Telecom Churn Analysis and Recommendations

Predicting churn and ensuring customer retention

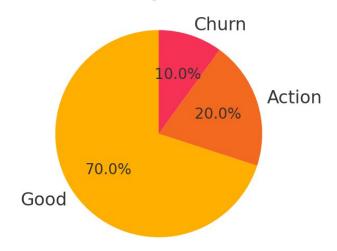
From: Manish Singh/Sanjay Poi/Syed Mohammad Raza

Summary of Findings

- 1. Telecom industry faces a churn rate of 15-25% annually.
- 2. Retaining customers is 5-10 times more cost-effective than acquiring new ones.
- 3. Churn is defined as zero usage over a specific period.
- 4. Machine learning models predict churn with accuracy (85.63%).
- 5. Customer lifecycle: Good, Action, and Churn phases.

<u>Customer Lifecycle Distribution</u>

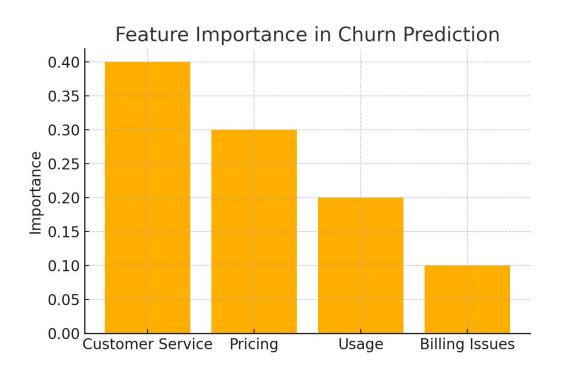
Customer Lifecycle Distribution



<u>Recommendations</u>

- 1. Focus on customers in the Action Phase for proactive retention.
- 2. Create personalized offers to match competitor pricing and enhance service.
- 3. Invest in infrastructure to improve service quality.
- 4. Use predictive analytics to guide retention strategies.

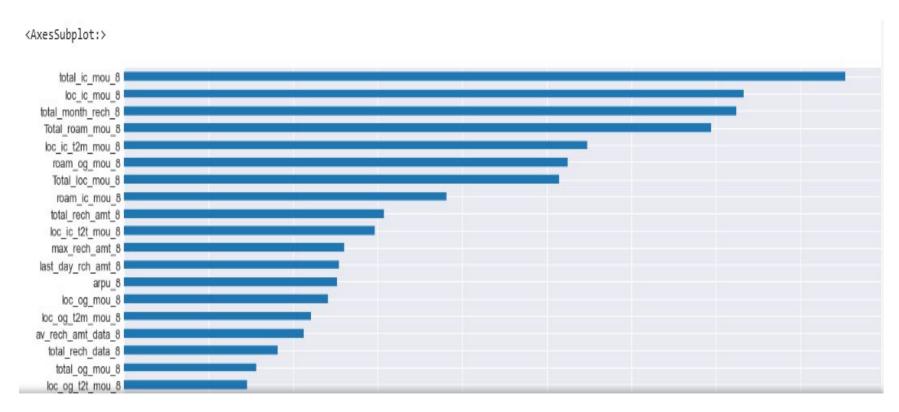
Feature Importance



Final Choice Model

Roc_auc_score 82.11% 81.21% 81.21% Specificity 77.75% 78.02% 79.54% 25.04% Precision 79.54% 76.57% 76.77% 78.03% 75% 75% 75.03% 75%	Model/Metrics	Train	Test
Sensitivity/Recall 86.48% 84.40% Specificity 77.75% 78.02% precision 79.54% 25.04% DecisionTree (cut-off = 0.4) Roc_auc_score 82.41% 76.57% Sensitivity/Recall 89.79% 78.13% Specificity 75.03% 75% precision 78.24% 21.38% Random Forest (cut-off = 0.45) Roc_auc_score 85.60% 96.53% Sensitivity/Recall 88.70% 77.57% Specificity 82.50% 81.73% precision 83.52% 26.97% GBC Roc_auc_score 96.11% 80.84% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) Roc_auc_score 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% Specificity 94.49% 85.38% SyM (linear C = 1000) Roc_auc_score 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 79.91% 78.40%	Logistic Regression (cut-off = 0.45)		
Specificity 77.75% 78.02% precision 79.54% 25.04% DecisionTree (cut-off = 0.4) 82.41% 76.57% Roc_auc_score 82.41% 76.57% Sensitivity/Recall 89.79% 78.13% Specificity 75.03% 75% precision 78.24% 21.38% Random Forest (cut-off = 0.45) 85.60% 96.53% Roc_auc_score 85.60% 96.53% Sensitivity/Recall 88.70% 77.57% Specificity 82.50% 81.73% grecision 83.52% 26.97% GBC 80.24% 80.84% Specificity 92.21% 81.81% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) 80.76% 80.76% Specificity 94.49% 85.38% Specificity 94.49% 85.38% SVM (linear C = 1000) 81.33% 82.62% Sensitivity/Recall 79.91%	Roc_auc_score	82.11%	81.21%
precision 79.54% 25.04% DecisionTree (cut-off = 0.4) Roc_auc_score 82.41% 76.57% Sensitivity/Recall 89.79% 78.13% Specificity 75.03% 75% precision 78.24% 21.38% Random Forest (cut-off = 0.45) Roc_auc_score 85.60% 96.53% Sensitivity/Recall 88.70% 77.57% Specificity 82.50% 81.73% grecision 83.52% 26.97% GBC Roc_auc_score 96.11% 80.84% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) Roc_auc_score 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) Roc_auc_score 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 79.91% 78.40%	Sensitivity/Recall	86.48%	84.40%
DecisionTree (cut-off = 0.4) Roc_auc_score	Specificity	77.75%	78.02%
Roc_auc_score 82.41% 76.57% Sensitivity/Recall 89.79% 78.13% Specificity 75.03% 75% precision 78.24% 21.38% Random Forest (cut-off = 0.45) 85.60% 96.53% Roc_auc_score 85.60% 96.53% Sensitivity/Recall 83.70% 77.57% Specificity 82.50% 81.73% precision 83.52% 26.97% GBC 96.11% 80.84% Sensitivity/Recall 100.00% 79.87% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) 80.76% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) 80.24% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	precision	79.54%	25.04%
Sensitivity/Recall 89.79% 78.13% Specificity 75.03% 75% precision 78.24% 21.38% Random Forest (cut-off = 0.45) 85.60% 96.53% Roc_auc_score 85.60% 96.53% Sensitivity/Recall 82.50% 81.73% Specificity 82.50% 81.73% GBC 96.11% 80.84% Sensitivity/Recall 100.00% 79.87% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	DecisionTree (cut-off = 0.4)		
Specificity 75.03% 75% precision 78.24% 21.38% Random Forest (cut-off = 0.45) 85.60% 96.53% Roc_auc_score 85.60% 96.53% Sensitivity/Recall 82.50% 81.73% Specificity 82.50% 81.73% GBC 96.11% 80.84% Sensitivity/Recall 100.00% 79.87% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) 80.76% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	Roc_auc_score	82.41%	76.57%
precision 78.24% 21.38% Random Forest (cut-off = 0.45) Roc_auc_score 85.60% 96.53% Sensitivity/Recall 88.70% 77.57% Specificity 82.50% 81.73% precision 83.52% 26.97% GBC Roc_auc_score 96.11% 80.84% Sensitivity/Recall 100.00% 79.87% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) Roc_auc_score 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) Roc_auc_score 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 94.49% 85.85%	Sensitivity/Recall	89.79%	78.13%
Random Forest (cut-off = 0.45) Roc_auc_score	Specificity	75.03%	75%
Roc_auc_score 85.60% 96.53% Sensitivity/Recall 88.70% 77.57% Specificity 82.50% 81.73% precision 83.52% 26.97% GBC GBC Roc_auc_score 96.11% 80.84% Sensitivity/Recall 100.00% 79.87% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) 80.76% Roc_auc_score 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% SVM (linear C = 1000) 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	precision	78.24%	21.38%
Sensitivity/Recall 88.70% 77.57% Specificity 82.50% 81.73% precision 83.52% 26.97% GBC 96.11% 80.84% Sensitivity/Recall 100.00% 79.87% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) 80.24% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	Random Forest (cut-off = 0.45)		
Specificity 82.50% 81.73% precision 83.52% 26.97% GBC GBC Roc_auc_score 96.11% 80.84% Sensitivity/Recall 100.00% 79.87% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) 80.76% Roc_auc_score 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% SVM (linear C = 1000) 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	Roc_auc_score	85.60%	96.53%
precision 83.52% 26.97% GBC Roc_auc_score 96.11% 80.84% Sensitivity/Recall 100.00% 79.87% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) Roc_auc_score 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) Roc_auc_score 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	Sensitivity/Recall	88.70%	77.57%
GBC Roc_auc_score 96.11% 80.84% Sensitivity/Recall 100.00% 79.87% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) Roc_auc_score 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) Roc_auc_score 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	Specificity	82.50%	81.73%
Roc_auc_score 96.11% 80.84% Sensitivity/Recall 100.00% 79.87% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) 80.76% Roc_auc_score 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	precision	83.52%	26.97%
Sensitivity/Recall 100.00% 79.87% Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) 80.76% Roc_auc_score 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	GBC		
Specificity 92.21% 81.81% precision 92.78% 28.52% XGB (cut-off = 0.2) 80.76% Roc_auc_score 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	Roc_auc_score	96.11%	80.84%
precision 92.78% 28.52% XGB (cut-off = 0.2) Roc_auc_score 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) Roc_auc_score 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	Sensitivity/Recall	100.00%	79.87%
XGB (cut-off = 0.2) Roc_auc_score 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) Roc_auc_score 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	Specificity	92.21%	81.81%
Roc_auc_score 97.24% 80.76% Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	precision	92.78%	28.52%
Sensitivity/Recall 99.99% 76.13% Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	XGB (cut-off = 0.2)		
Specificity 94.49% 85.38% precision 94.78% 32.13% SVM (linear C = 1000) 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	Roc_auc_score	97.24%	80.76%
precision 94.78% 32.13% SVM (linear C = 1000) Roc_auc_score 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	Sensitivity/Recall	99.99%	76.13%
SVM (linear C = 1000) Roc_auc_score	Specificity	94.49%	85.38%
Roc_auc_score 81.33% 82.62% Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	precision	94.78%	32.13%
Sensitivity/Recall 79.91% 78.40% Specificity 82.75% 86.85%	SVM (linear C = 1000)		
Specificity 82.75% 86.85%	Roc_auc_score	81.33%	82.62%
	Sensitivity/Recall	79.91%	78.40%
precision 82.25% 35.14%	Specificity	82.75%	86.85%
	precision	82.25%	35.14%

Some of the top main predictors of churn are the monthly KPI features for the action phase (3rd month August).



Top 25 Features Affecting High-Value Customer Churn Action Month (Month 8) Insights:

Top Features:

- 1. 1. Total_ic_mou_8: Total incoming minutes of usage.
- **2.** Loc_ic_mou_8: Local incoming minutes of usage.
- **3.** Total_month_rech_8: Total recharge amount.
- **4. Total_roam_mou_8**: Total roaming minutes (incoming + outgoing).
- **5.** Loc_ic_t2m_mou_8: Local incoming minutes to other operators.
- **6.** Roam_og_mou_8: Outgoing roaming calls.
- 7. Total_loc_mou_8: Total local minutes of usage.
- **8.** Roam_ic_mou_8: Incoming roaming calls.
- **9.** Total_rech_amt_8: Total recharge amount.
- **10.** Loc_ic_t2t_mou_8: Local incoming calls from same operator.
- **11.** Max_rech_amt_8: Maximum recharge amount.
- **12.** Last_day_rch_amt_8: Most recent recharge amount.
- **13. Arpu_8**: Average revenue per user.
- **14.** Loc_og_mou_8: Local outgoing calls minutes.
- **15.** Loc_og_t2n_mou_8: Local outgoing to other operators.
- **16.** Av_rech_amt_data_8: Average data recharge amount.
- 17. Total_rech_data_8: Total data recharge (MB).
- **18.** Total_og_t2t_mou_8: Outgoing calls to same operator.
- **19.** Total_rech_num_8: Total number of recharges.
- 20. Total_rech_amt_data_8: Total data recharge amount.
- 21. Max_rech_data_8: Maximum data recharge (MB).
- 22. Avg_rech_amt_8: Average recharge amount.
- 23. Fb_user_8: Usage of Facebook-like services.
- 24. Vol_data_mb_8: Data volume consumed (MB).
- **25.** Count_rech_2g_8: Number of 2G recharges.
- -> Additional Insight: Loc_og_to_ic_mou_8 (Outgoing to Incoming ratio).

Strategies to manage customer churn:

- **1.**It is a fact that it costs 5-10 times more to acquire a new customer than to retain an existing one, customer retention has now become even more important than customer acquisition.
- 2. For many incumbent operators, retaining high profitable customers is the number one business goal.

Monitoring Drop in Usage:

- 1. Customer churn seems to be well predicted by drop in usage.
- 2. Aside from using the Machine Learning model for predicting churn, the telecom company should pay close attention to drop in MoU, ARPU and data usage (2g and 3g) month over month. If feasible, the company should track these numbers week over week. Since billing cycles are typically monthly, a drop in usage numbers will give the company time to react when tracked at weekly level.
- 3. Contact these customers proactively to find out what's affecting their experience. Perhaps, offer them coupons or other incentives to continue to use the services, while the company fixes the issues reported.
- 4. Marketing team must come up with campaigns which targets these high-value tobe churner.

Churners initially had higher outgoing usage, but it dropped over time, possibly due to dissatisfaction with tariffs, call quality, or network coverage—an area for further investigation by the provider.

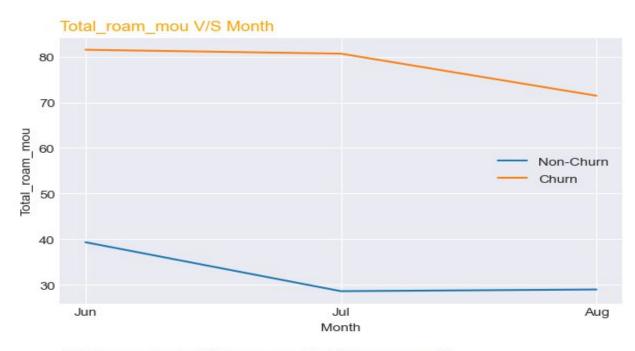


Strategy Suggestions:

- 1.Investigate and revise outgoing tariffs to stay competitive.
- 2.Launch campaigns targeting high outgoing users, e.g., discounted rates during specific hours or free minutes for X MOU usage.
- 3.Offer free monthly outgoing minutes based on past roaming usage.

Strategy Suggestions:

- 1. Investigate and optimize roaming tariffs and service quality.
- 2. Ensure quality guarantees with roaming partners.
- 3. Launch targeted campaigns, e.g., discounted roaming rates during specific hours or free monthly roaming minutes based on past usage.



Total_roam_mou_6 Total_roam_mou_7 Total_roam_mou_8

churn			
0	39.360033	28.643301	29.016734
1	81.504156	80.651973	71.443623

Business Implications

- 1. Improved customer retention increases revenue.
- 2. Proactive retention strengthens competitive positioning.
- 3. Focused spending on high-value customers optimizes costs.
- 4. Predictive analytics ensures long-term strategy effectiveness.