

CHURN PREDICTION

A MACHINE LEARNING PROJECT

~By Victoria Nabea

Project Overview:

- SyriaTel Communications is a Telecommunications company that is looking to predict and prevent customer churn. Customer churn is when a customer leaves/discontinues their service with SyriaTel. Customer churn is a major problem for many service-based companies because it is so expensive. Not only does the company lose the customer's monthly/yearly payment, but they also incur a customer acquisition cost to replace that customer. The project aims to analyze a telecommunications dataset to predict customer churn. The dataset contains information about customers' demographics, calling behavior, and service plans. The goal is to develop a predictive model that can accurately identify customers who are likely to churn, enabling the company to take proactive measures to retain them

Business Understanding:

- The telecommunications company is facing a significant issue of customer churn, where customers are discontinuing their services. This poses a threat to the company's revenue and market share. To address this problem, the company wants to leverage the available customer data and build a predictive model that can identify customers at a high risk of churn. By identifying these customers in advance, the company can implement targeted retention strategies to improve customer satisfaction and reduce churn rates.

Modelling:

- We use different modeling algorithms to compare and choose the one that is most accurate to predict our data.
- Models involved:
 - **Logistic Regression**
 - **Decision Trees**
 - **Random Forest**

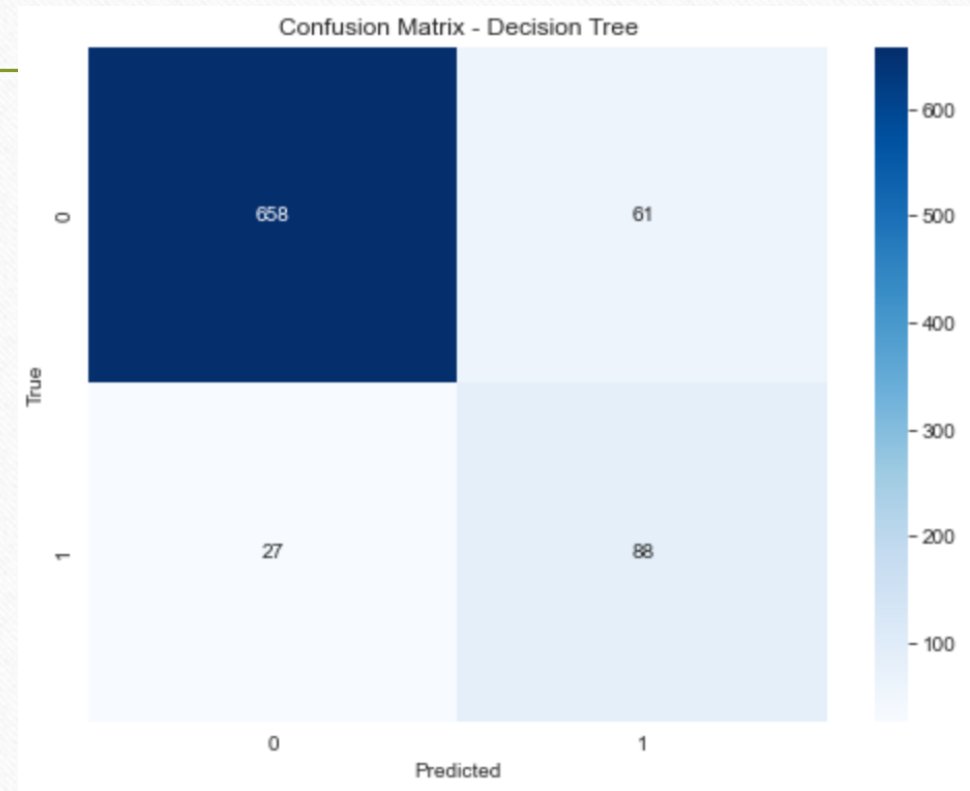
Modelling: Logistic Regression

- It is most commonly used when the data in question has binary output, so when it belongs to one class or another, or is either a 0 or 1.
- This model gave an accuracy metric of 75.2% which is not so bad, but we can have better



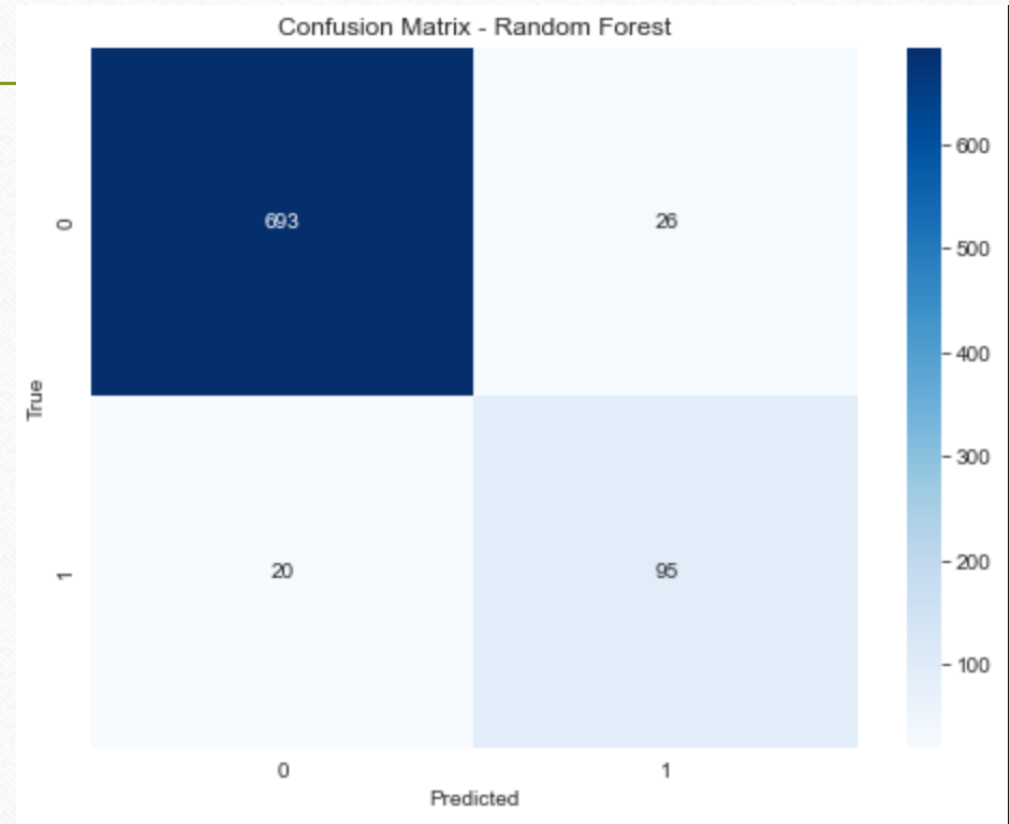
Modelling: Decision Trees

- The logic behind a decision tree can be readily understood due to its tree-like structure, which visually represents the sequence of decisions and their outcomes.
- This model gave an accuracy metric of 89% which is improved, but can we have better? hmm



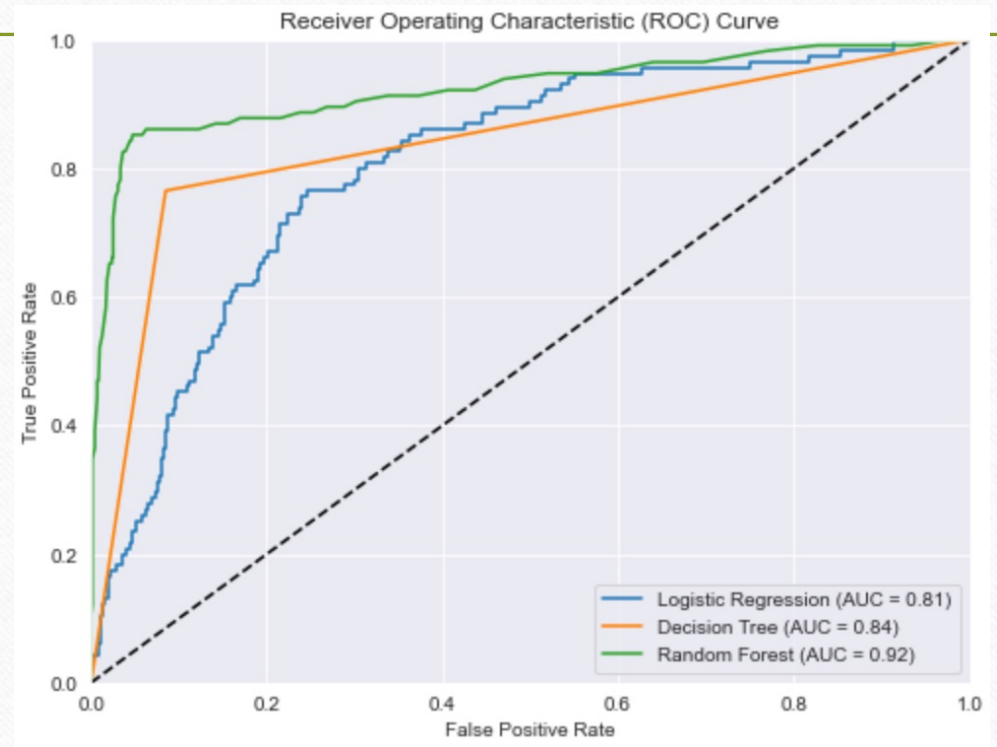
Modelling: Random Forest

- In random forests, decision trees are built using randomly selected data samples. Each tree provides a prediction, and the final solution is determined through a voting process.
- This model happened to give us the best accuracy results of 94%. wow



Evaluation

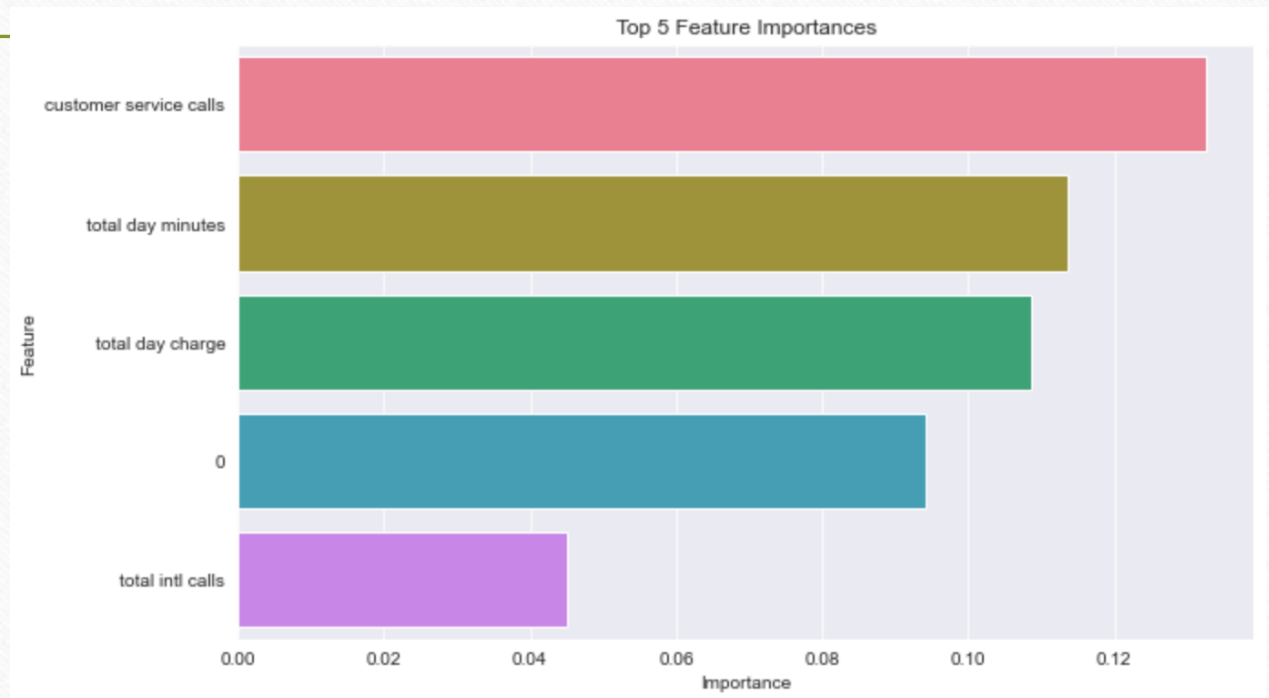
- We evaluate our models using the ROC curve and comparing the accuracy scores. Seems like we have a winner already from both comparisons. THE RANDOM FOREST MODEL!



Findings:

Our analysis showed that we can predict churn with an accuracy of 95.55 % . Quite impressive

The diagram by the side whowes the most important features affecting churn that we should be keen about.



Recommendations

- The telecommunication company should prioritize efforts to enhance customer satisfaction through improved customer service experiences, enhancements in service offerings, and the implementation of strategies to reduce day-time charges.
- it is crucial for the telco to monitor and analyze the number of customer service calls, as a high volume of such calls can be indicative of customer dissatisfaction.
- By focusing on these insights and taking proactive measures, the telecommunication company can strategically address the identified pain points, thereby reducing customer churn, improving customer retention, and fostering higher levels of customer satisfaction.

Next steps:

- 1. Occasional re-evaluation and checking of the models performance
- 2. Gather more data in states with high churn rate
- 3. Considering feedback and insights from domain experts, stakeholders, or end-users of the model
- 4. Gather more data on cell signals both locally and international calls
- 5. Audit customer service department

THANK YOU

- ~Victoria Nabea
 - Data scientist
 - Email: nabeamvictoria@gmail.com
- LinkedIn: <https://www.linkedin.com/in/victoria-nabea-10b0b3202>
 - Github: [git@github.com:VikkieN](https://github.com/VikkieN)