PREDICTING PATTERNS IN CUSTOMER CHURN AT SYRIATEL MOBILE TELECOM

The use of Machine learning to predict churn

OVERVIEW¶

- Customer retention is important in sustaining business growth and Success in an ever evolving and competitive business environment.

- In alignment with industry insights, which emphasize the high cost of acquiring new customers compared to retaining existing ones, Syriatel is committed to leveraging machine learning to predict patterns in customer churn. By proactively identifying customers at risk of leaving, Syriatel aims to implement targeted retention strategies that enhance customer satisfaction and loyalty.

PURPOSE OF THE RESEARCH

- **Research Objectives:**

1. Identify the pivotal factors influencing customer churn for Syriatel Mobile Telecom.
2. Determine the most optimal machine learning model for predicting customer churn.
3. Develop a comprehensive customer retention strategy to mitigate churn rates.

- **Research Questions:**

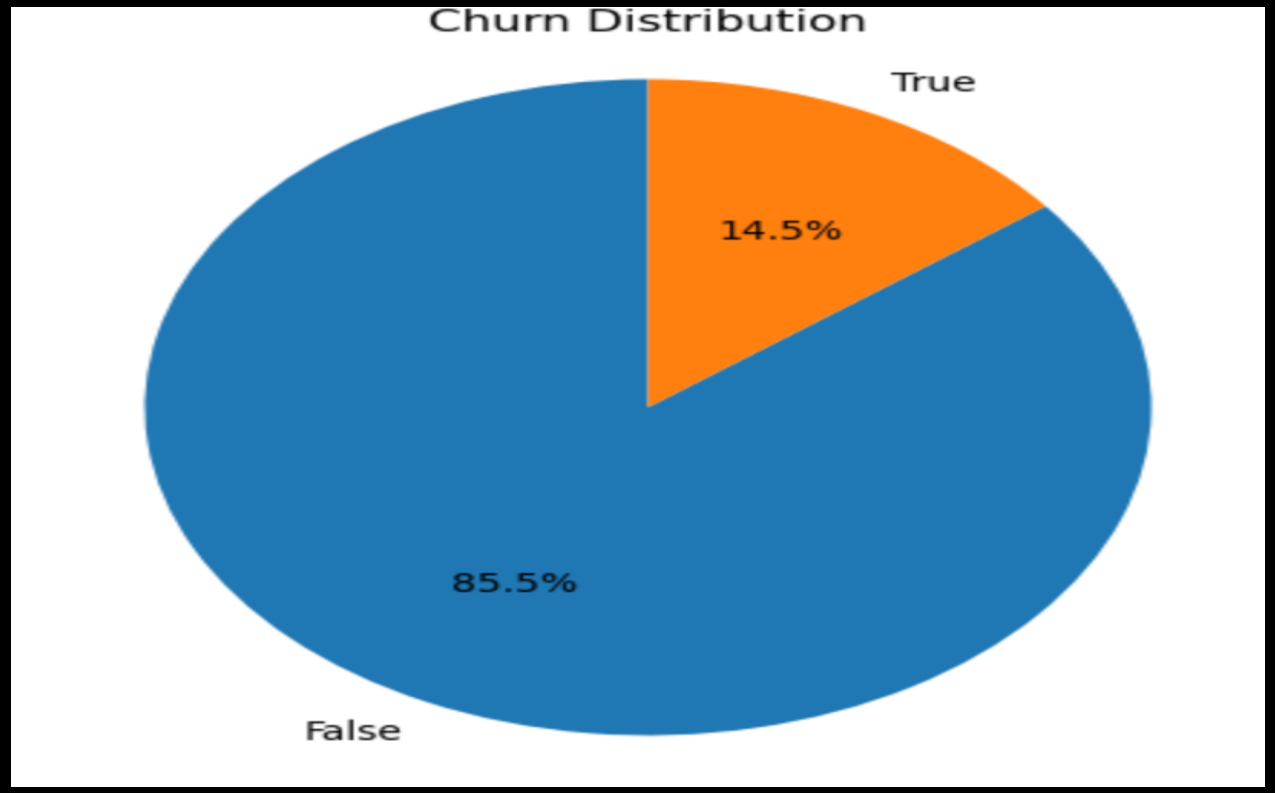
1. What are the primary predictors of customer churn for Syriatel Mobile Telecom?
2. Which machine learning model exhibits the highest efficacy in predicting customer churn?
3. What strategic initiatives can Syriatel Mobile Telecom undertake to foster customer retention and diminish churn rates?

DATA UNDERSTANDING

- The "Churn in Telecom" dataset from Kaggle and it aims to provide insights into customer behavior and predict customer churn for a Syria telecommunications firm.
- The dataset contains 3333 entries and 21 columns, including various attributes related to customer activity and subscription cancellation status.
-

EXPLORATORY DATA ANALYSIS (EDA)¶

- **Univariate Analysis**



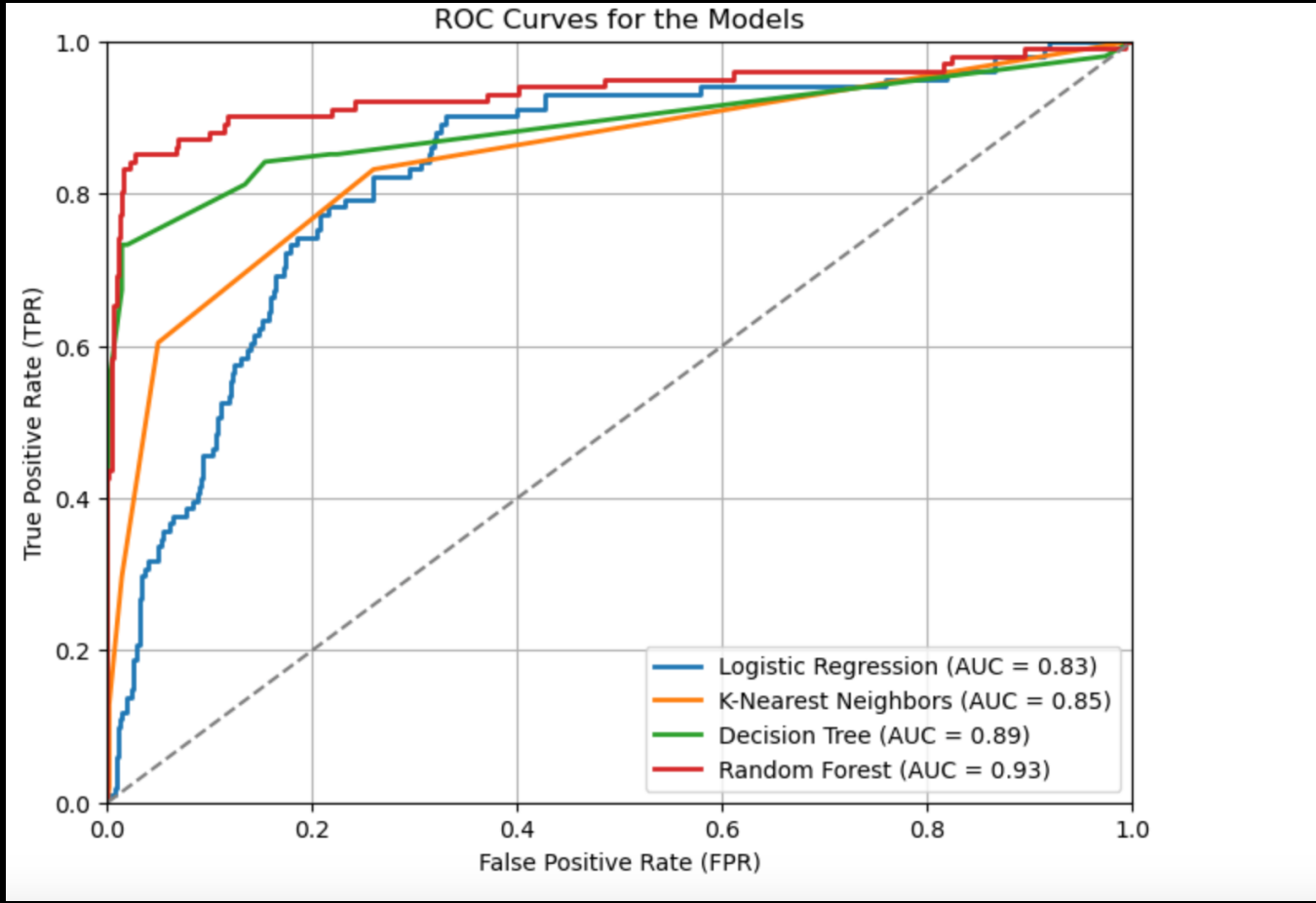
MODELING

- The intention in this case is to find the model and parameters that perform the best. I trained and evaluate the following models:
- Logistic Regression Model
- Decision Trees,
- Random Forests.
- K – Nearest Neighbour

ANALYSIS OF EACH MODEL PERFORMANCE

Model	Accuracy (Test Set)	Precision (Test Set)	Recall (Test Set)	F1 Score (Test Set)	
0	Logistic Regression	0.780000	0.390000	0.780000	0.520000
1	Decision Trees Classifier	0.940000	0.940000	0.680000	0.790000
2	Random Forest Classifier	0.950000	0.930000	0.680000	0.780000
3	K-Nearest Neighbors Classifier	0.880000	0.770000	0.270000	0.410000

ROC CURVE AND AUC -SCORES



- The ROC curves for Logistic regression, K-Nearest neighbour, Decision tree and Random Forest were analyzed.
- The Random Forest Classifier out-performed the others showing a higher Area Under the Curve (AUC)and better classification performance, making it the most effective model for the given task.

CONCLUSION

- The ROC curves for Logistic Regression , K-Nearest Neighbors , Decision Tree , and Random Forest models were analyzed. The Random Forest model outperformed the others, showing a higher Area Under the Curve (AUC) and better classification performance, making it the most effective model for the given task.

RECOMMENDATIONS

1. Syriatel to make use of the Random Forest Classifier as the primary model for predicting customer churn.
2. Create better pay packages or incentives to customers who use the services during the day.
3. Customer behavior during different times of the day and international calling patterns strongly influence churn.
4. Syriatel comes up with strategies to reduce on Customer Service calls, as this is among the top features that would likely lead to Customer Churn. Example: come up with an AI feature that responds to customers concerns in a timely and efficient manner.
5. Address Customer Service Issues:’ customer_service_calls’ has high importance, indicating that the number of customer service calls is a strong predictor of churn.

NEXT STEPS

- Further Analysis of patterns in customer day calls and those on international plan to identify the main causes of churn for these particular set of customers.

THANK YOU

QUESTIONS AND FEEDBACK SESSION.