

SYRIATEL CUSTOMER CHURN

Author: Scholar Chepkirui

Project overview

This project uses machine learning algorithms to build a model that can accurately predict customers who will churn based on the information available in the dataset. Classification algorithms are used to build the predictive model since the target variable, 'churn', is a categorical variable.

Recall is used to evaluate the model's performance.

Business problem



It is less expensive to keep a current customer than acquire a new one.

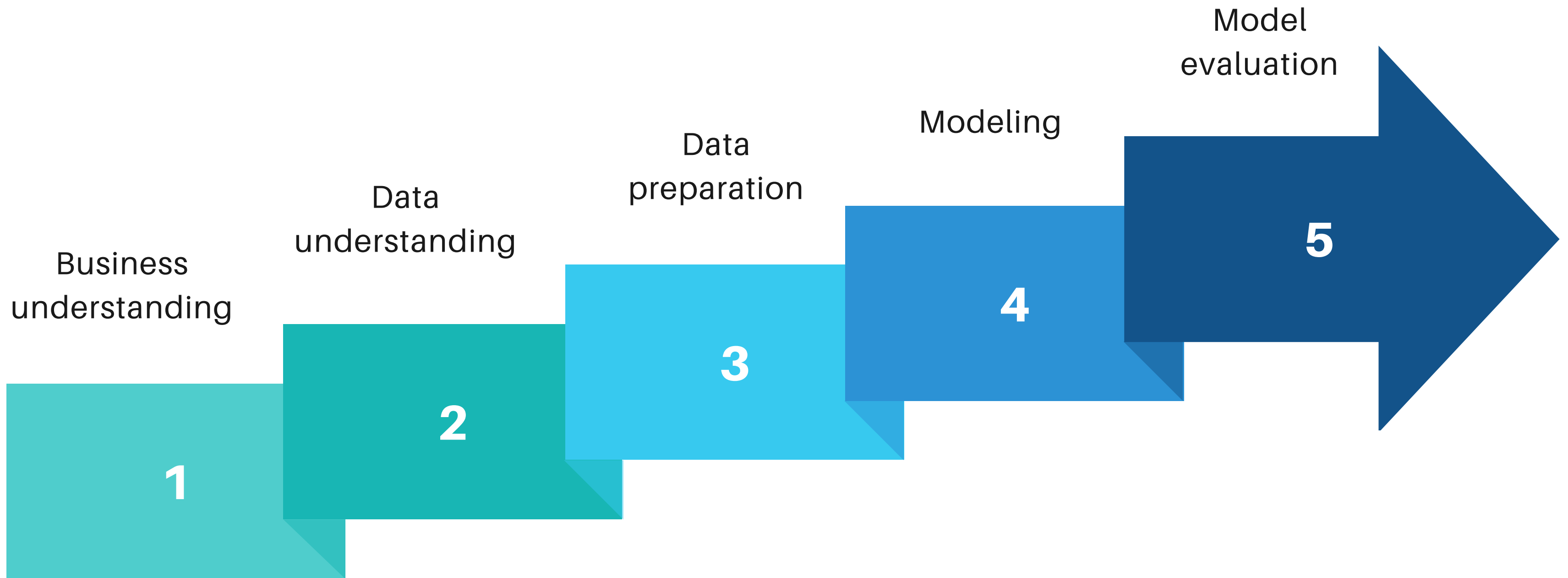


5 times more used in acquiring a new customer.

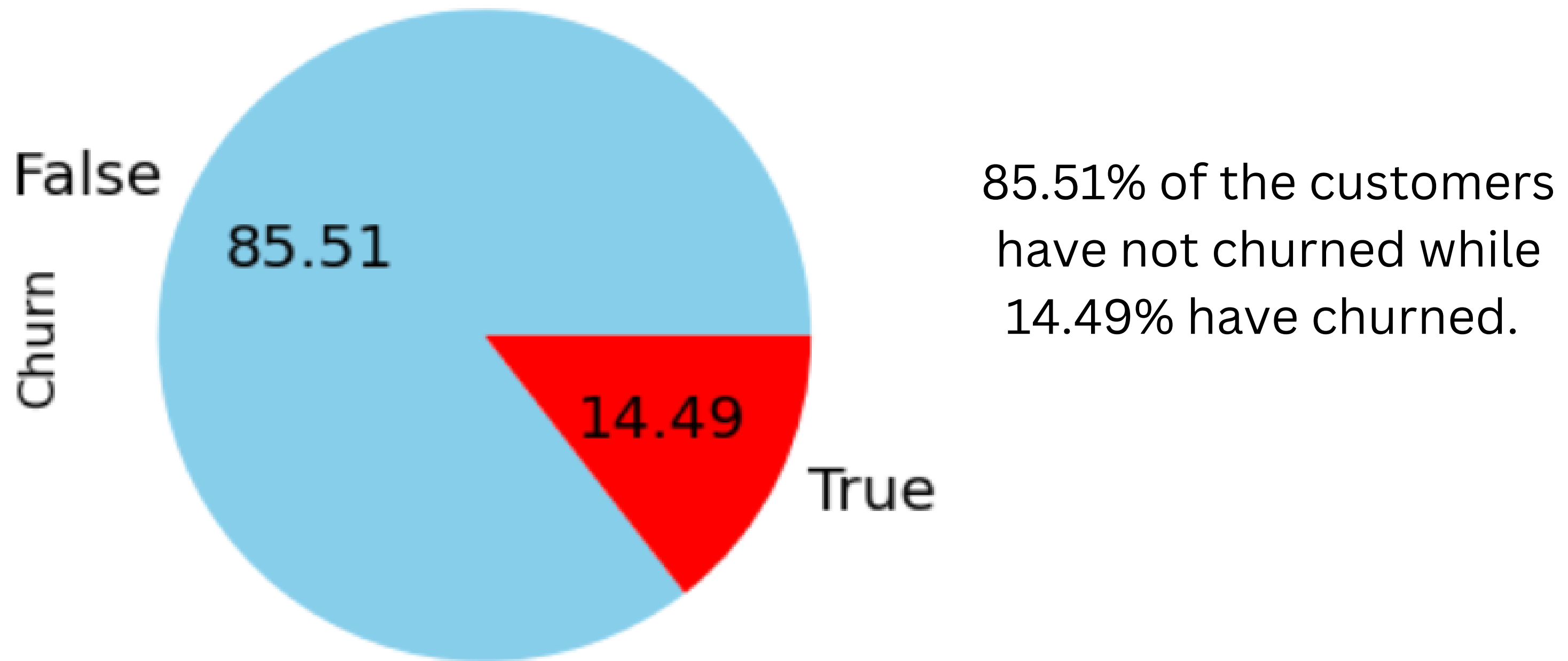


5% increase in customer retention can grow profit by 75%

Method



Analysis of target variable

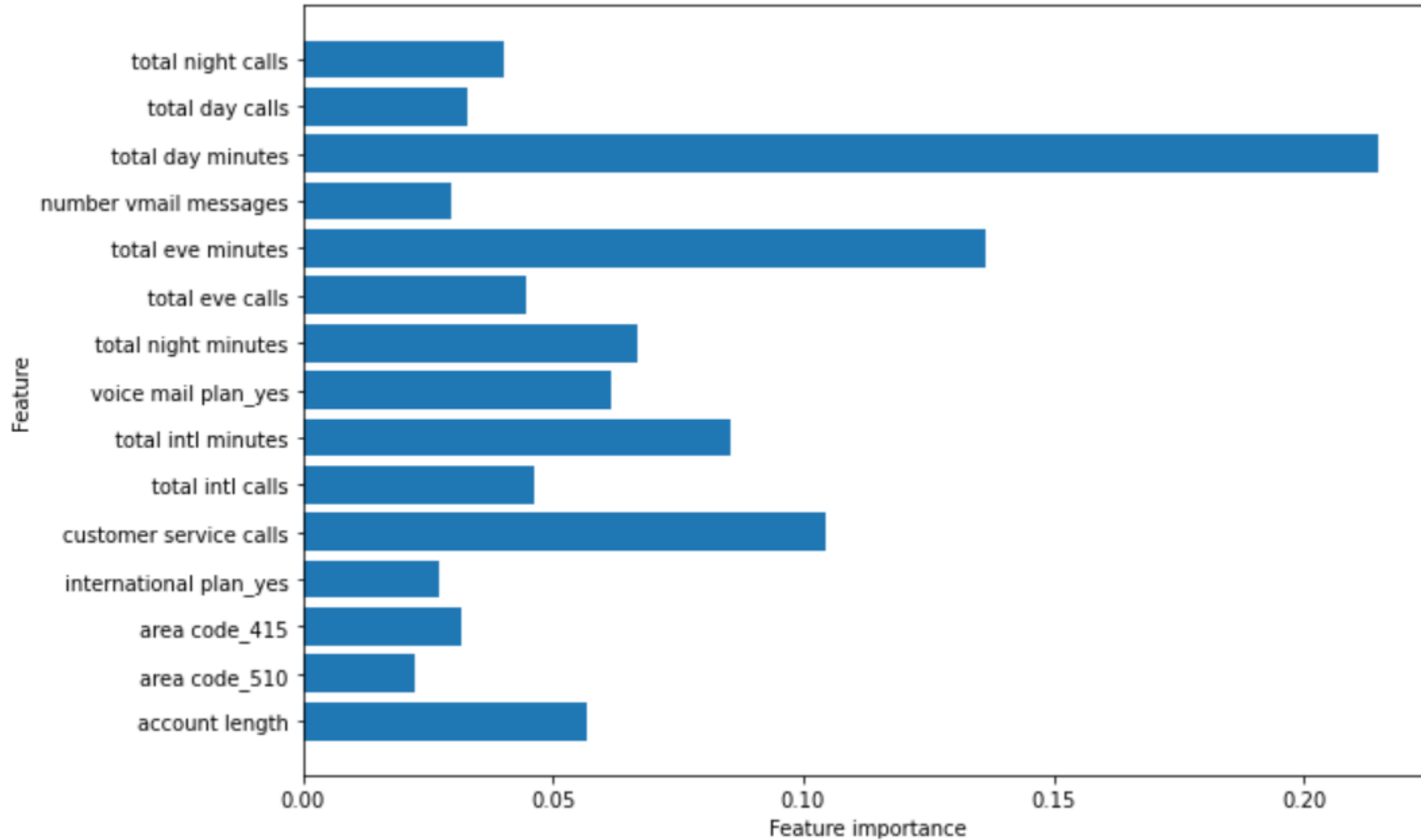


Results

Model	Recall	Precision	Accuracy
Logistic regression	0.25	0.51	0.87
Decision tree	0.75	0.49	0.86
Decision tree(tuned)	0.79	0.52	0.88
Random forest	0.67	0.69	0.92
Random forest(tuned)	0.68	0.67	0.91

Decision tree with tuned hyperparameters is the model with the best performance. It has the highest recall score.

Results



The most important features for predicting customer churn:

- total day minutes
- total evening minutes
- customer service calls
- total international minutes

Conclusions

Syriatel should ensure effective customer service so as to meet customers' expectations and analyze customer interactions.

The company should look into the call charge rates in comparison to the competitors, and consider if they should lower the charges of calls per minute. This can prevent other customers from churning.

Next Steps

The training data size should be increased to reduce overfitting and hence improve model performance.

