SyriaTel Co Ltd.

Churning Model for Telcom Company

Contributors: Brian-Waweru
GitHub: rurigi-waweru
email: rurigiwaweru@gmail.com

Overview

We aims to predict customer churn for *SyriaTel* using a sample of their historical customer data. By building a binary classification model we were able to identify patterns and factors that influence whether a customer will leave the company. The model will assist the company in targeting at-risk customers with *retention strategies*, thereby reducing customer attrition and preserving revenue.

Business Understanding

Customer churn is a critical business challenge for telco companies.

In a _highly competitive_ and _saturated market_, **_retaining existing customers_** is often more cost-effective than acquiring new ones. Churn not only impacts immediate revenue but also affects long-term customer lifetime value, **brand loyalty**, and **operational efficiency**.

Understanding why customers leave — and more importantly, identifying who is likely to leave — can empower SyriaTel to take timely, targeted actions. These may include personalized marketing campaigns, service improvements, or tailored retention offers.

Objectives
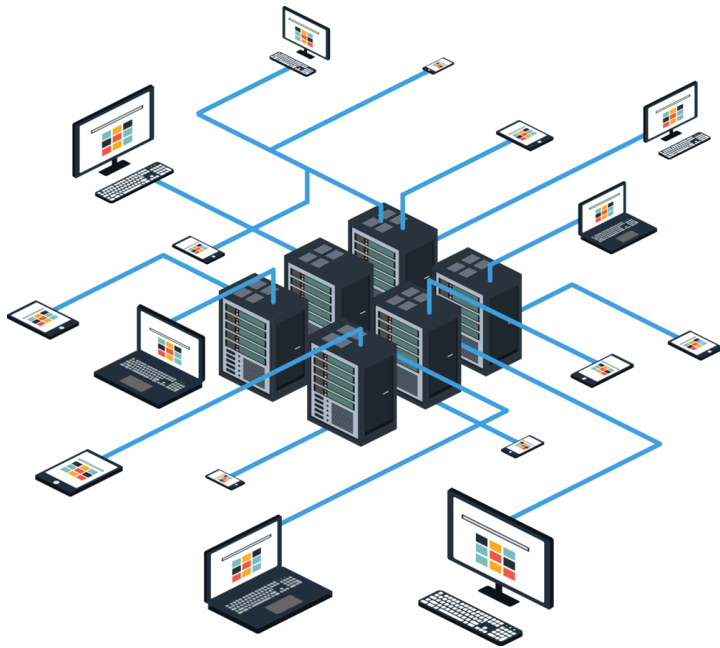
- Primarily we aimed at building a predictive model that accurately flags at-risk customers enabling **SyriaTel** to shift from reactive to proactive customer retention.

- Here are the Specific Objectives:-

  1. **Build a Predictive Model for Churn**

  2. **Improve Churn Prediction Accuracy**

  3. **Develop a Repeatable ML Pipeline**: Build a clean and modular workflow that can be reused with updated customer data in the future.

  4. **Then Communicate Findings Clearly**: Present model insights

Models

- We leveraged, trained and evaluated several models including **Logistic Regression**, **Random Forest**, and **KNN**. We also addressed class imbalance via *SMOTE* and *class weighting*.

- We used classification metrics (***accuracy***, ***precision***, ***recall***, ***F1-score***) and **ROC-AUC** to assess performance. Then focused on identifying the model's ability to detect churned customers accurately.

- We ran improvements to capture more accuracy and robustness of predictions.

Interpretation

ROC Curve Comparison
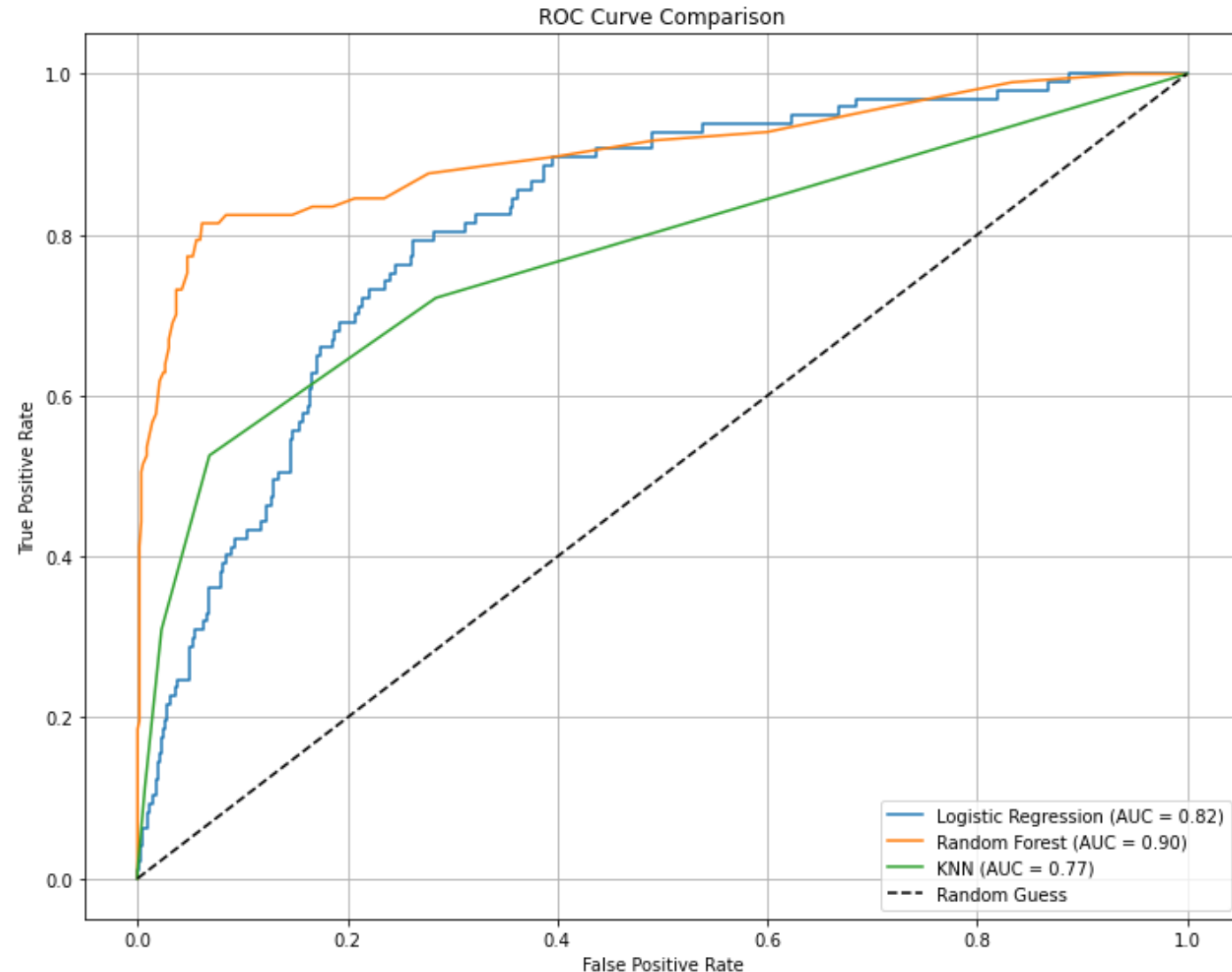
- The Random Forest model ran well in capturing potential churn costumers.
- This model will serve the company well so that it can leverage better customer retention strategies.

Recommendation

The ROC curve showing how your classifier performs across different thresholds. The AUC value, 0.90 summarizing overall performance for the **Random Forest** points the best performing model. This means that it is the closest to perfect classification. From this we infer and conclude that the Random Forest Model is the most accurate and reliable classifier among the three.

Thank you

The End!