



Customer Churn Data

FARNAZ GOLNAM

In this project we are going to build a classifier to predict whether a customer will ("soon") stop doing business with SyriaTel, a telecommunications company.

We are presenting a predictable pattern to the telecom company to prevent them from losing money on current/future customers who don't stick around very long.

Finally , we offer a financial cost/benefit analysis with regard to leaving/staying costumers rate to help them make the most benefit out of their offered services.

INTRODUCTION

Main questions to answer:

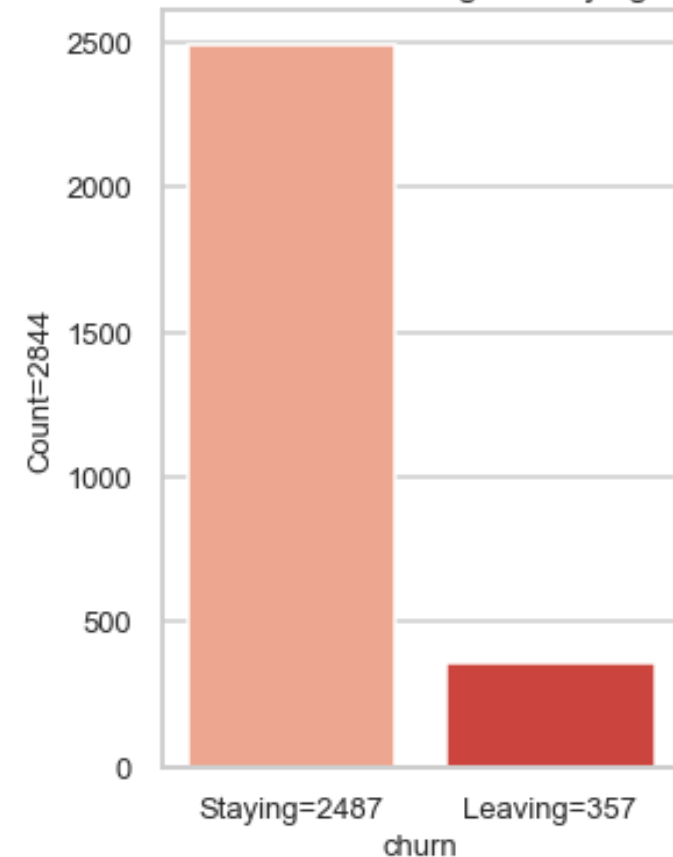
1. what are the most important factors that cause costumers to leave?
2. with the same offers and services, does the location of customers have any impact on their leaving ?
3. what is the current rate of customers leaving the company and what is the predicted rate for future?
4. how to offer services and discounts to keep making benefits despite the leaving customers?

Current rate of leaving customers

Out of 2844 customer observations:

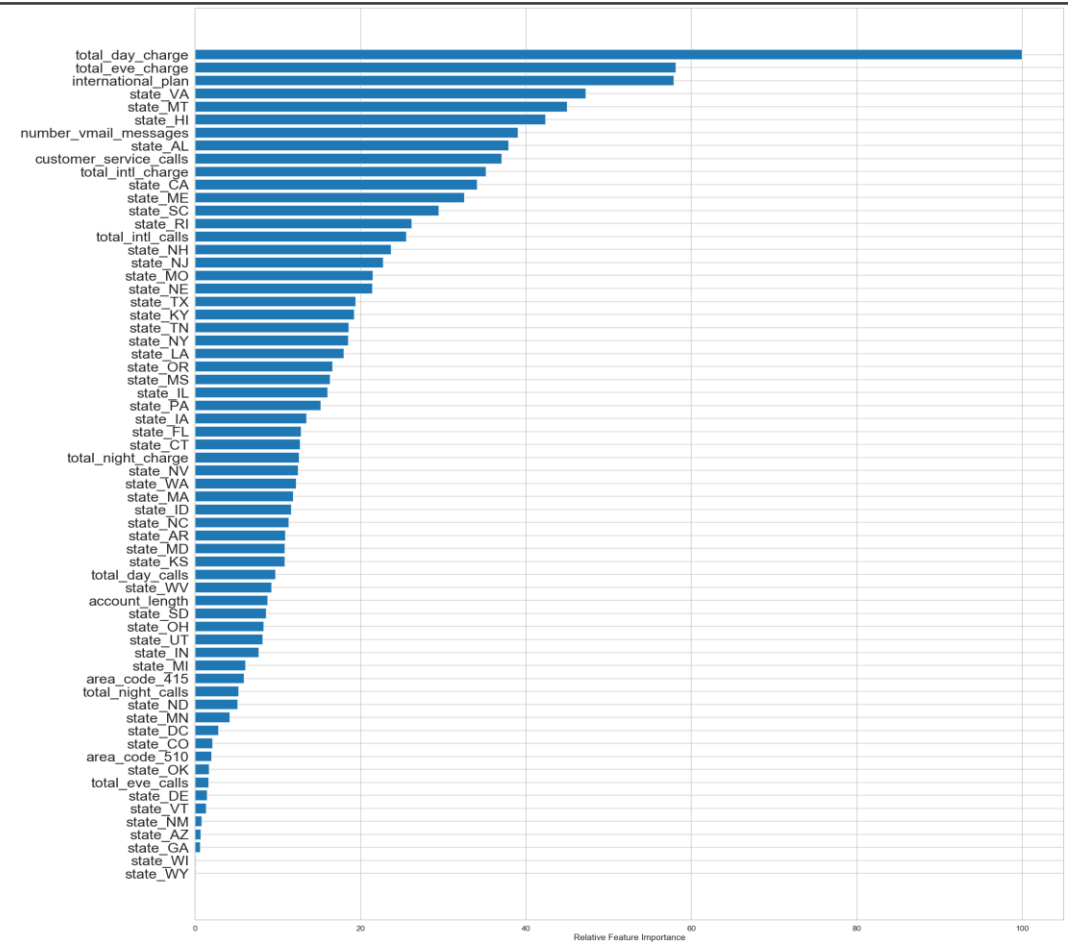
1. 357 customers , around 12.5% are cancelling their contract with the telecom company.
2. 2487 customers, around 87.5% are continuing their business with the telecom company.

Number of Customers Leaving vs Staying in Business

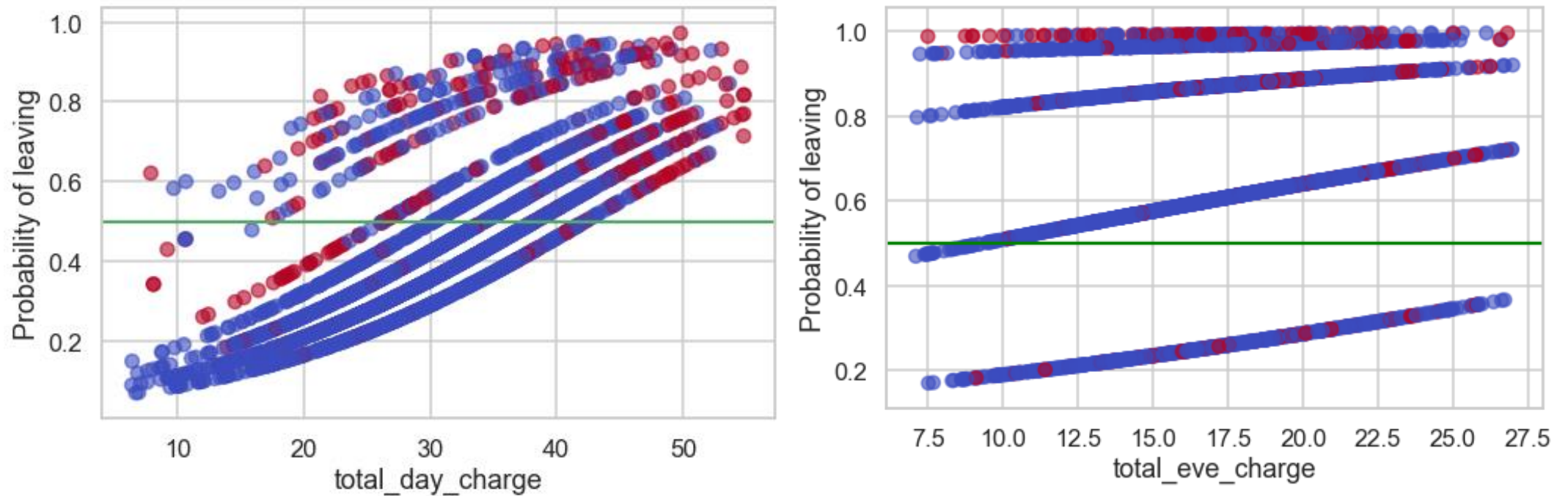


Relative Feature importance :

1. International plan
2. Total day charge
3. Total evening charge
4. Customer service calls
5. Number vmail messages
6. Total international charge
7. State/location

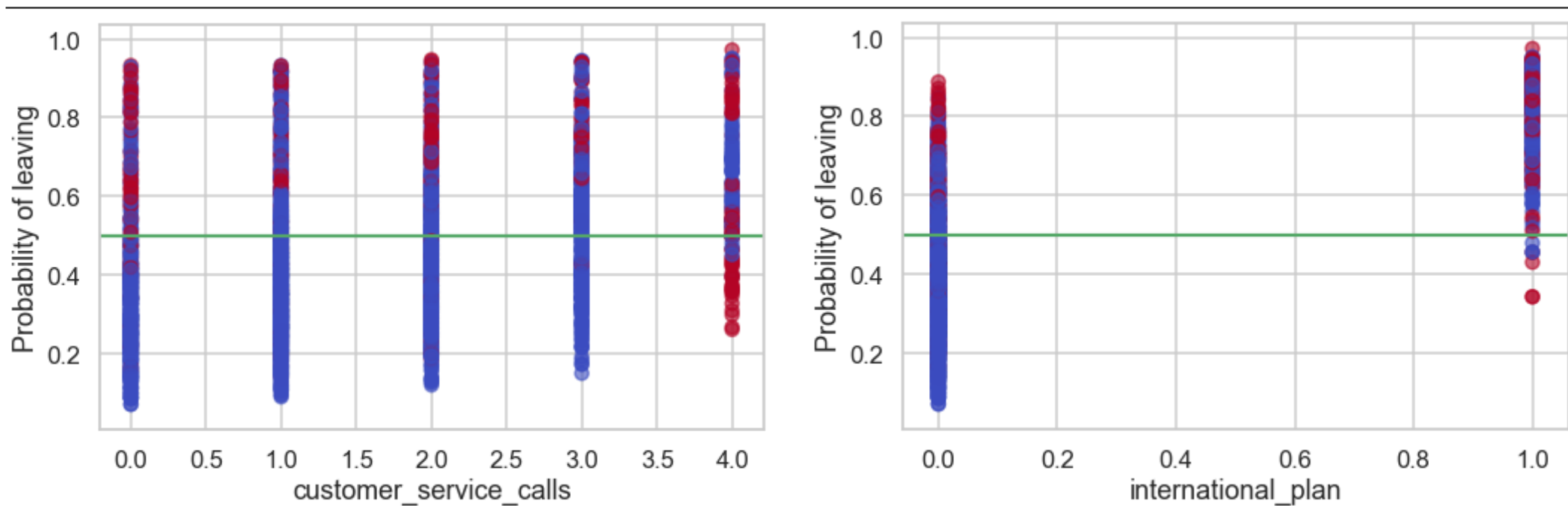


Visual comparison of the relationship between different features and the probability of leaving:



As shown in the scatter plots, the probability of cancellation shown by red dots goes higher as the day charge rate goes higher, but the evening rate does not have as strong affect as day charge rates on cancelling the service.

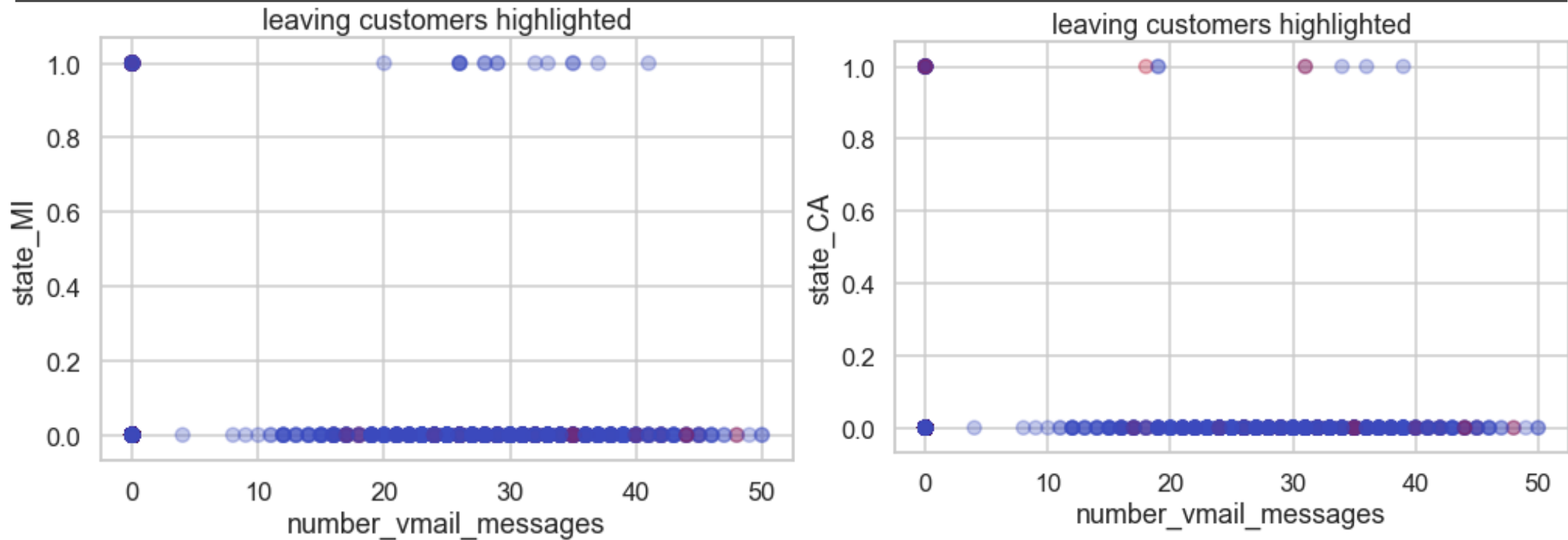
Visual comparison of the relationship between different features and the probability of leaving:



As shown in the scatter plots, the probability of cancellation shown by red dots has a reverse relation with the number of customer service calls.

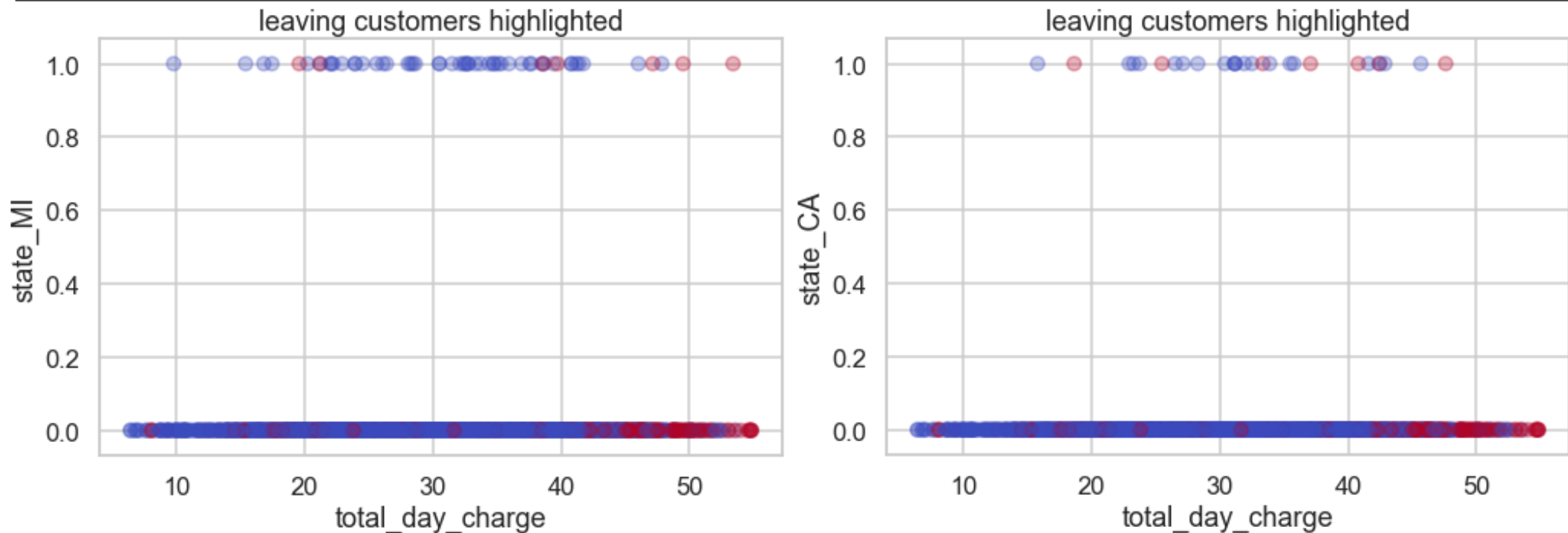
Also, having no international plan increases the chance of leaving the company.

For customers using the same plans and services, does location plays a role in leaving the company?



Just to compare two random state, it can be seen that, offering the same number of vmail messages, customers in state CA have a higher probability of leaving the company compare to state MI.

For customers using the same plans and services, does location plays a role in leaving the company?



To mention another important feature, as shown above, offering the same rate for day charge, customers in California have a higher probability of leaving the company compare to state MI.

The financial impact of the predicting model on the SyriaTel, telecommunications company and our Recommendations:

The default assumption is that we don't know the budget of company to distribute promotions and offers or their ability to change any charge rates, thus we suggest our formula based on the raw input.

The **policy** can be offering promo plans to potential leaving customers to motivates them to stay:

Promo Plan to keep customer motivated = [(potential yearly promo) + (12* potential monthly discount)]

1. True Negative predictions of model(TN) : cost

Revenue loss/year = TN x (\$monthly contract rate*12)

The company will lose its customer through canceling the service and discontinuing the subscription.

2. False Positive predictions of model(FP) : cost

Revenue loss/yearly = FP x (\$monthly contract rate*12)

The model predicts that SyriaTel costumer will continue doing business with the company but they don't ! In this case, the model will instruct the website not to offer a promo plan or discount, thus it will lose the opportunity to keep making revenue from a potential customer.

The financial impact of the predicting model on the SyriaTel, telecommunications company and our Recommendations:

3. False Negative (FN) : Benefit

Benefit on revenue & loss on promo plans = FN x [(\$ monthly contract rate*12) - (\$ promo plan)]

SyriaTel company will earn money in revenue if the model predicts a customer will be leaving , thus a promo plan or discount is offered to motivate the customer to stay. the customer is actually staying so the company does not lose the revenue but it will lose money on the promo plans and discounts.

4. True Positive (TP) : Benefit

Revenue gain/year = TP x (\$ monthly contract*12)

SyriaTel company will earn in revenue if our model correctly predicts a customer continues doing business with them, thus no discount or promo is offered.

The financial impact of the predicting model on the SyriaTel, telecommunications company and our Recommendations:

Benefit equation:

$$\text{TP} \times (\$ \text{ monthly contract} \times 12) - \text{TN} \times (\$ \text{ monthly contract} \times 12) - \text{FP} \times (\$ \text{ monthly contract} \times 12) + \text{FN} \times [(\text{monthly contract} \times 12) - (\text{promo plan})]$$

So the predicting model will best benefit the company if limits the false positive predictions considering most important features as below:

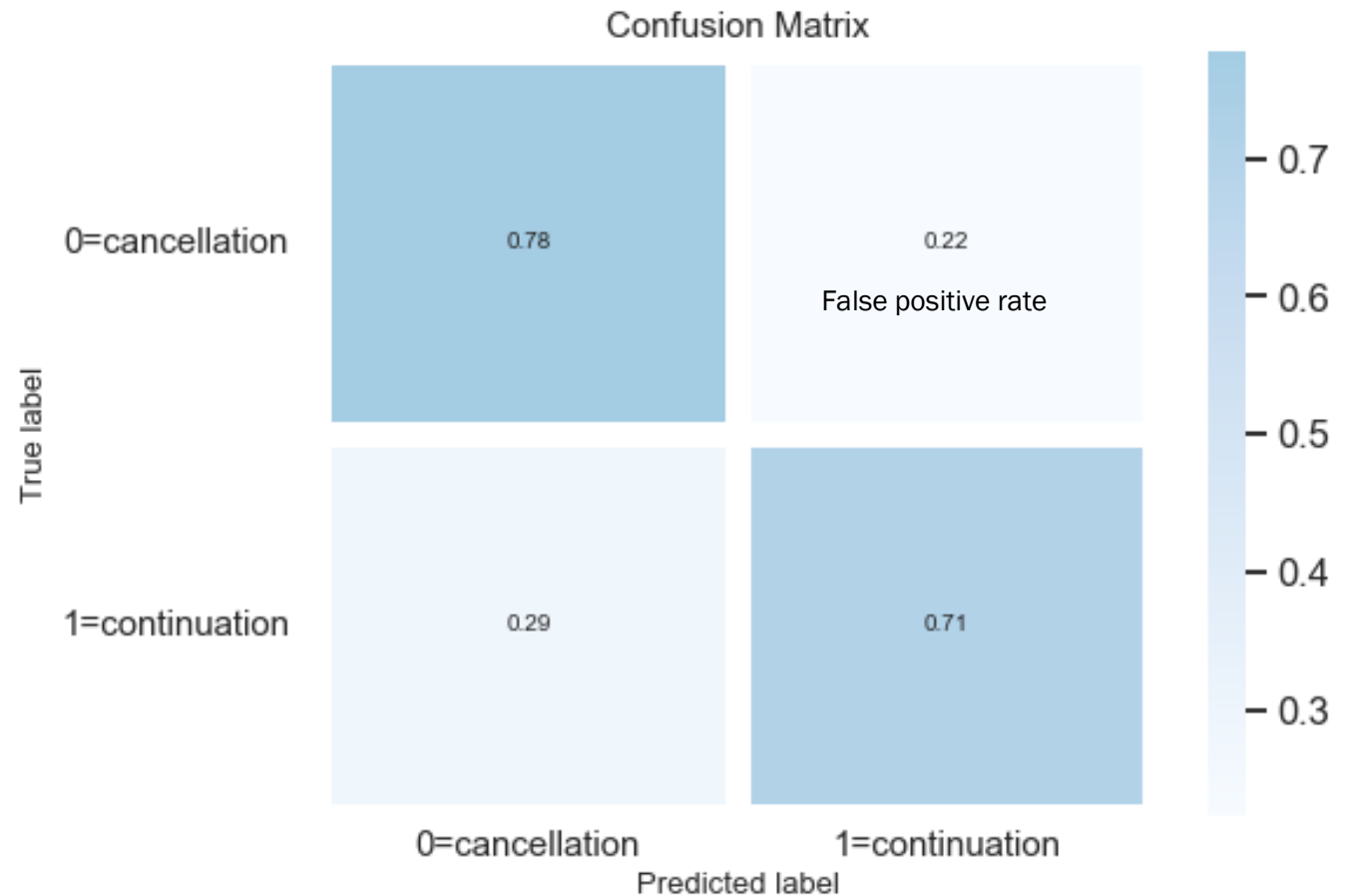
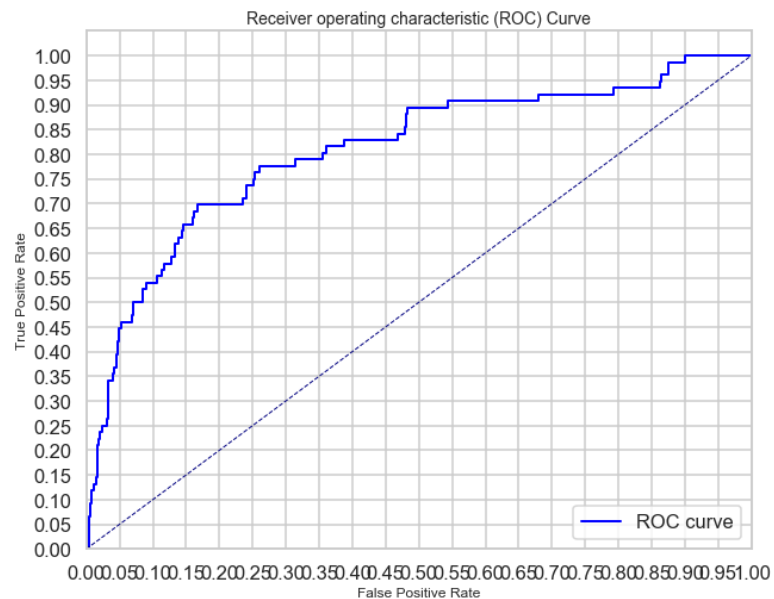
1. International plan
2. Total day charge
3. Customer service calls
4. Number vmail messages

Suggested Predicting model

Model Test Data Precision: 0.345

Model Test Data Accuracy: 0.738

Model Test Data AUC: 0.815



Future Work:

1. Present other classification algorithms and compare the results to our existing predicting model. Decision Trees, Random Forests, and Support Vector Machines are a few other classifiers to consider testing.
2. Defining a function to show the affect of location and each specific state on the leaving probability when all other features are the same .
3. Identifying the best promotions to offer to higher the benefit and lower the leaving probability.

The End

Thank you for your attention

Farnaz Golnam