SyriaTel Churn Analysis

Introduction

SyriaTel is a telecommunications company in Syria. They have been informed that some of their customers have started to discontinue their service. This analysis will determine what features will indicate if a customer will discontinue their service.

PROBLEM STATEMENT

Syria Tel is encountering a significant issue as some customers are discontinuing their services, leading to instability in the company's performance.

To tackle this, they aim to forecast when customers might leave and comprehend the reasons behind their departure, striving to prevent churn and maintain overall customer satisfaction.

OBJECTIVES

1. Analyze Syria Tel's customer churn data using classification techniques to understand and quantify the factors influencing customer attrition.
2. Data Exploration: Understand relationships between variables and churn.
3. Build Classification Model: Develop models (e.g., Logistic Regression, Random Forest) for churn prediction.
4. Feature Analysis: Determine critical features impacting churn predictions.Model Evaluation: Assess model performance using accuracy, precision, and recall metrics.

DATA UNDERSTANDING

Data Source:

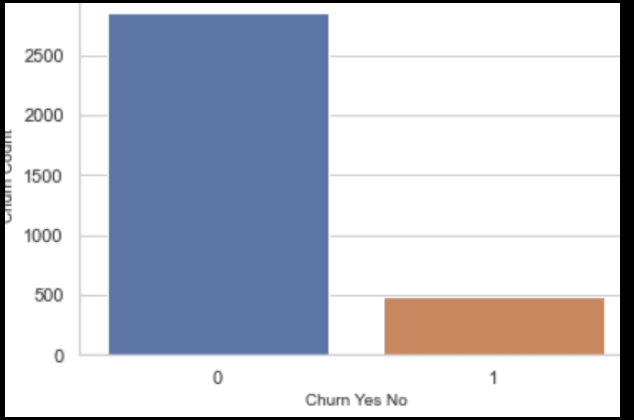
The data was obtained from Kaggle.

It has 3333 rows, 21 columns which include diverse customer-related information like state, account length, area code, phone number, voice mail plan, and more.

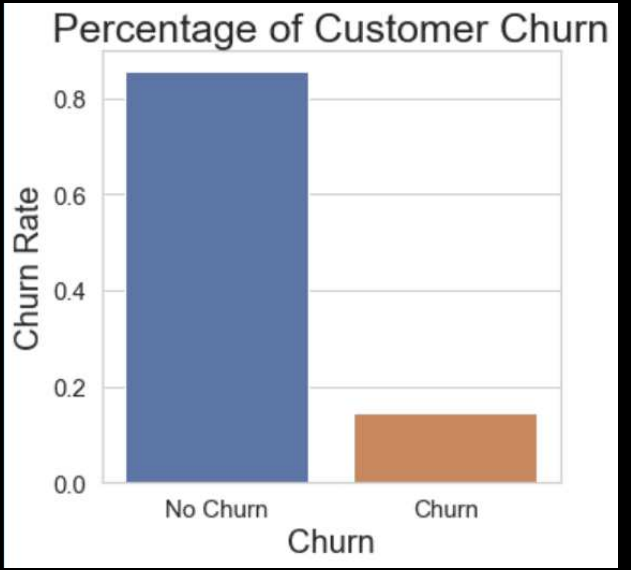
Some of the Key columns are:

1. Churn:Indicates if a customer terminated their contract.
2. International Plan: Indicates whether the customer has an international plan.
3. Customer Service Calls: Number of calls made to customer service.

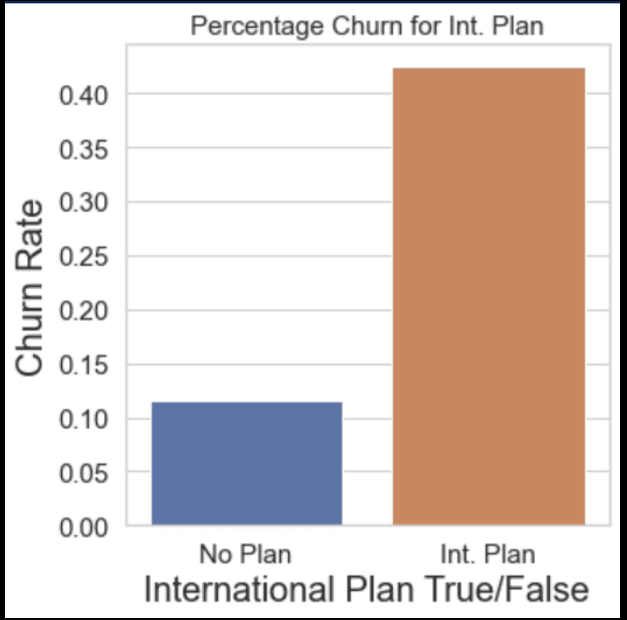
DATA ANALYSIS



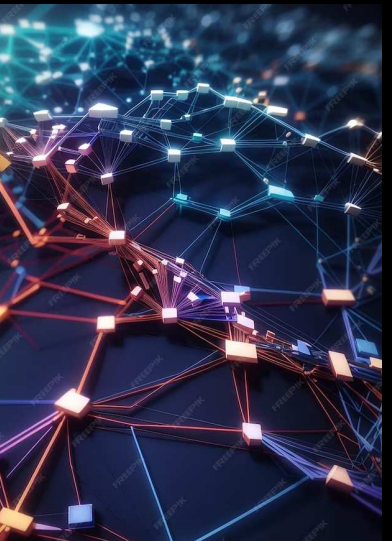
1. These shows that 2850 customers are likely to be retained while 483 are likely to churn



From the visualization, 85% of the customers are retained while 15% are likely to churn.



1. This shows that roughly 42% of customers with an international plan end up churning.



MODELING

The analysis encompassed various classification models, including Logistic Regression, Random Forest (optimized through hyperparameter tuning), and XGBoost, each contributing distinct approaches to predicting customer churn.

MODEL RESULTS

MODEL	ACCURACY	RECALL	F1 - SCORE
BASELINE MODEL	0.87	0.58	0.61
RANDOM-FOREST	0.89	0.66	0.70
XGBOOST	0.91	0.73	0.78

MODEL INTERPRETATIONS

The interpretations for the models:

1. Baseline Model: Achieved an accuracy of 87%. It correctly identified 58% of churn cases. Its F1-Score, a balance between precision and recall, was 61%.

2. Random Forest: Improved accuracy to 89%, identifying 66% of churn cases accurately. It showed a balanced F1-Score of 70%, indicating a reliable prediction balance.

3. XGBoost: Performed the best with an accuracy of 91%. It identified 73% of churn cases and displayed the highest F1-Score of 78%, indicating robust predictive capability.

CONCLUSIONS

1. **Model Performance**: XGBoost led with 91% accuracy, indicating strong predictive capability.
2. **Churn Prediction Insights**: Models help identify at-risk customers, aiding proactive retention strategies.
3. **Key Churn Factors**: Customer service calls, international plan, and day minutes were influential factors.

RECOMMENDATIONS

- **Tailored Retention Strategies**: Implement personalized strategies for high-risk churners.
- **Enhance Customer Service**: Improve service quality to boost satisfaction and loyalty.
- **Model Refinement**: Continuously update models for reliability and accuracy.
- **Focus on International Plans**: Explore targeted strategies for retaining international plan users

NEXT STEPS

- **Implement Strategies**: Roll out tailored retention initiatives based on model insights
- .**Enhance Service Quality**: Invest in better customer service to boost loyalty.
- **Model Refinement**: Continuously update models with fresh data for accuracy.
- **International Plan Focus**: Develop strategies to retain international plan users effectively.

Thanks!