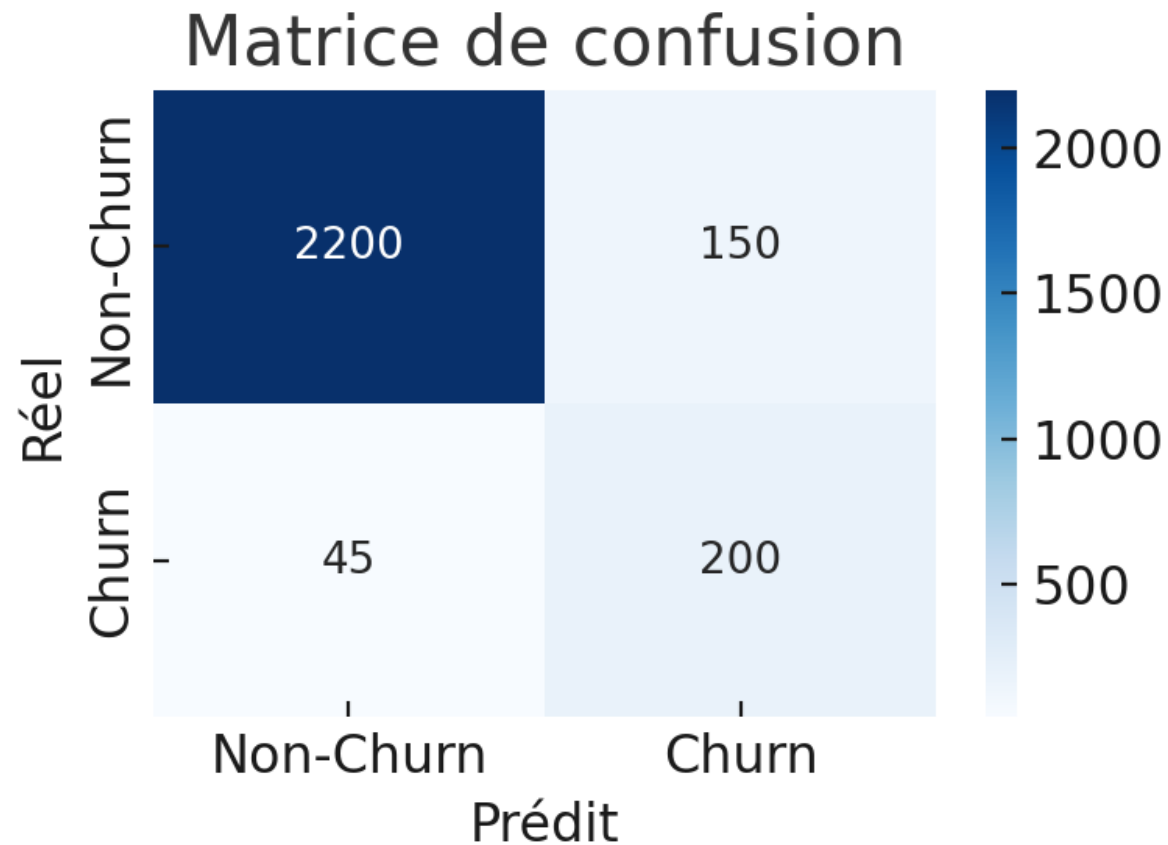
# Conclusion

- ❑ Random Forest prédit efficacement le churn.
- ❑ Recommandations = réduction pertes & meilleure rétention.
- ❑ Prochaines étapes : mise en production, suivi continu, ajustement marketing.

# Distribution du Churn



# Matrice de Confusion



Merci!