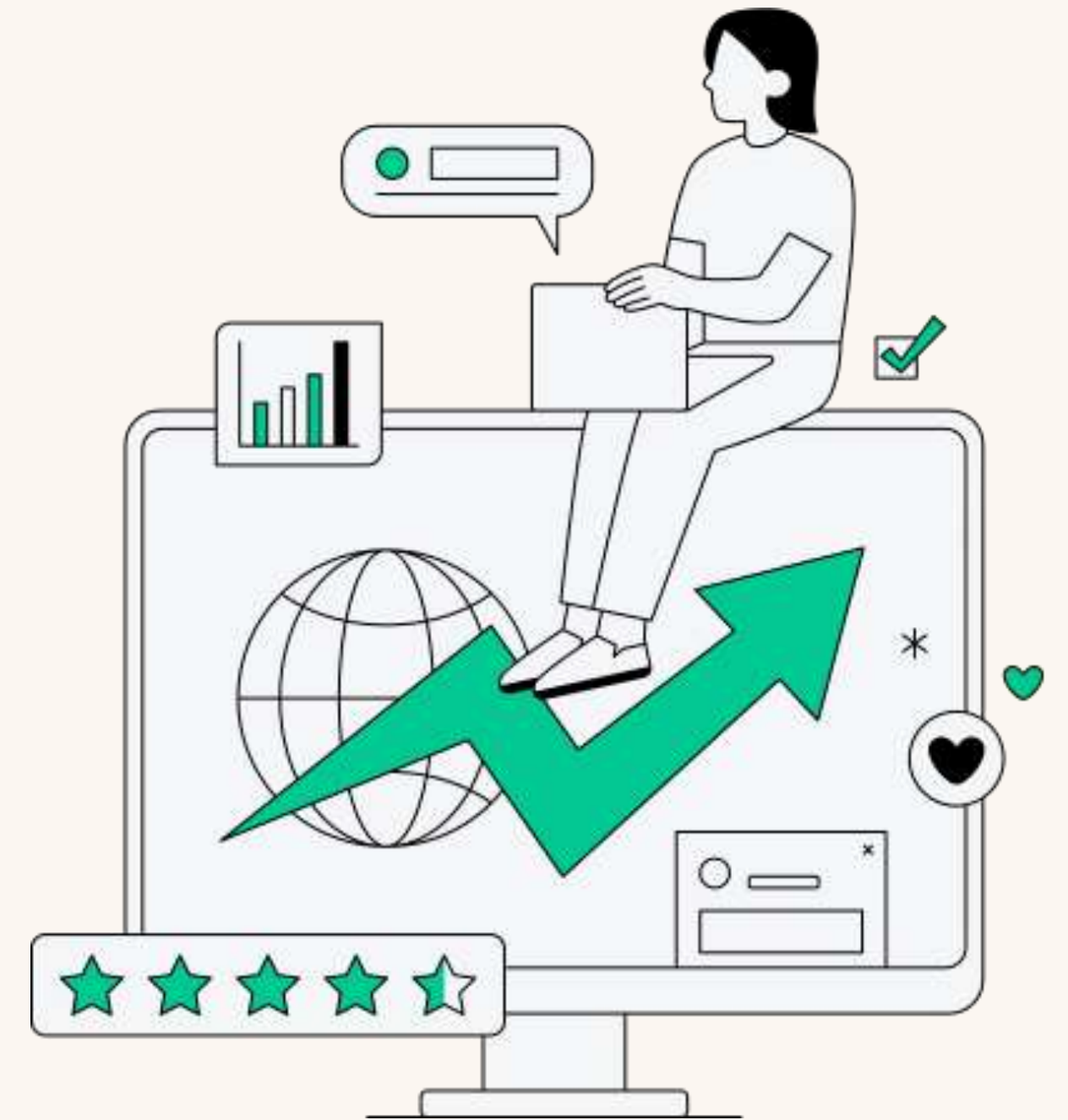Presented by Salahudin Mahamed

SyriaTel Customer Churnig Analysis

1. Business Understanding

SyriaTel company is a telecommunication company focused on being one the best in providing quality services to its customers and therefore they need to be competitive and be aware of there customers dynamics and preferences.

Although such efforts have been fruitful over the years, the company needs to increase its commitment by reducing customer churning rates, which might threaten its market position, profitability, and overall growth. Retaining customers will increase the company's net profits by reducing costs involved in losing customers.

Stakeholders of the Project

1. SyriaTel company
2. Management and Shareholders
3. Employees
4. Customers

Comparison

Research Questions

- What is the main reason Customers leave?
- What are the retention Strategies that are currently in place
- What is the financial impact on financial position

Research Objective

- Improve customer searvice
- Identify Key drivers to churn
- Reduce cost associated with customer churning by being able to predict in advance

2. Data Understanding
Features

- State: The state the customer lives in
- Account Length: The number of days the customer has had an account.
- Area Code: The area code of the customer
- Phone Number: The phone number of the customer
- International Plan: True if the customer has the international plan, otherwise false.
- Voice Mail Plan: True if the customer has the voice mail plan, otherwise false.
- Number Vmail Messages: the number of voicemails the customer has sent.
- Total Day Minutes: total number of minutes the customer has been in calls during the day.
- Total Day Calls: total number of calls the user has done during the day.
- Total Day Charge: total amount of money the customer was charged by the Telecom company for calls during the day.
- Total Eve Minutes: total number of minutes the customer has been in calls during the evening.
- Total Eve Calls: total number of calls the customer has done during the evening.
- Total Eve Charge: total amount of money the customer was charged by the Telecom company for calls during the evening.
- Total Night Minutes: total number of minutes the customer has been in calls during the night.
- Total Night Calls: total number of calls the customer has done during the night.
- Total Night Charge: total amount of money the customer was charged by the Telecom company for calls during the night.
- Total Intl Minutes: total number of minutes the user has been in international calls.
- Total Intl Calls: total number of international calls the customer has done.
- Total Intl Charge: total amount of money the customer was charged by the Telecom company for international calls.
- Customer Service Calls: number of calls the customer has made to customer service.
- Churn: true if the customer terminated their contract, otherwise false

3. Data Preparation

01.
checking the shape for our dataset before we start cleaning it

02.
checking for missing values and for our case there were no missing values

03.
checking for duplicates and for our case there were no duplicates to be treated

Explaratory Data Analysis

After treating our dataset, we removed majority of the outliers and performed binary and multivariable analysis to see what our data communicates before building the prediction models

This graph shows the churning rate at every state with Nj having the highest rate and Hi having the lowest

Churn Rate by State
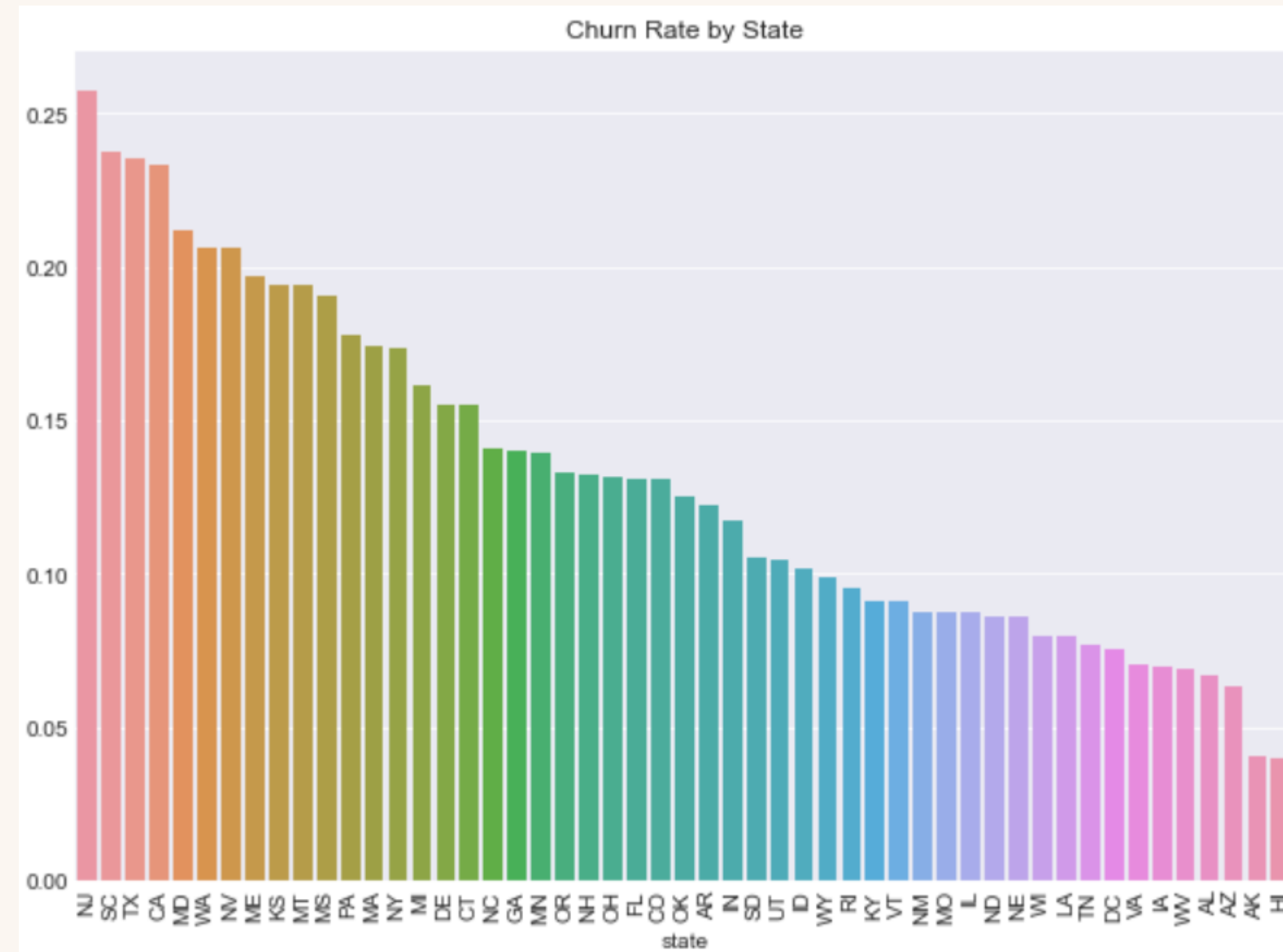
Multivariate Analysis

- **We performed a multivariate analysis by checking at the correlation between our Features of the dataset**
- **This shows how every feature affect the change in the other with 1 and -1 showing the most change**
- **if the correlation is 0 then it means a change in one variables is not affected bt the other**

4. Modeling
(a) Logistic Regression

For the first model we a built a base logistic Regression that had a test accuracy score of of 88% but in terms of predicting the minority(Churned class) it was not one of the best

We had to improve by addressing the class imbalance which did well in predicting the minority class but test accuracy dropped to 76%
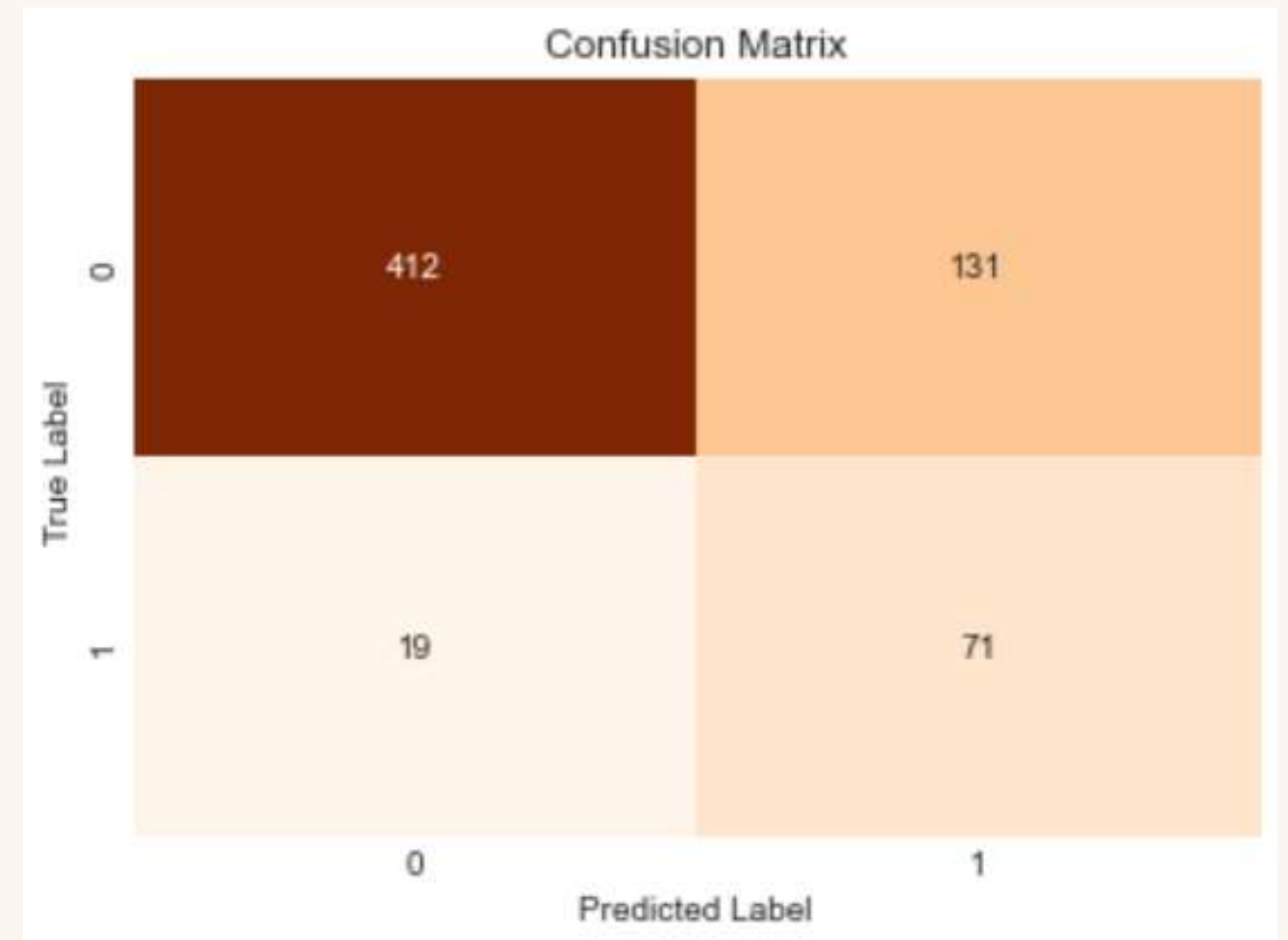
(b) K-Nearest Neighbor

For the second model we a built a base K-Nearest Neighbor that had a test accuracy score of of 87% which is quite nice

We tried to improve by introducing Grid search and hyperparameter tuning to get the best parameters for our model. We tested a new KNN model with the best parameters and the new accuracy rose to 88% which was amazing

Confusion Matrix
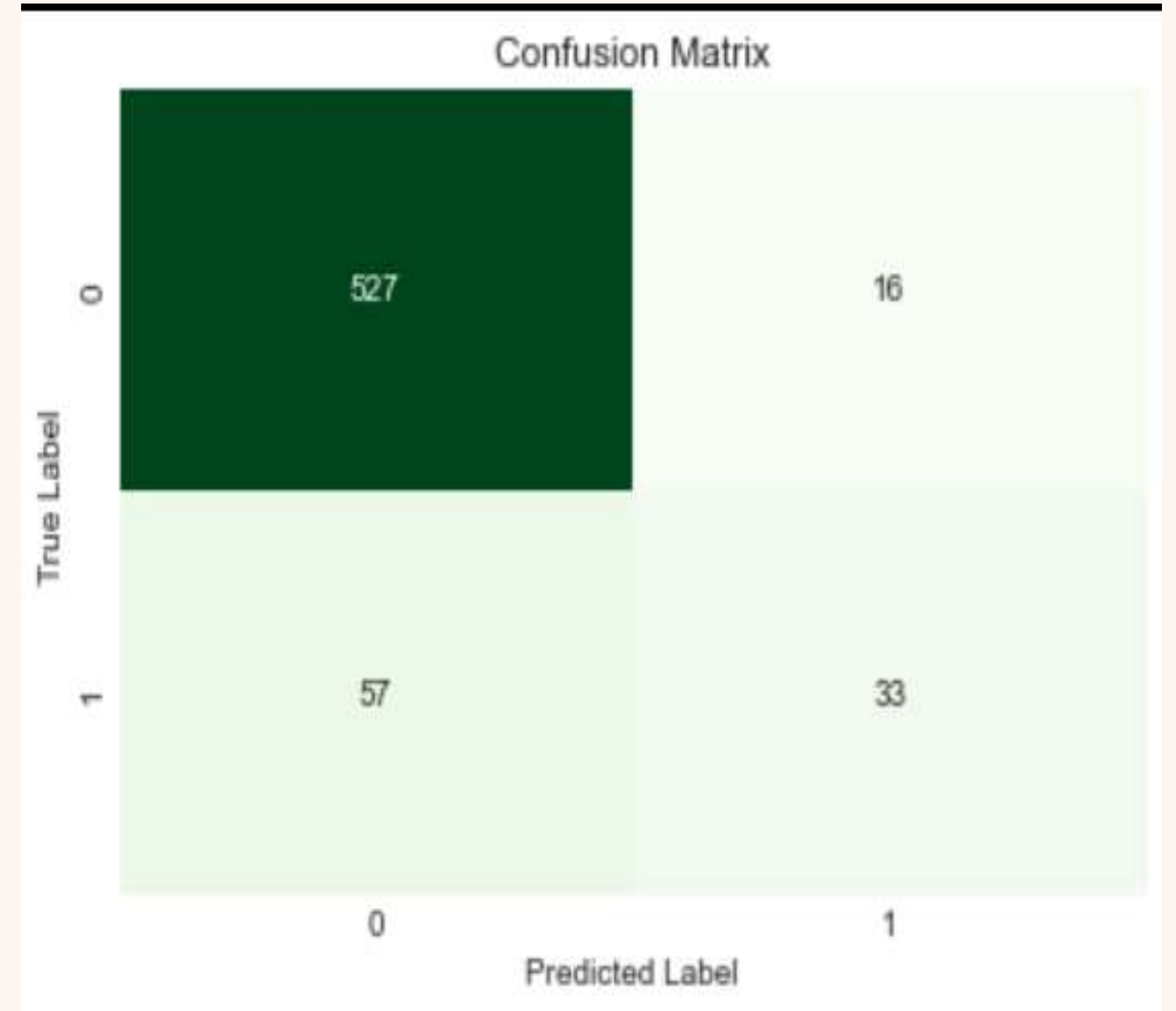
(b)Decision Tree Classifier

For the Third model we a built a base Decision Tree classifier that had a test accuracy score of of 91% which is the best compared to the previous 2 model

We tried to improve by introducing Grid search and hyperparameter tuning to get the best parameters for our model. We tested a new Decision Tree classifier model with the best parameters and the new accuracy rose to 95% which was amazing
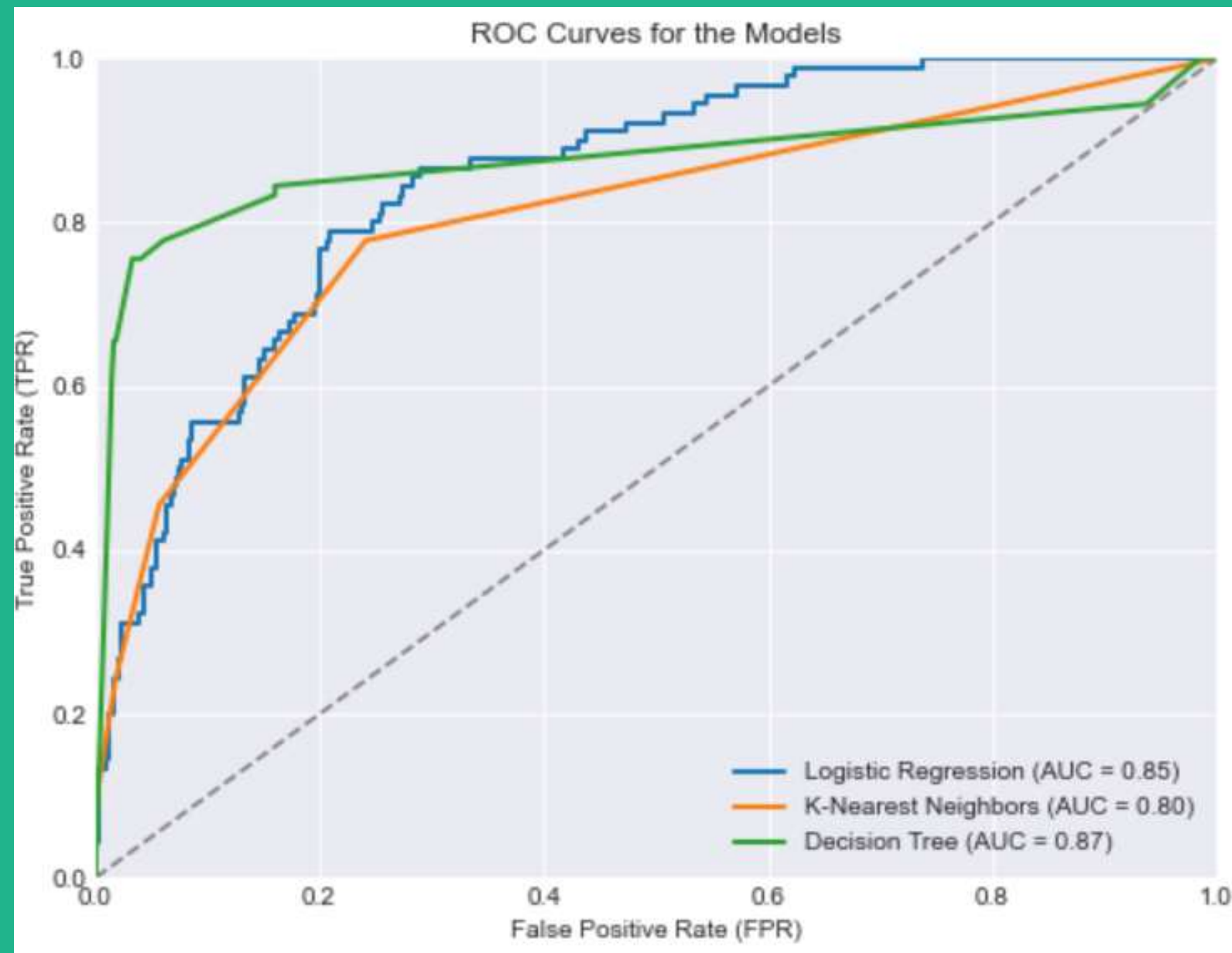

Confusion Matrix

95%

Comparing the Three Models

Now that we have done base module and improved module for each of our 3 modules, we can compare them using the accuracy score and see which module fits our data the most when it comes to prediction. From the test score, Module 3(Decision Tree Classifier) is the best in predicting whether a customer will churn or not

Comparison

ROC Curves for the Models

Logistic Regression (AUC = 0.85)
K-Nearest Neighbors (AUC = 0.80)
Decision Tree (AUC = 0.87)

ROC curves

Again after plotting the roc-curve, Decision Tree performed better than Logistic regression and K-nearest Neighbor since it has a largest area under the curve.

This was to prove that Decision Tree is the best module out of the 3 modules we built in predicting the whether a customer is going to churn or not

5. Recommendation and conclusion

01.

I would recommend SyriaTel communication to use Decision tree Model in predicting whether a customer will churn or not therefore reducing the cost of trying to retain customers

02.

Put new strategies in place such as marketing and advertising in states like NJ, SC and TX to reduce churning rate and increase market reach

03.

Maintain the same strategies of retaining customers in states like HI, AK and AZ since churning rate is very low

Presented by Salahudin Salat

Thank you very much!