MACHINE LEARNING;

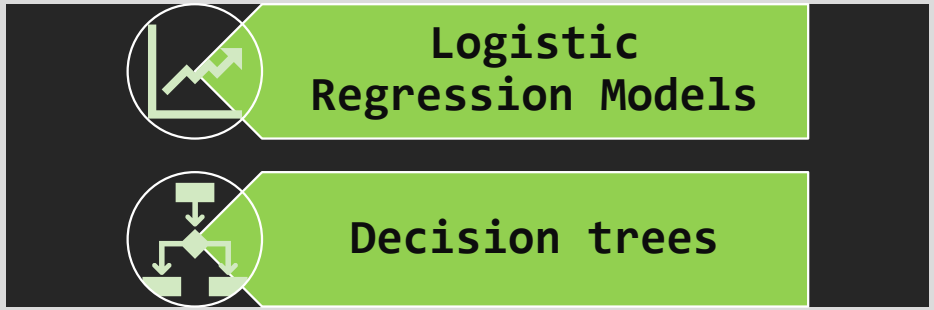
SYRIATEL CUSTOMER CHURN PREDICTION

INTRODUCTION

In the telecommunications industry, a lot of factors contribute to customer churn. Some of the factors include;

- quality of service provision
- customer service effectiveness
- competition from service providers

Employing predictive models such as;



can help service providers anticipate churn and implement measures to counteract client loss easily.

BUSINESS PROBLEM

What is Customer Churn;

- Customer churn refers to the phenomenon where a customer stops using a company's services

Impact of Customer Churn;

- Churn poses a significant threat to the growth and profitability of telecommunications firms.

Goal

- Be able to predict customer churn rates for remediation purposes

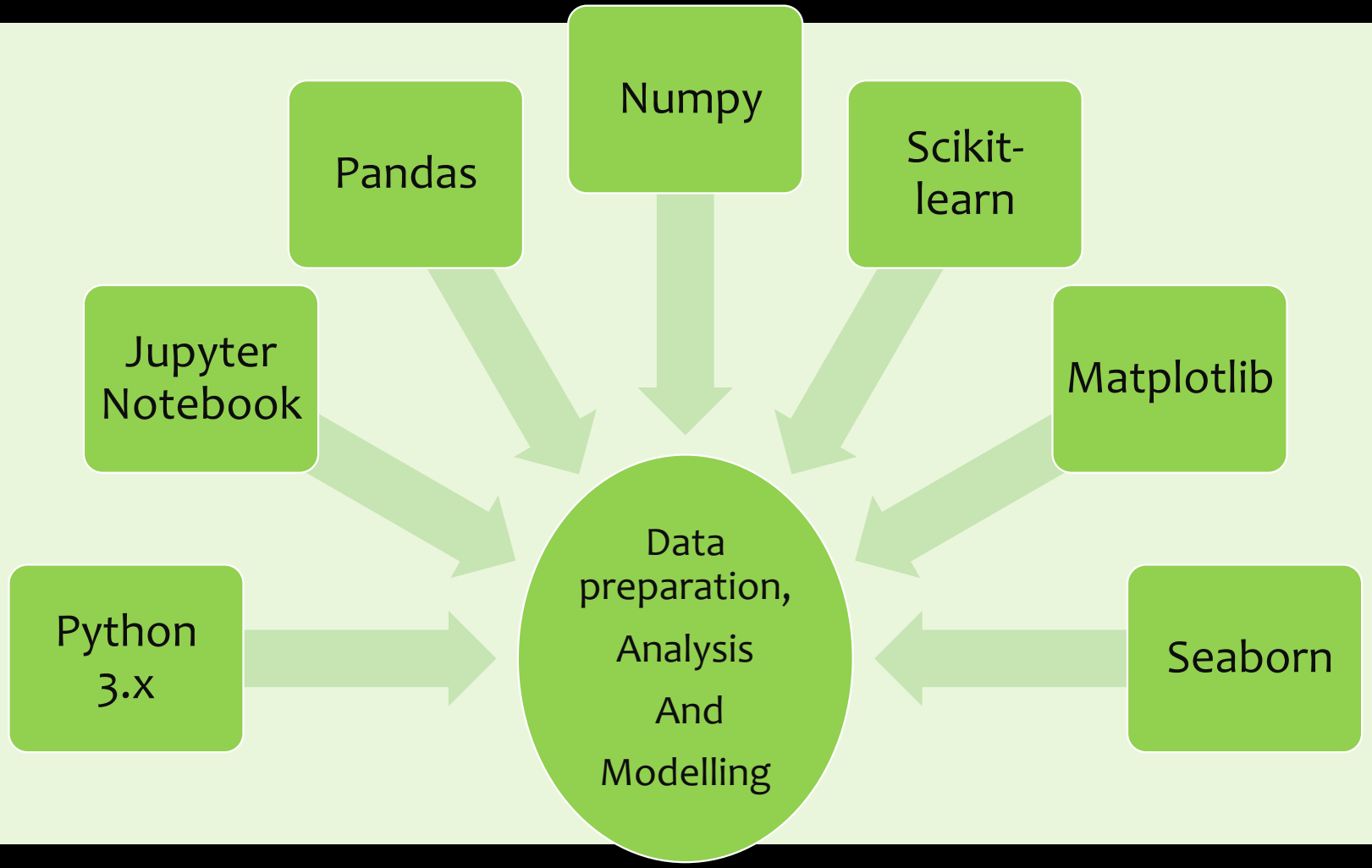
Methodology

- Using a data-driven approach to predict the likelihood of customer churn based on various customer characteristics and behaviors.

OBJECTIVES

- Understand the customer churn rate.
- Estimate the churn effects on the company's revenue.
- Develop a predictive model that can accurately predict the likelihood of customers churn based on their past behavior and characteristics.

PROJECT DEPENDENCIES



PROCESS

Data Preparation

- Loading Data
- Checking For missing values

Exploratory Data Analysis

- Exploring Churn Rates
- Estimating Customer Churn Impact on Revenue

Preprocessing

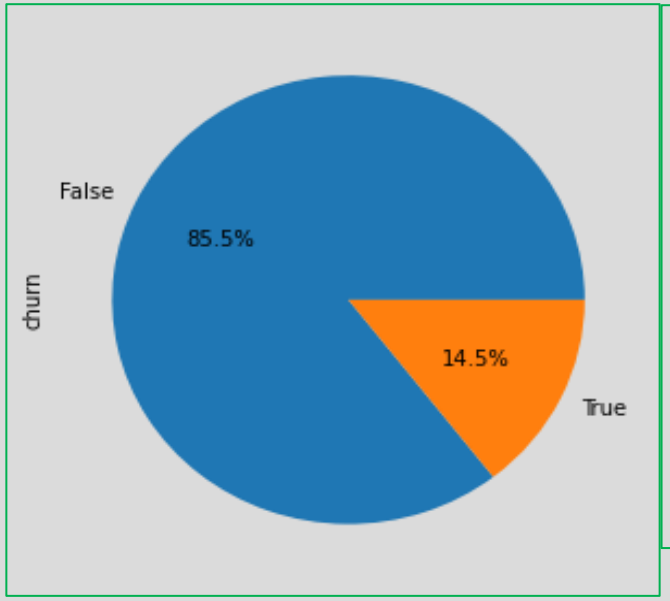
- Performing test train split; to prevent data leakage in the modelling process
- Dealing with Categorical Variables ; One hot encoding, scaling
- Creating processed dataset for modelling through merging

Modelling; Iterative approach

- Fitting the baseline logistic regression model
- Fixing class imbalance through SMOTE
- Training the second model on the resampled data; the model overfits
- Regularizing the model; Fitting Ridge and Lasso Model; The evaluation Metrics are still substandard hence need to explore another modelling technique
- Fitting A decision Tree.

CUSTOMER CHURN ANALYSIS

CUSTOMER CHURN RATES



Churn status;

False: Still a customer (85.5%)

True : Not a customer (14.5%)

EFFECT ON CUSTOMER REVENUE

- In every 100 clients, 14 leave the company
- Total revenue from charge(198,145.93)
- Total contributed by the lost clients (31566.93 units) 15% of the Revenue
- The 14% of the churn amounts to 15% loss of the firm's revenue obtained through total charge

CUSTOMER CHURN PREDICTION MODEL

- Models fitted
- Several predictive models are built and evaluated, including logistic regression and decision trees, to identify the most effective model for predicting customer churn.
- 1. The baseline Model; class imbalance in the data is detected in the data. The model has poor F1 , precision, and recall scores. It is necessary to conduct SMOTE (Synthetic Minority Oversampling Technique)
- 2. The second Model is fitted on the resampled data. The Model overfits as it has better F1, Precision, and recall score on train data but performs poorly on test data.
- 3. To deal with the overfitting problem, Model Regularization is used hence fitting Ridge and Lasso Models. The results do not differ from the second model.
- 4. Decision Tree is finally fitted on resampled data. The accuracy of training data is 71% and 91% of testing data. This stands to be the best perming model hence recommended for customer churn prediction.

RESULTS

Objective 1:

Understand the customer churn rate

- The SyriaTel customer churn rate is approximately 14%, which means that out of every 100 customers, 14 are expected to leave the company.

Objective 2 :

Estimate the Customer churn effects on the company's revenue.

- The 14% churn rate approximately results to approximately 15% of revenue loss.
- This is a significant number and there is a need for immediate attention to improve customer retention.

Objective 3:

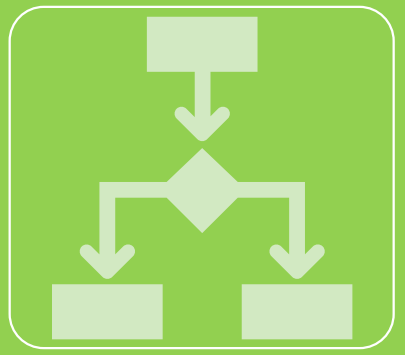
Develop a predictive model that can accurately predict the likelihood of customer churn based on their past behaviour and characteristics

- The Decision Tree model stands out to be one of the most accurate model to predict customer churn.
- The model has an accuracy of 71% on training data and 91% on the test data showing its ability to generalize well on new data.

RECOMENDATIONS

14% client loss with an estimated loss of the company's 15% revenue could be detrimental to the company on both profit margins and sustainability.

- There is need to Identify the most influential factors contributing to customer churn from the correlation matrix. There is need for collecting data that is more client-centred so as to explore characteristics of clients who are likely to be lost for remediation purposes.



Based on the results, we recommend the use of a Decision Tree model for predicting customer churn

- This model has a high accuracy score, making it a reliable choice for this specific problem. The model can be fine-tuned and regularized to improve its performance


Thank
You!