Case Study – Customer Churn Prediction

Q1. Clearly state the business case (who/problem/solution/payoff matrix).

Business Case –
Who - Retention team at the telecommunications provider company.
Problem - Customers switching to another telecommunications provider for better deals or services. Telecommunication companies want to minimize this churn rate which currently is around 14% as they cannot just give offers to everyone.
Solution - Giving the customers offers or incentives for continuing with the same telecommunications provider.

The dataset used for modeling has 4250 records. The models are required to predict whether a customer will change telecommunications provider or not, also known as churning.

Payoff matrix –

Assumptions –

1. Average plan is for 12 months (1 year) meaning that the customer stays for 1 year.
2. Average cost of the telecommunication provided plan for customers is $25/month.
3. Retention team offers a discount of 25% on its yearly plan to its customers to make them stay. They offer package of '3 months free on purchase of 9-month plan.'
4. Success rate for retention is 100% which means that all customers to whom the offer is made accept it and stay with the same provider.
5. Cost associated with making calls to offer the discount to customers is $60/hour and they can make 5 calls per hour. Therefore, each such call costs them $12.

True Positive (TP) - Model predicts customer would churn and hence company needs to act like giving them discount to make them stay which in our case is set to 25% on the yearly plan. And under the assumption that success rate is 100%, the customer stays therefore adding revenue to the company.
TP = (cost of plan) * (average plan duration) * (1 - discount provided) – (cost of calling the customer)
TP = 25 * 12 * 0.75 - 12 = 213

True Negative (TN) – Model predicts that the customer wont churn and hence takes no action. The customer still stays with the same provider.
TN = (cost of plan) * (average plan duration) = 25 * 12 = 300

False Positive (FP) – Model predicts that the customer would churn even though the customer was not planning to switch and hence the company provides discount to the customer. The customer accepts the discount and hence the company still makes revenue from this though it is lower than the original one.
FP = (cost of plan) * (average plan duration) * (1 - discount provided) – (cost of calling the customer)
TP = 25 * 12 * 0.75 - 12 = 213

False Negative (FN) – The model predicts that the customer would not churn and hence no outreach is made. There is no associated cost or revenue.
FN = 0

True Positive	False Positive
213	213
True Negative	False Negative
300	0

Q2. Perform exploratory data analysis, pre-process the data as necessary.

Target variable – churn(yes/no)
Churn value of yes is associated with a negative outcome which means that the customer leaves the current provider for some other provider and a churn value of no is associated with a positive outcome meaning that the customer stays with the current service provider.

Features selected for modeling – We have 4 categorical features and 16 numeric features.

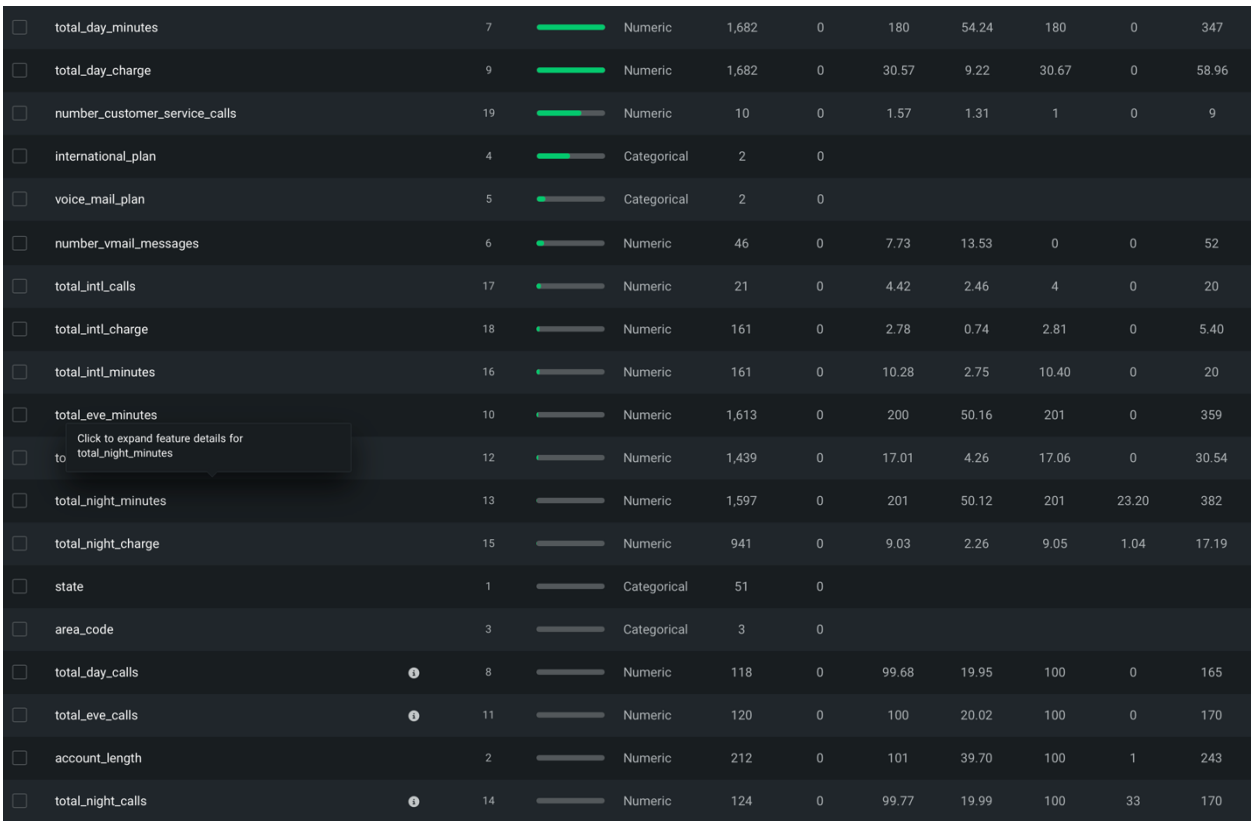


Fig: - Feature List

Missing values – There are no columns in the data with missing values.

Cardinality – None of the 4 categorial features has high cardinality with respect to the number of records available and hence we are using them for modeling.
Categorical features -

1. international_plan – 2 unique values
2. voice_mail_plan – 2 unique values
3. area_code – 3 unique values
4. State – 51 unique values – even though this number is large it is still relatively much less in comparison to the number of records which is 4250.

Q3. Make use of all the modeling techniques that you know to build models to forecast customer churn. Show a table with a summary of model performance metrics: Recall, Precision, F1, ROC AUC, max payoff. Identify the best performing model. Explain your choice of metric.

The below models were run on the above feature list –

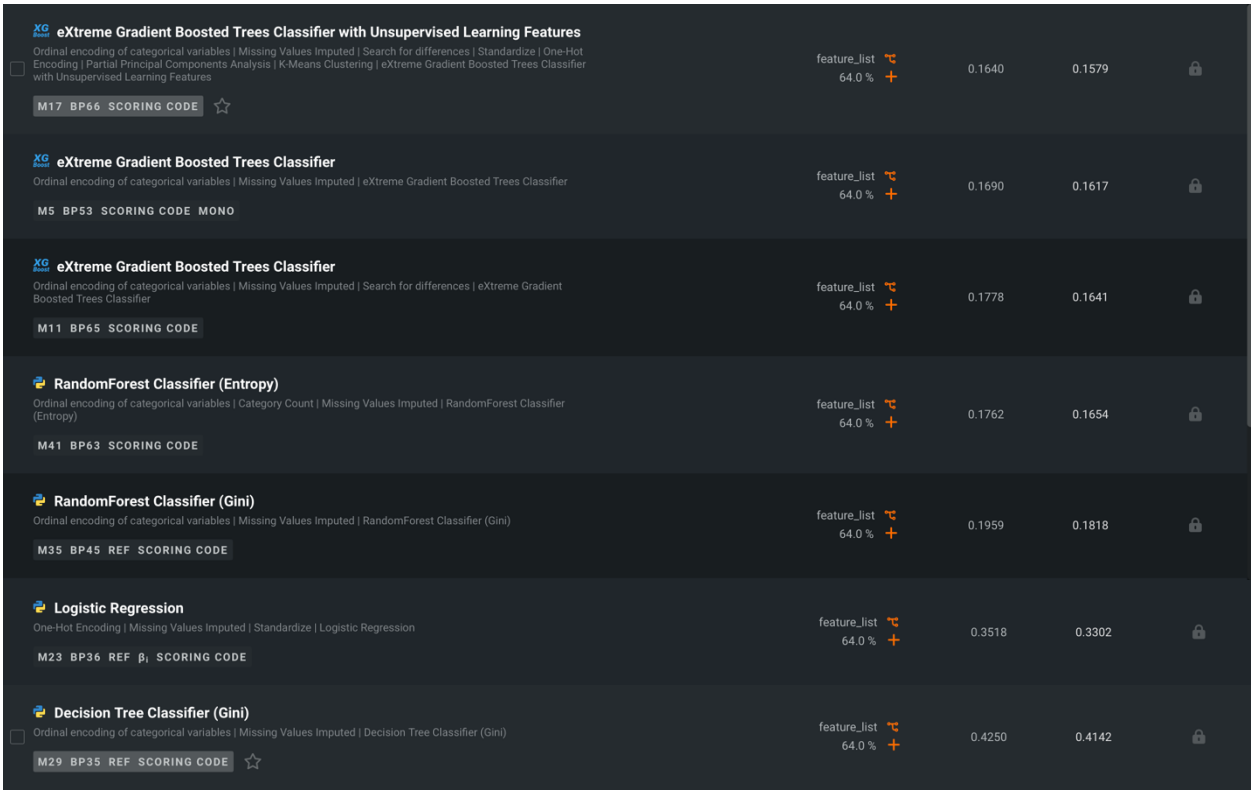
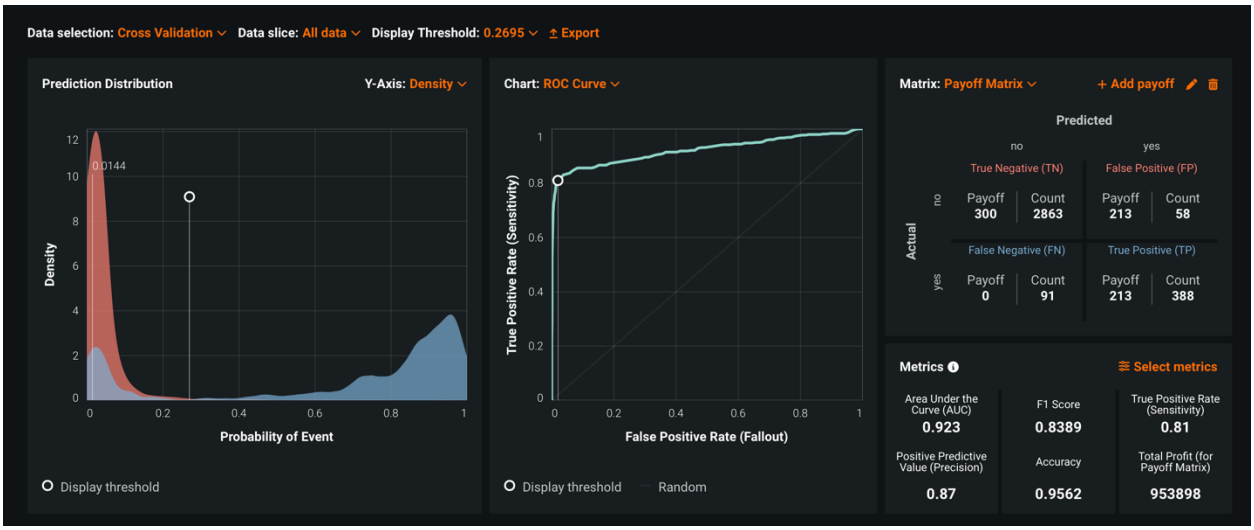
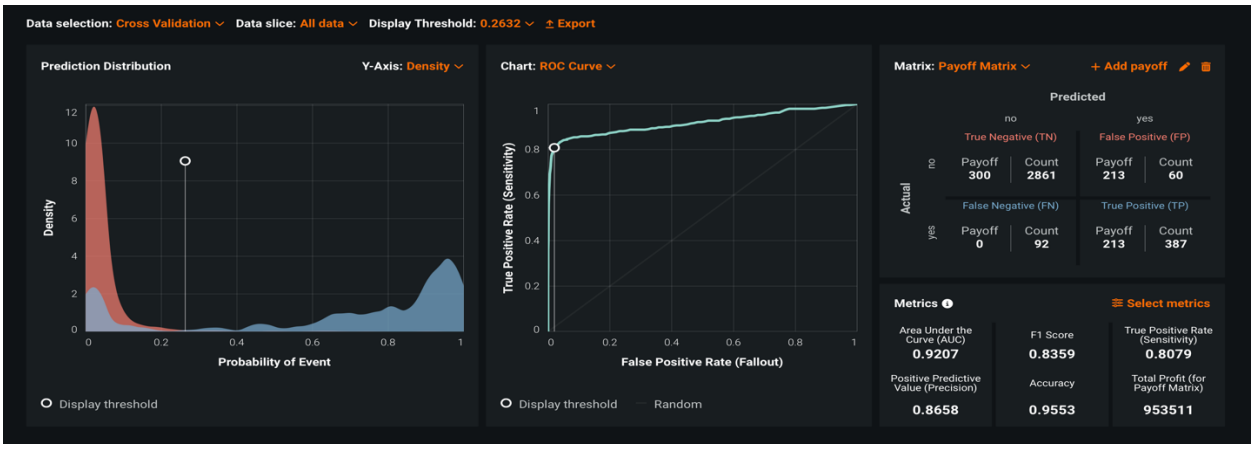


Fig: - Models trained on the data

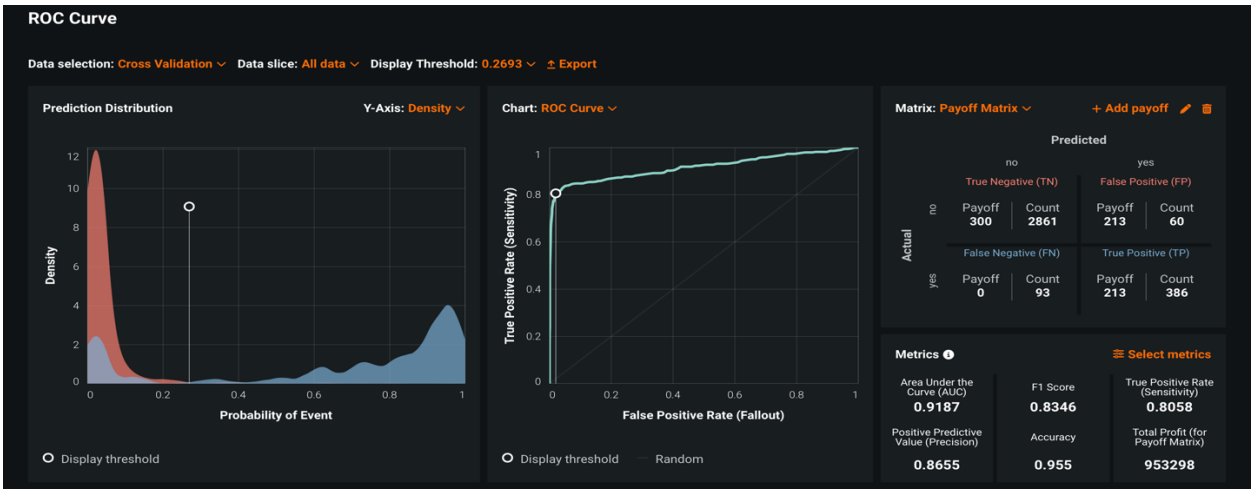
1. Model – eXtreme Gradient Boosted Trees Classifier with Unsupervised Learning Features



2. Model – eXtreme Gradient Boosted Trees Classifier BP53

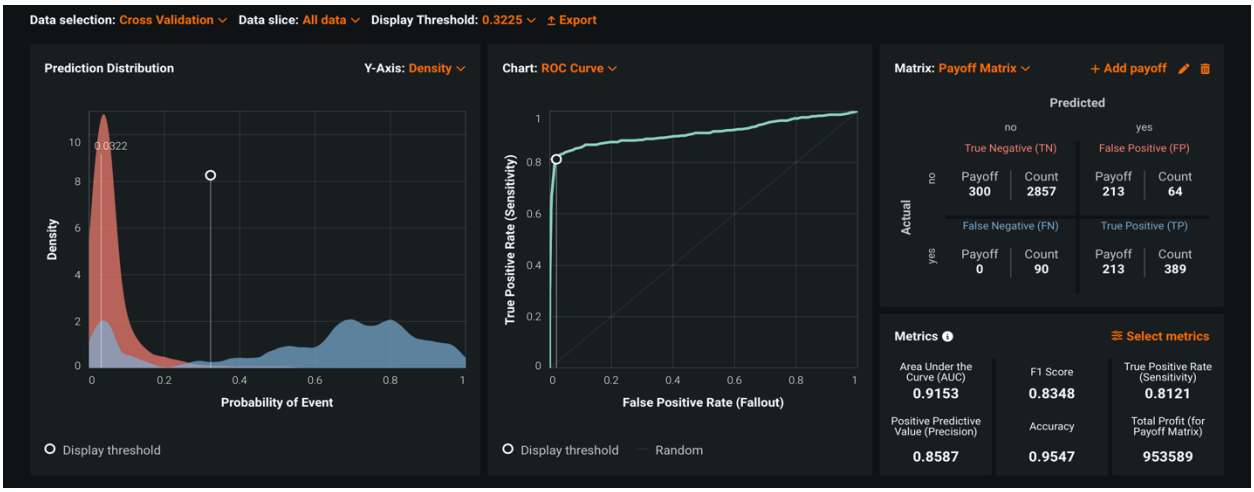


3. Model – eXtreme Gradient Boosted Trees Classifier BP65

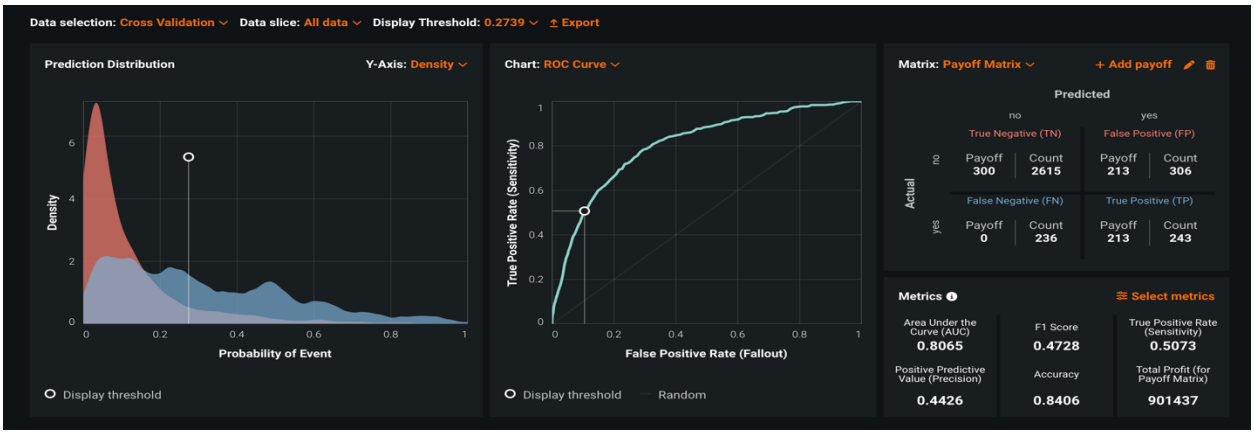


4. Model – RandomForest Classifier (Entropy)

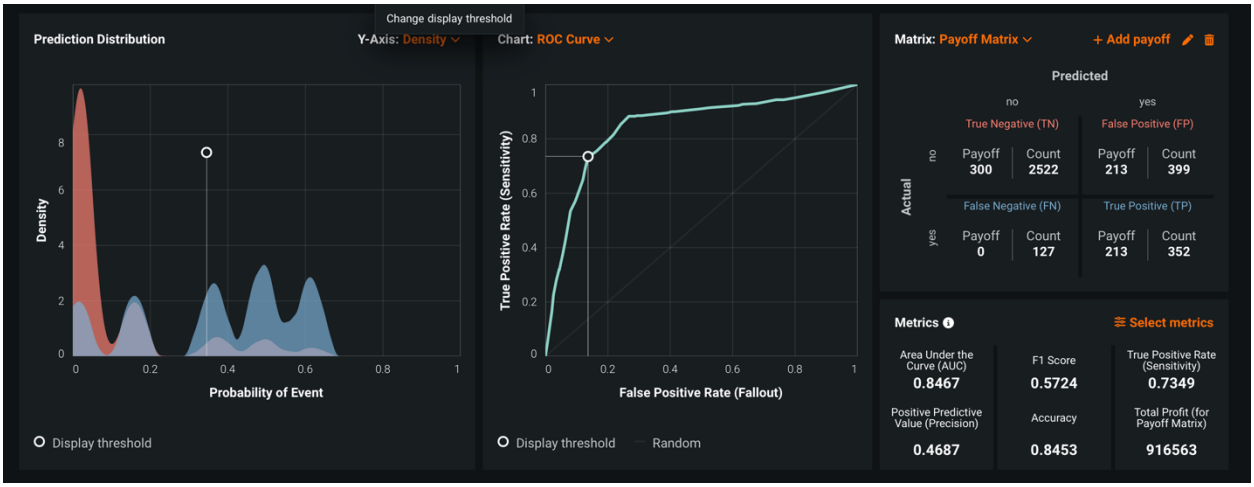
5. Model – RandomForest Classifier (Gini)



6. Model – Logistic Regression



7. Model – Decision Tree (Gini)



(All the below metrices are reported at threshold of maximum profit)

Models	Recall	Precision	F1-score	ROC AUC	Accuracy	Maximum payoff
eXtreme Gradient Boosted Trees Classifier with Unsupervised Learning Features	0.81	0.87	0.8389	0.923	0.9562	953898
eXtreme Gradient Boosted Trees Classifier BP53	0.8079	0.8658	0.8359	0.9207	0.9553	953511
eXtreme Gradient Boosted Trees Classifier BP65	0.8058	0.8655	0.8346	0.9187	0.955	953298
RandomForest Classifier (Entropy)	0.8246	0.8351	0.8298	0.9152	0.9524	953649
RandomForest Classifier (Gini)	0.8121	0.8587	0.8348	0.9153	0.9547	953589
Logistic Regression	0.5073	0.4426	0.4728	0.8065	0.8406	901437
Decision Tree (Gini)	0.7349	0.4687	0.5724	0.8467	0.8453	916563

The best model is the Boosted Tree Classifier as it has the maximum payoff of $953898. The metric chosen here is the maximum payoff as we want to maximize the revenue for the company by trying to retain the maximum number of possible customers. Our business case requires us to minimize the churn rate which means minimizing the number of customers who want to switch to other telecommunication providers for better services or deals by providing them lucrative offers. The payoff matrix assigns costs and revenues to each of the predictions (TP/TN/FP/FN) which helps us to evaluate the associated cost and revenue. The Boosted Tree Classifier helps to target the customers in a manner that has the maximum gains for the company hence making it the best model for our case.

Q4. Visualize the effects of the top 4 predictors of customer churn. Summarize the effects in 1-2 sentences.

For the best model, below is the list of features and their importance percentage in predicting customer churn –

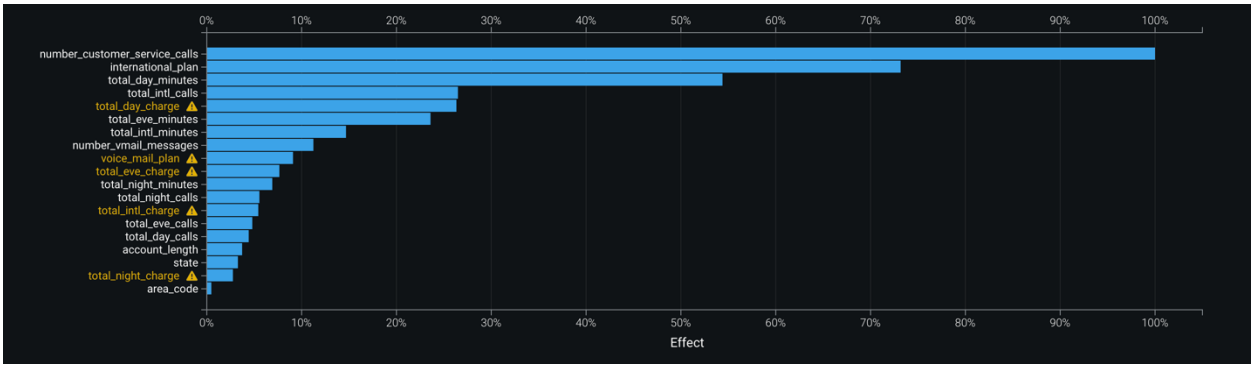
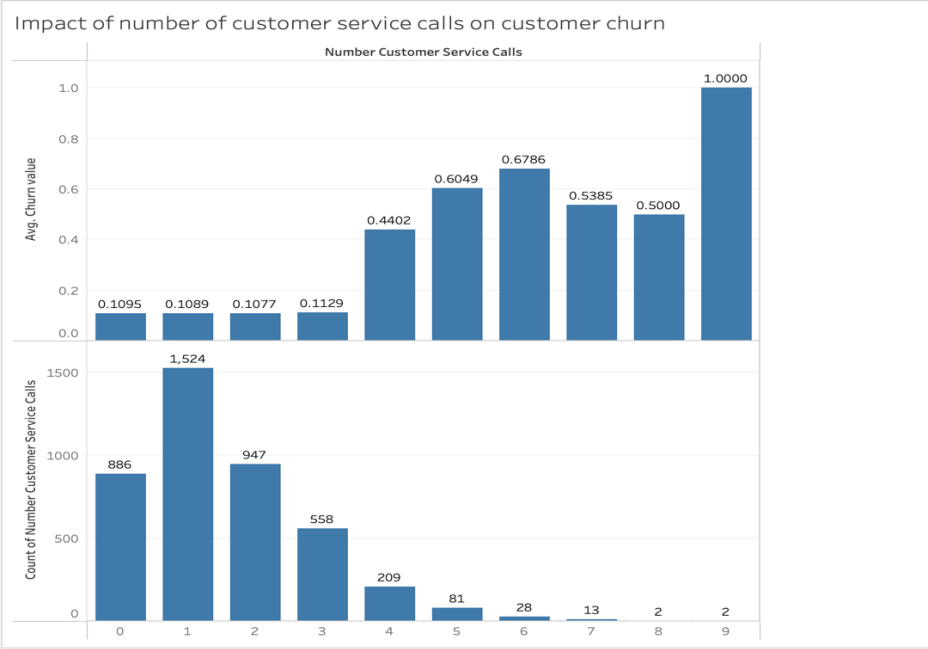


Fig: - Feature effect on churn rate

The top 4 predictors are –

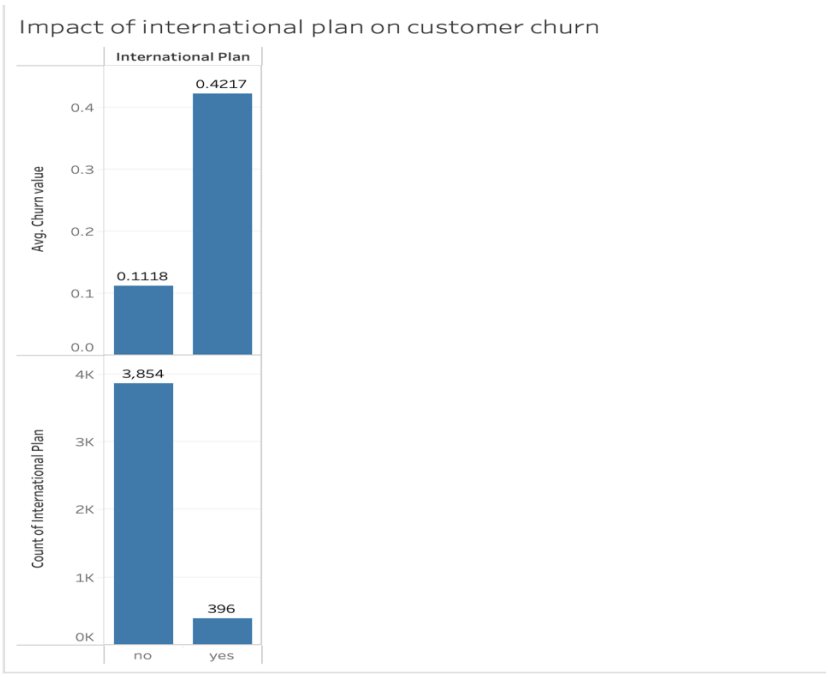
1. number_customer_service_calls - Number of calls to customer service
2. international_plan - The customer has international plan or not.
3. total_day_minutes - Total minutes of day calls.
4. total_intl_calls - Total number of international calls.

1) Effect of number_customer_service_calls on churn rate



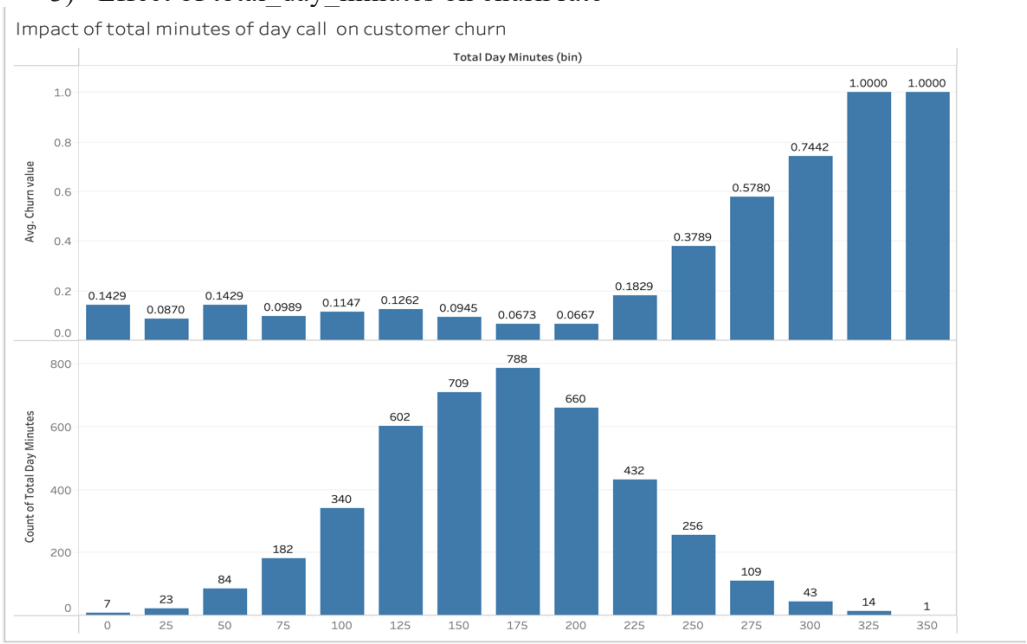
Churn rate is low for cases that have count of customer service calls within the range of 0-3 (almost 11.29%). But this increases from that to 67.8% as the number of calls increases from 3 to 6. The churn rate again decreases to around 50% if the number of calls made is 7 or 8. However, for 9 calls it comes up as 100% but we don't have enough data to justify this.

2) Effect of international_plan on churn rate



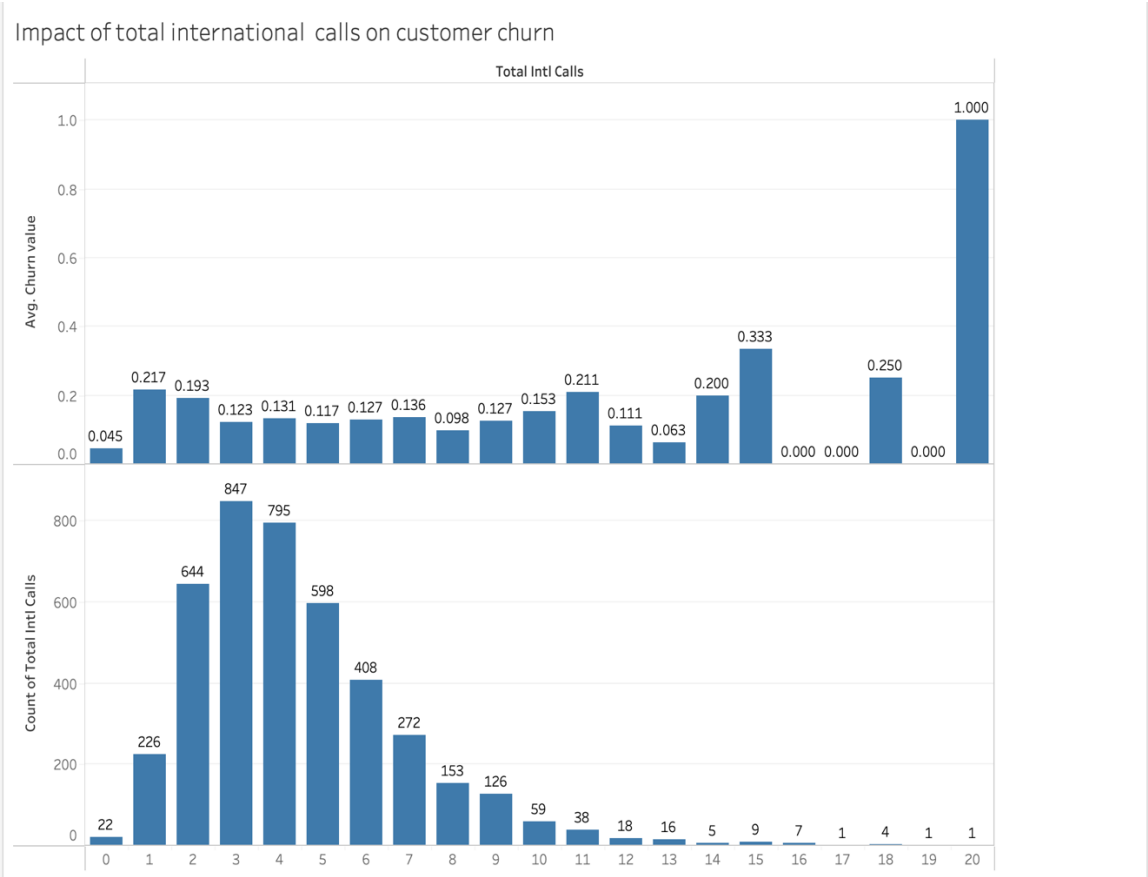
Even though the count of customers who have not opted for international plan is almost 9.7 times more as compared to those who have, the churn rate for these costumers is much lower than the customers who have an international plan. The churn rate for customers who don't have an international plan is 11.18% which is almost 4 times lower than the customers who have an international plan (42.17%)

3) Effect of total_day_minutes on churn rate



As the total minutes during the day for the call increases, the churn rate also increases. For the bins of 0 to 200 minutes the average churn rate is in the range of 6.7% - 14.29%. It even goes to be 100% for values in the range of 325-350 minutes but we don't have enough data in this range to infer from it.

4) Effect of total_intl_calls on churn rate



There is no major trend of increase or decrease visible in the churn rate based on the count of international calls made by the customer. It generally lies in the range of 4.5% to 33% for 0-15 international calls even though the count of 2-6 international calls is much larger as compared to others. We do observe a 100% churn rate for 20 international calls, but we only have one such data point which is not enough to infer any pattern. Also, we get a churn rate of 0% for 16, 17 and 19 international calls.

Q5. For each of the top 4 predictors, formulate actionable recommendations based on the observed effects. If the observed effect cannot be reasonably made actionable, please state so.

Features	Observed Effects	Actionable Recommendations
Number of customer service calls	Churn rate is low for cases that have count of customer service calls within the range of 0-3 (almost 11.29%). But this increases from that to 67.8% as the number of calls increases from 3 to 6. The churn rate again decreases to around 50% if the number of calls made is 7 or 8.	**Customer issues should be given priority and the customer service should be able to provide quick and good response to the issues.** If a customer must contact the customer service multiple times for any issue it would increase the churn rate which we have observed also.
International plan	Even though the count of customers who have not opted for international plan is almost 9.7 times more as compared to those who have, the churn rate for these costumers is much lower than the customers who have an international plan. The churn rate for customers who don't have an international plan is 11.18% which is almost 4 times lower than the customers who have an international plan (42.17%)	**The company should provide better international plans for its customers** as we can see that smaller count of customers with international plans have a much larger churn rate. This indicates that the customers are not finding the plans good.
Total minutes of day call	As the total minutes during the day for the call increases, the churn rate also increases. For the bins of 0 to 200 minutes the average churn rate is in the range of 6.7% - 14.29%. It even goes to be 100% for values in the range of 325-350 minutes but we don't have enough data in this range to infer from it.	**The telecommunication provider should come up with better plans for customers who have longer calls during the day.** The plans can be made focused to the call durations or the total minutes of call during the day.
Total international calls	There is no major trend of increase or decrease visible in the churn rate based on the count of international calls made by the customer. It generally lies in the range of 4.5% to 33% for 0-15 international calls even though the count of 2-6 international calls is much larger as compared to others. We do observe a 100% churn rate for 20 international calls, but we only have one such data point which is not enough to infer any pattern.	There is no specific trend observed in this therefore we have no such actionable recommendation for this.