



Predicting Customer Churn in SyriaTel

BY: CHRISTINE MALINGA

Business Understanding

- ▶ SyriaTel, a company in the telecommunication industry, has approached us with a pressing challenge: high customer churn rates are leading to financial losses and decreased customer satisfaction. To combat this issue, SyriaTel aims to accurately predict which customers are likely to leave, allowing them to take proactive measures to retain them. However, traditional methods for predicting customer churn have proven to be unreliable, with high rates of false positives and false negatives.
- ▶ In light of this, SyriaTel has tasked us with developing a machine learning classification model that can accurately predict customer churn based on historical usage patterns and demographic information. Our goal is to create a model that will help SyriaTel minimize financial losses and improve customer satisfaction.

Problem Statement

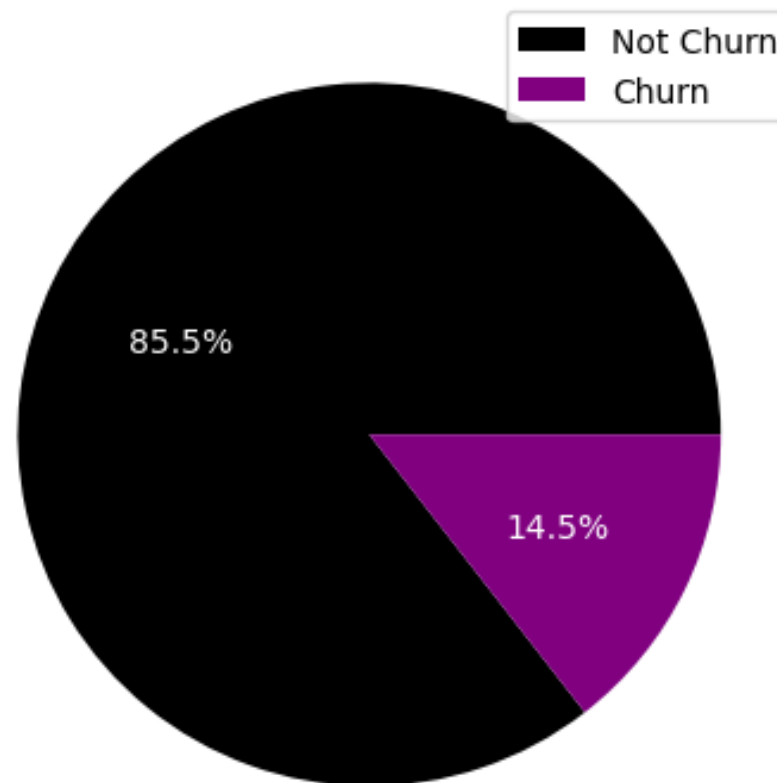
- ▶ The telecommunication industry is facing a challenge of high customer churn rates, leading to financial losses and decreased customer satisfaction. Traditional methods for predicting customer churn have high false-positive and false-negative rates, making it difficult for companies such as SyriaTel to accurately identify which customers are at risk of leaving and take proactive measures to retain them.

Objectives

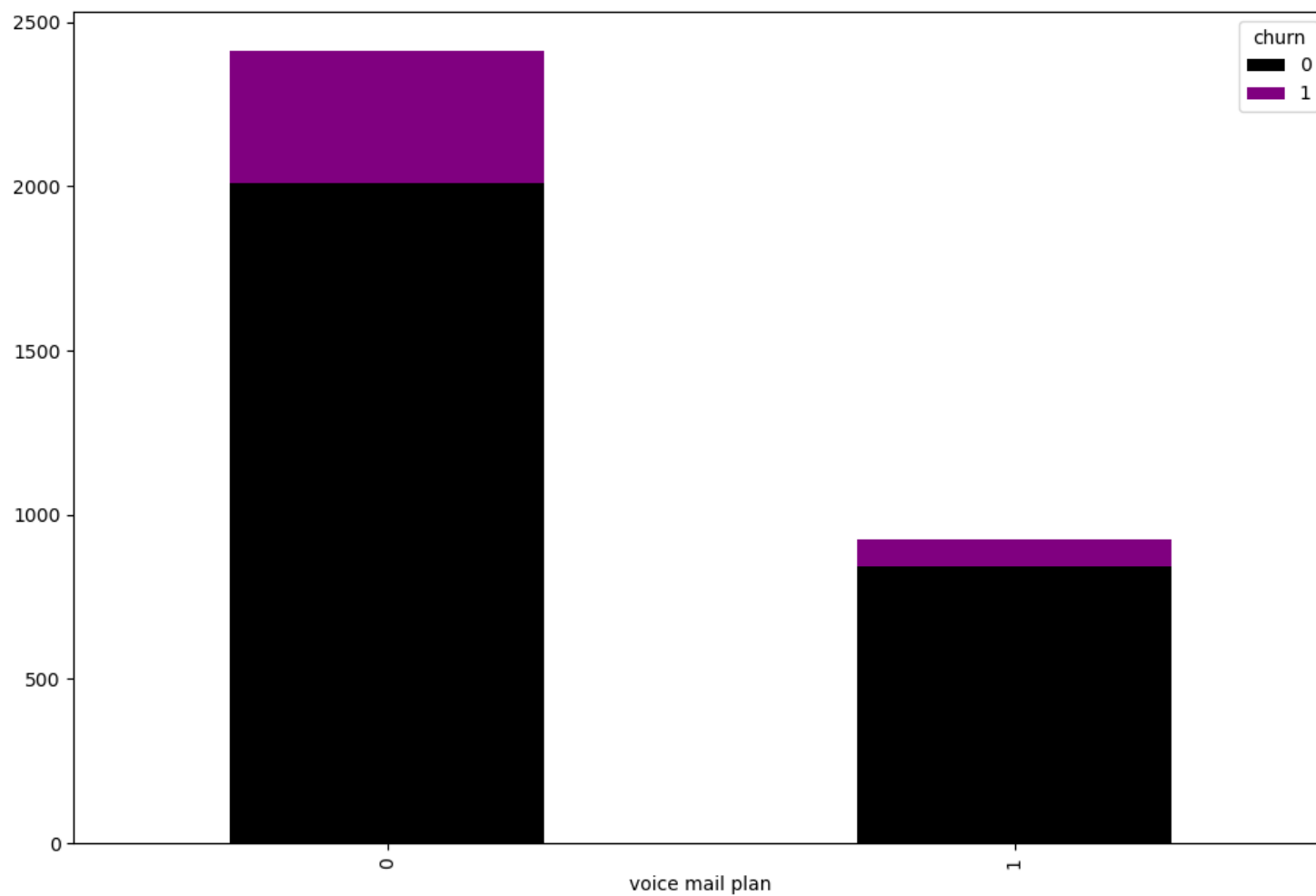
- ▶ The primary objective of this project is to reduce customer churn, which is the loss of customers to competitors. By predicting which customers are at risk of leaving
- ▶ Identify which features/predictor variables affect the attrition of customers.
- ▶ Build 3 Classification models and evaluate the best one for classifying and predicting the churn rate

Churn Breakdown

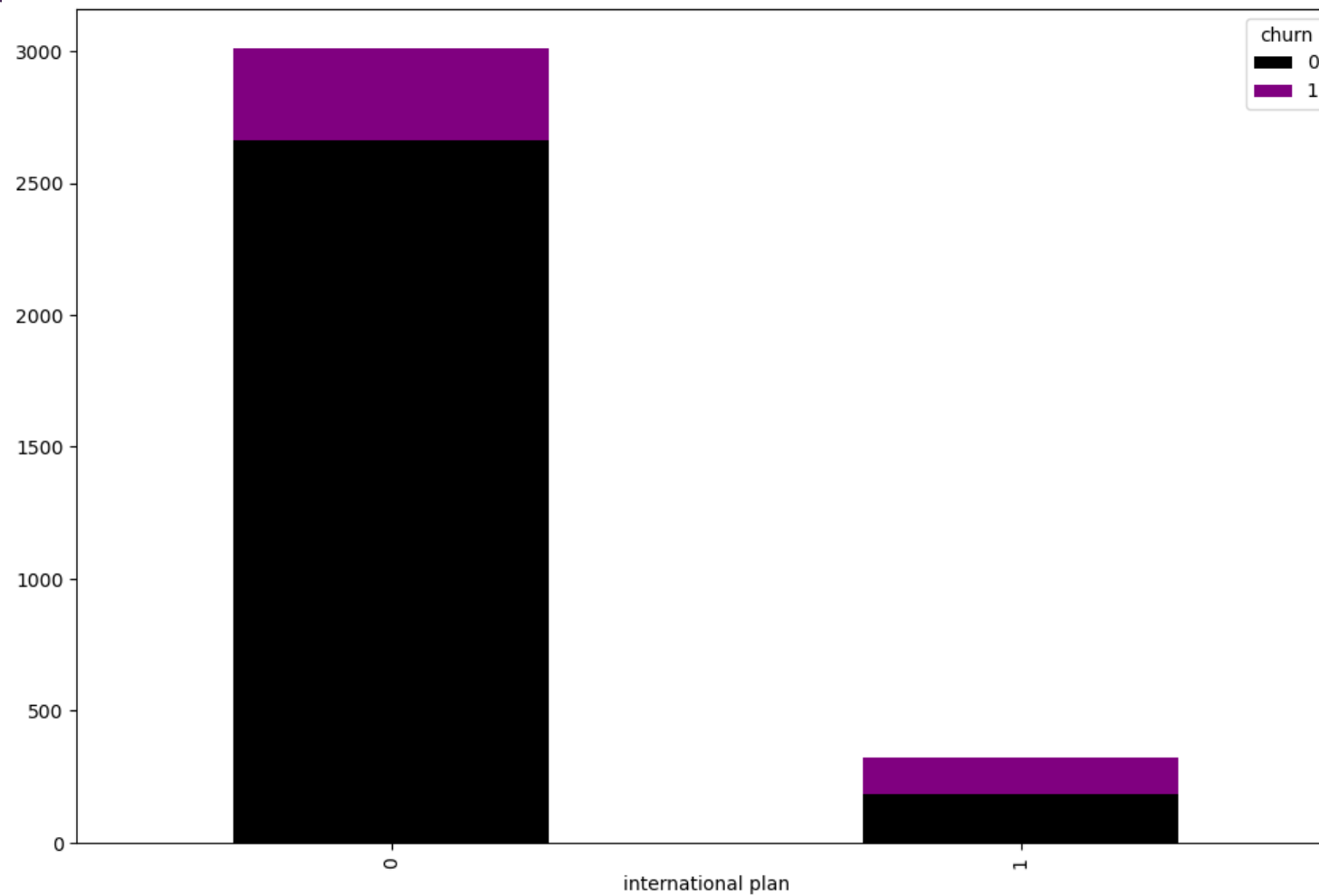
Customer Churn



Churn by voicemail plan



Churn by international plan



Models tested

- ▶ Logistic regression classifier model
- ▶ Adaboost Classifier model
- ▶ Gradient Boosting classifier mode
- ▶ K-Nearest Neighbors
- ▶ Decision Tree
- ▶ Random forest classifier

Best performing model

Cross validated Random Forest

Mean Training Score: 91.38%

Mean Test Score: 94.90%

The model was accurate over 90% of the time

Conclusion

These results indicate that the random forest model performed well in terms of accuracy. The model's performance in predicting class 0 (not churned) was slightly better compared to class 1 (churned), with a higher precision.

Overall, the random forest model was considered the best model for predicting the churn rate.

Recommendations

- ▶ Feature Engineering:
It's possible to enhance the prediction of churn by creating new features from the existing data. For example, a feature for the average daily or monthly charge, a feature for the average call duration or a feature for the region covered instead of many states could be generated. These new features could offer deeper insights into customer behavior and contribute to a better prediction of churn.
- ▶ Data collection:
Undertake a customer satisfaction survey to collect info on how the customers rate Network Coverage and Quality, customer service and improve on these areas.