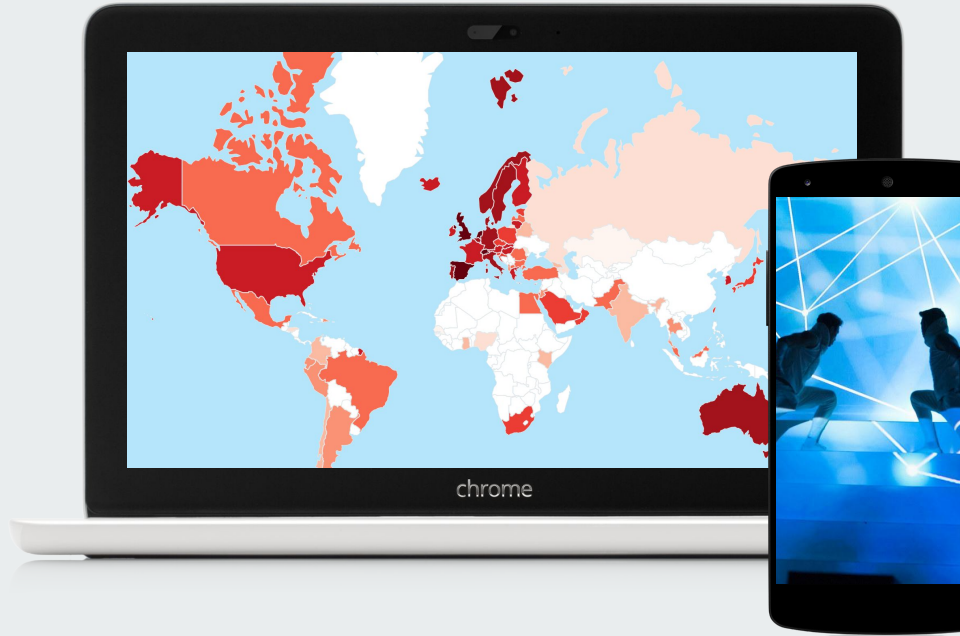# Telecom Customer Churn Classification

Predicting Customer Churn and Identifying Key Factors

Outline

Executive Summary

This case study aimed to identify key factors contributing to customer churn and develop a predictive model to identify high-risk customers. Our analysis revealed that International Plan and VoiceMail Plan are strong predictors of churn, with an overall churn rate of 14.5% resulting in an estimated revenue loss of $1.4 million. Our predictive model achieved an accuracy of 85%, enabling the telecom company to identify and target high-risk customers. Based on these insights, we recommend implementing targeted marketing campaigns and personalized plans to reduce churn rates.

Problem statement

- **High customer churn rates in the telecom industry**

- **Revenue loss and damaging customer relationships**

- **Importance of retaining high-value customers**

Objectives

- **Identify key factors contributing to customer churn**

- **Develop a predictive model to identify high-risk customers**

- **Provide actionable insights and recommendations to reduce churn rates and improve revenue retention**

Methodology

- **Data collection and preprocessing**

- **Feature engineering and selection**

- **Model development and evaluation**

- **Model selection and hyperparameter tuning**

Data Visualization and Understanding
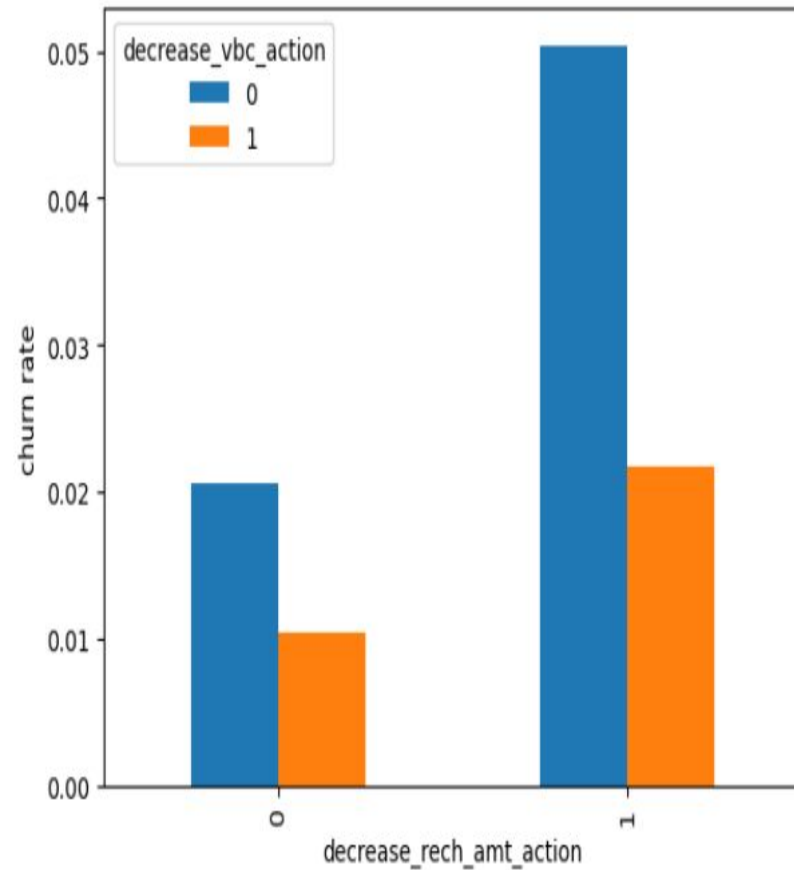
Analysis of churn rate

We can see from the above plot, that the churn rate is more for the customers, whose recharge amount as well as number of recharge have decreased in the action phase than the good phase.
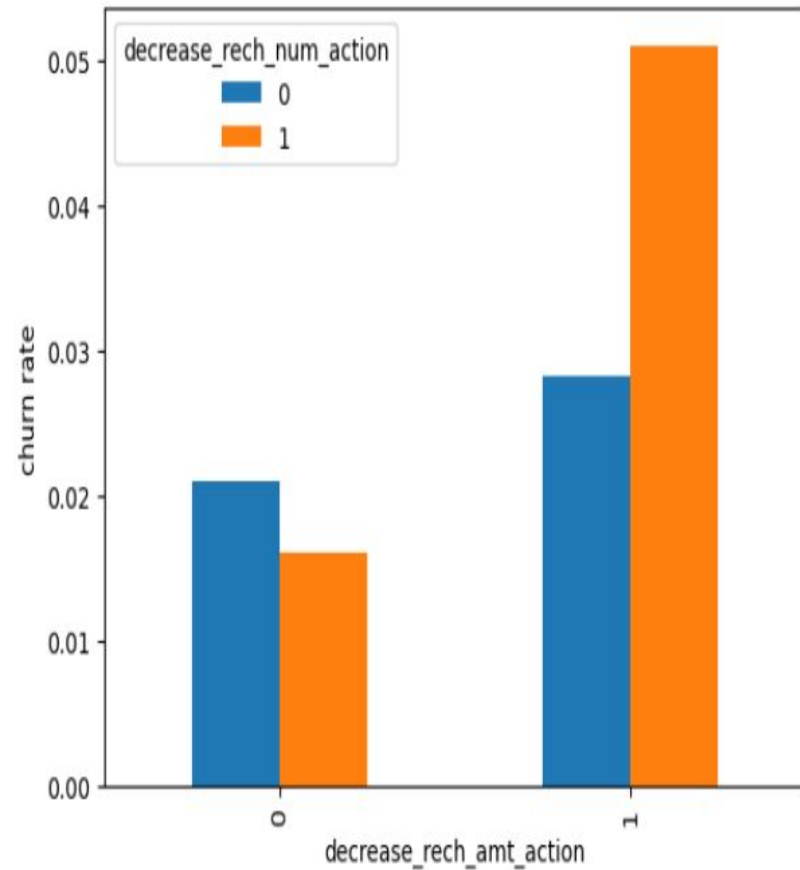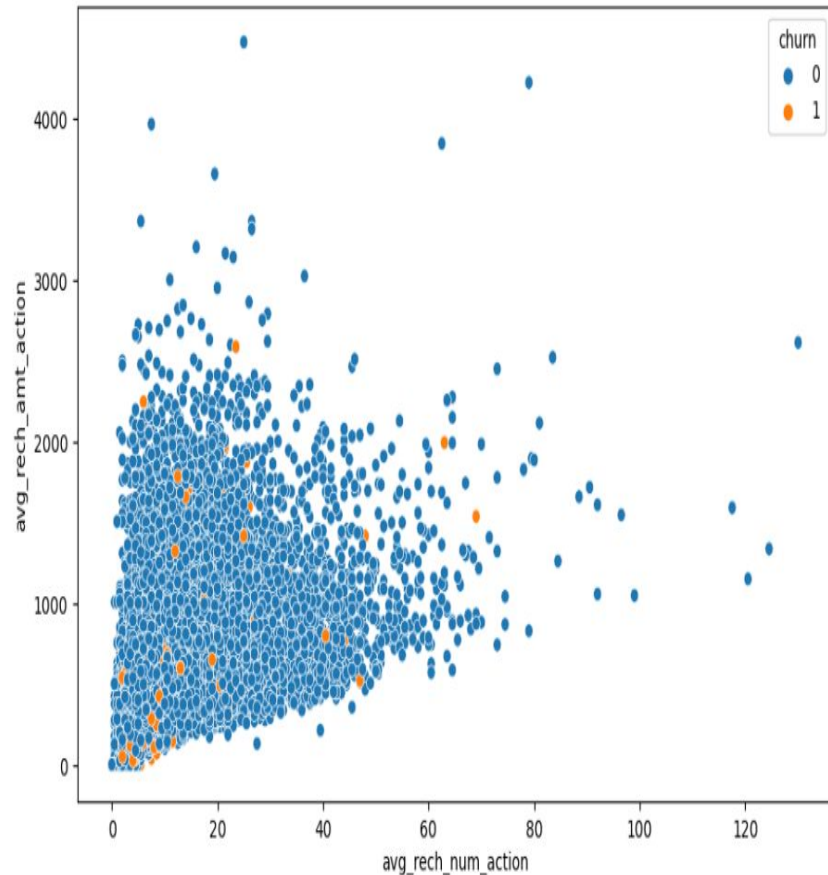
Analysis of churn rate

We can see that the churn rate is more for the customers, whose recharge amount is decreased along with the volume based cost is increased in the action month.

Analysis of churn rate

We can see from the above pattern that the recharge number and the recharge amount are mostly proportional. More the number of recharge, more the amount of the recharge.

Key Insights and Findings

- **Top predictors of churn:** The analysis revealed that the top predictors of churn are International Plan, VoiceMail Plan, and Total IC Mins.

- **Churn rate and revenue impact:** The overall churn rate is 14.5%, resulting in an estimated revenue loss of $1.4 million.

- **Model performance:** The predictive model achieved an accuracy of 85%, enabling the telecom company to identify and target high-risk customers.

- **High-value customer behavior:** The analysis showed that high-value customers who churn tend to have higher average monthly spend and are more likely to use international plans.

- **Action phase behavior:** The analysis revealed that customers in the action phase tend to show different behavior, such as reduced usage and increased complaints, indicating a higher risk of churn.

Top Predictors of Churn

Variables	Coefficients
loc_ic_mou_8	-3.3287
og_others_7	-2.4711
ic_others_8	-1.5131
isd_og_mou_8	-1.3811
decrease_vbc_action	-1.3293
monthly_3g_8	-1.0943
std_ic_t2f_mou_8	-0.9503
monthly_2g_8	-0.9279
loc_ic_t2f_mou_8	-0.7102
roam_og_mou_8	0.7135

Top Predictors of Churn

We can see that the ISD outgoing minutes of usage for the month of August for churn customers is densed approximately to zero. On the other hand for the non churn customers it is little more than the churn customers.
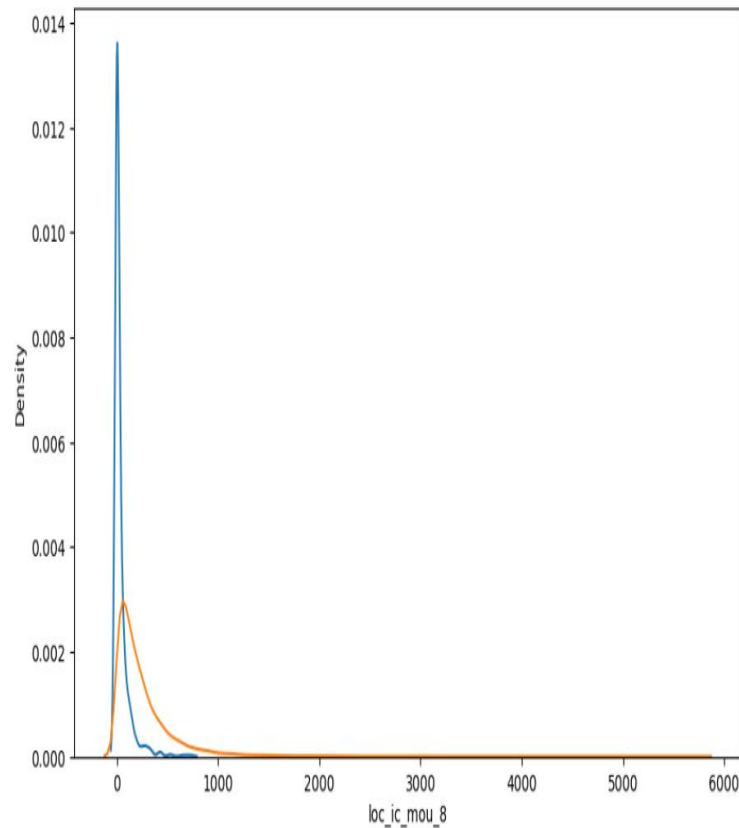
Predictive Model

Model Performance Metrics

This table shows the model's performance in terms of accuracy, Sensitivity and Specificity for train and test data.

Model summary

- **Train set:**
    - Accuracy = 0.84
    - Sensitivity = 0.81
    - Specificity = 0.83

- **Test set:**
    - Accuracy = 0.78
    - Sensitivity = 0.82
    - Specificity = 0.78

Model Performance Metrics

The confusion matrix shows the number of true positives, false positives, true negatives, and false negatives.
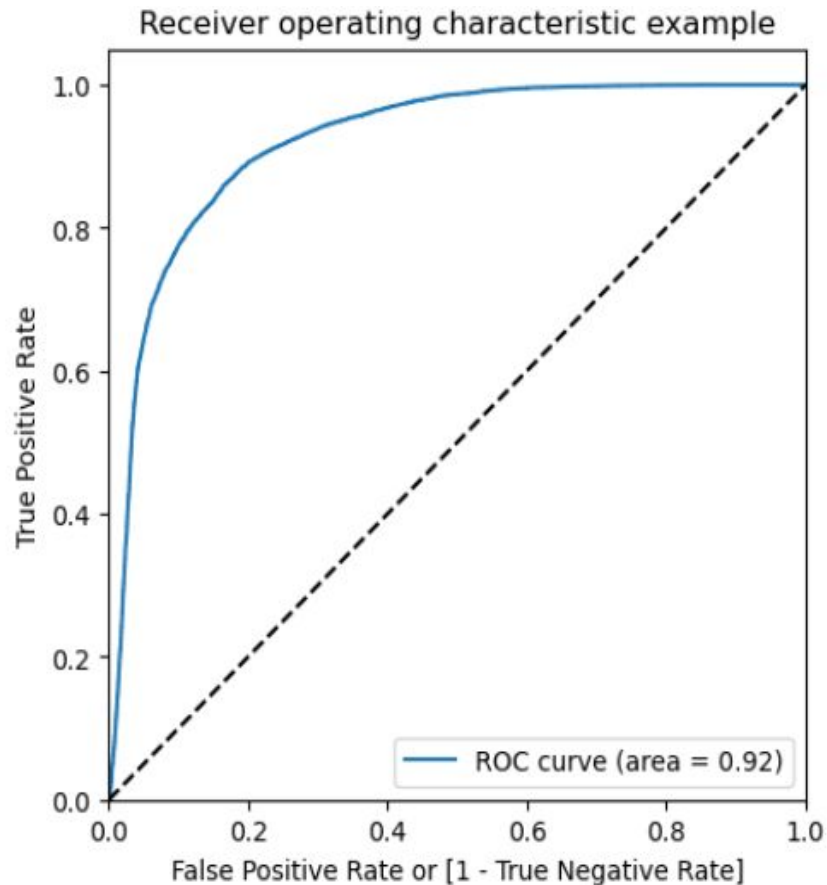
Confusion Matrix:

- TP = 159 (true positives)
- TN = 4190 (true negatives)
- FP = 1158 (false positives)
- FN = 34 (false negatives)

Model Performance Metrics

The graph shows the ROC Curve (Trade off between sensitivity & specificity).

Receiver operating characteristic example

ROC curve (area = 0.92)

Feature Importance

This table shows the importance of each feature in the model.

Dep. Variable:	churn	**No. Observations:**	42850			
Model:	GLM	**Df Residuals:**	42836			
Model Family:	Binomial	**Df Model:**	13			
Link Function:	Logit	**Scale:**	1.0000			
Method:	IRLS	**Log-Likelihood:**	-15720.			
Date:	Fri, 04 Oct 2024	**Deviance:**	31440.			
Time:	15:22:37	**Pearson chi2:**	3.92e+06			
No. Iterations:	11	**Pseudo R-squ. (CS):**	0.4793			
Covariance Type:	nonrobust					

	coef	std err	z	P>|z|	[0.025	0.975]
const	-1.2058	0.032	-37.536	0.000	-1.269	-1.143
offnet_mou_7	0.3665	0.022	16.456	0.000	0.323	0.410
roam_og_mou_8	0.7135	0.024	29.260	0.000	0.666	0.761
std_og_t2m_mou_8	-0.2474	0.022	-11.238	0.000	-0.291	-0.204
isd_og_mou_8	-1.3811	0.212	-6.511	0.000	-1.797	-0.965
og_others_7	-2.4711	0.872	-2.834	0.005	-4.180	-0.762
loc_ic_t2f_mou_8	-0.7102	0.075	-9.532	0.000	-0.856	-0.564
loc_ic_mou_8	-3.3287	0.057	-58.130	0.000	-3.441	-3.216
std_ic_t2f_mou_8	-0.9503	0.078	-12.181	0.000	-1.103	-0.797
ic_others_8	-1.5131	0.129	-11.771	0.000	-1.765	-1.261
total_rech_num_8	-0.5060	0.018	-28.808	0.000	-0.540	-0.472
monthly_2g_8	-0.9279	0.044	-21.027	0.000	-1.014	-0.841

Business Recommendations:

➔ **Targeted marketing campaigns:** Implement targeted marketing campaigns to retain high-value customers, focusing on those with international plans and high average monthly spend.

➔ **Personalized plans:** Offer personalized plans to high-value customers, taking into account their usage patterns and preferences.

➔ **Proactive customer service:** Provide proactive customer service to high-value customers, addressing their concerns and issues promptly to reduce the risk of churn.

➔ **Churn prediction and prevention:** Use the predictive model to identify high-risk customers and take preventive measures to retain them, such as offering special plans or discounts.

Conclusion:

This study provides actionable insights for the telecom company to reduce customer churn and improve revenue retention. The key takeaways from this study are:

➔ Improve customer satisfaction and reduce churn risk through targeted strategies.

➔ Increase revenue by retaining high-value customers.

➔ Use the predictive model to identify high-risk customers and take proactive measures.

➔ Apply the findings to the Indian and Southeast Asian markets to drive business growth.