Telecom Churn:
Case Study

Presented by: **Shayak Majumder | Satya Ranjan | Syed MD Umar Zayed**

Business Problem

- **Industry Context:** The telecom industry faces an **annual churn rate of 15-25%** due to multiple service provider options.
- **Business Impact:** Acquiring a new customer costs **5-10 times more** than retaining an existing one.
- **Key Challenge:** Identifying and retaining **high-value customers** who contribute **most revenue.**
- **Project Objective:** Analyze customer data, build predictive models, and identify key churn indicators to help telecom firms take proactive action.
- **Market Focus:** This study is based on the **Indian and Southeast Asian markets**, where prepaid customers dominate.
- **Churn Definition:**
  - **Usage-based churn**: Customers who have **zero usage (calls, data, SMS, etc.) over a period.**
  - **High-Value Churn:** Focusing on the customers who generate the majority of revenue.

Dataset Overview

- **Source:** Telecom customer dataset (99,999 records, 226 features).
- **Data Span:** Customer data collected over four consecutive months - **June, July, August, and September** (encoded as months 6, 7, 8, and 9).
- **Business Objective:** Predict churn in the last (9th) month using data from the first three months.
- **Customer Lifecycle Phases:**
  - **Good Phase (Months 6 & 7):** Regular customer behavior.
  - **Action Phase (Month 8):** Changes in behavior indicate possible churn.
  - **Churn Phase (Month 9):** Customer either churns or stays.
- **Key Features:** Recharge history, usage patterns, roaming behavior, data consumption.
- **Target Variable:** Churn (Yes/No).

Data Dictionary

(For Reference)

- **MOBILE_NUMBER:** Customer phone number
- **CIRCLE_ID:** Telecom circle area to which the customer belongs
- **LOC:** Local calls - within the same telecom circle
- **STD:** STD calls - outside the calling circle
- **IC:** Incoming calls, **OG:** Outgoing calls
- **T2T:** Operator T to T (within same operator)
- **T2M:** Operator T to other operator mobile
- **T2O:** Operator T to other operator fixed line
- **T2F:** Operator T to fixed lines of T
- **T2C:** Operator T to its own call center
- **ARPU:** Average revenue per user
- **MOU:** Minutes of usage - voice calls
- **AON:** Age on network - number of days on operator T's network
- **ONNET:** All calls within the same operator network
- **OFFNET:** All calls outside operator T's network
- **ROAM:** Indicates customer is in a roaming zone
- **SPL:** Special calls, **ISD:** ISD calls
- **RECH:** Recharge, **NUM:** Number, **AMT:** Amount in local currency
- **MAX:** Maximum, **DATA:** Mobile internet
- **VOL:** Mobile internet usage volume (in MB)
- **2G:** 2G network, **3G:** 3G network
- **AV:** Average, **PCK:** Prepaid service schemes
- **NIGHT:** Night-hour schemes, **MONTHLY:** Monthly service schemes
- **SACHET:** Short-term service schemes
- **FB_USER:** Social media data pack user
- **VBC:** Volume-based cost (pay-per-use data charges)
- ****.6, *.7, .8, .9:* KPI for June, July, August, and September respectively

Process/Flow

Exploratory Data Analysis (EDA)

Data Preprocessing &
Model Building

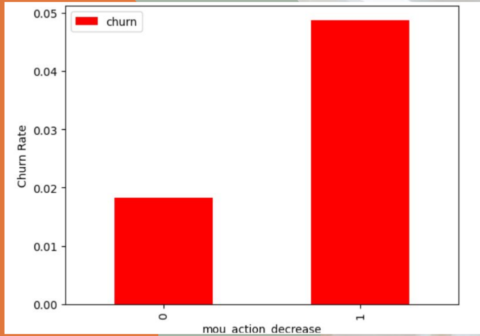
Model Selection

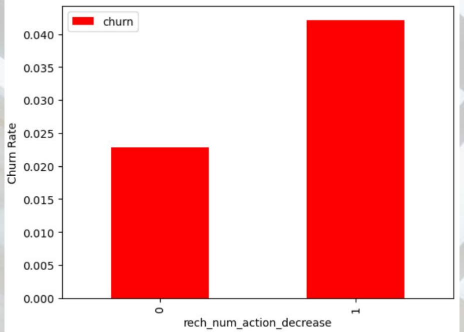
Model Performance &
Metrics

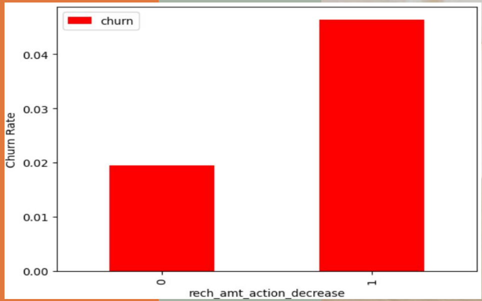
Deriving Business Insights
& Recommendations

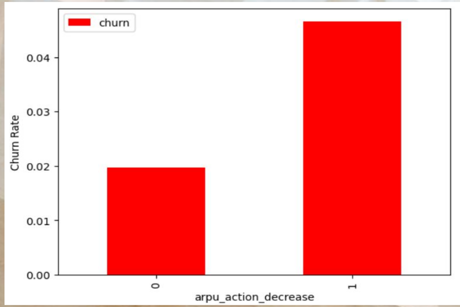
Top EDA Takeaways

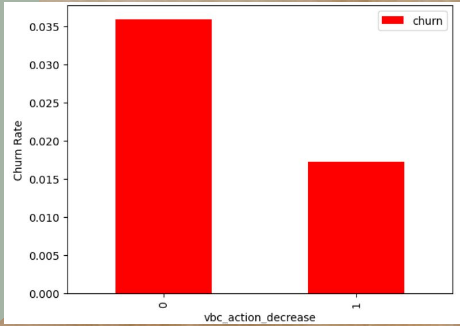
- **Churn Rate** is **higher** for customers **who registered a decrease in minutes of usage (mou)** during the **'action' phase.** *(fig. 1)*
- **Churn Rate** is **higher** for customers **who registered a decrease in number of recharges** during the **'action' phase.** *(fig. 2)*
- **Churn Rate** is **higher** for customers **who registered a decrease in recharge amount** during the **'action' phase.** *(fig. 3)*
- **Churn Rate is higher** for customers **who registered a decrease in arpu** during the **'action' phase.** *(fig. 4)*
- **Churn Rate** is **much lesser** for customers **who have actually seen an increase in volume based cost.** *(fig. 5)*











Data Preprocessing & Model Building

- Conducted **80-20 Train-Test Split**.
- Since **rate of churn was low**, around 5-10%, we **used SMOTE technique to deal with data imbalance.**
- **Standardized** numerical features **using StandardScaler.**
- To **extract the most important features to reduce dimensionality**, we created a model with **Principal Component Analysis (PCA).**

Logistic Regression

Using PCA

- Used **confusion matrix** in order to get the **True Positives and not miss a False Negative (churner).**
- **Focussed on 'Recall' more than Accuracy**, as it **helped us understand how our model actually identifies churners.**
- ***Performance:***

Train set:

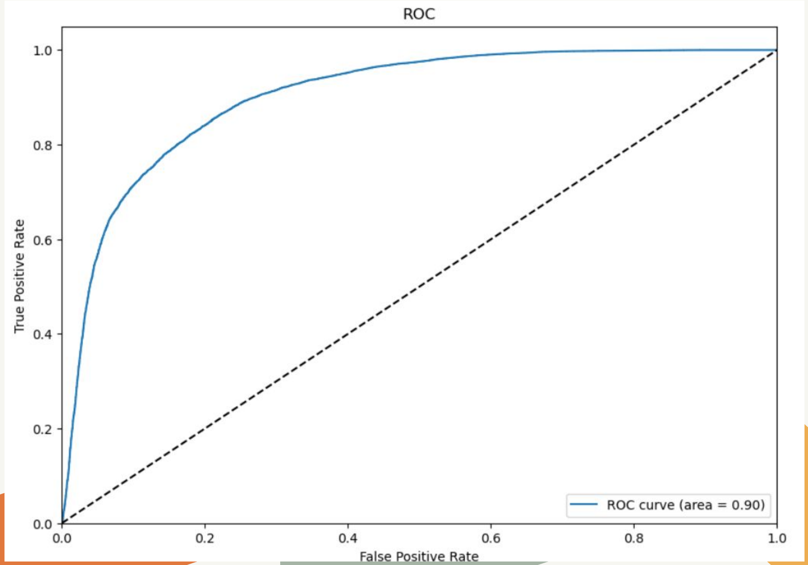
1. **Accuracy:** 86.01%
2. **Sensitivity:** 89.09%
3. **Specificity** 82.94%

Test set:

1. **Accuracy:** 82.84%
2. **Sensitivity:** 81.87%
3. **Specificity** 82.87%

Logistic Regression

Without PCA



- Used **Coarse Tuning** to deal with unwanted features.
- **First used RFE** and then **manually removed features based on VIFs.**
- Created a **total of 5 models.**

Train set:

1. **Accuracy:** 81.99%
2. **Sensitivity:** 87.47%
3. **Specificity** 76.52%

Test set:

1. **Accuracy:** 75.87%
2. **Sensitivity:** 80.83%
3. **Specificity** 75.69%

Model Selection

To recall, our model with PCA generated these results:

Train set:

1. **Accuracy:** 86.01%
2. **Sensitivity:** 89.09%
3. **Specificity** 82.94%

Test set:

1. **Accuracy:** 82.84%
2. **Sensitivity:** 81.87%
3. **Specificity** 82.87%

Conclusion on model choice:

As we can see, based on **performance on test set**, the **sensitivity of the PCA model is slightly better** than **model without PCA** (81.87% vs. 80.83%).

Since **sensitivity (recall) is our priority**, we should **choose Logistic Regression with PCA**, as it **captures more churn cases both in training and testing**.

Top Features

These are the top features:

1. **Number of recharges in August** (*total_rech_num_8*)
2. **Number of recharges in July** (*total_rech_num_7*)
3. **Difference in recharge count over months** (*difference_rech_num*)
4. **Total recharge amount in June** (*total_rech_amt_6*)
5. **Total recharge amount in July** (*total_rech_amt_7*)
6. **Local incoming call minutes in August** (*loc_ic_mou_8*)
7. **Local incoming calls to fixed lines in August** (*loc_ic_t2f_mou_8*)
8. **Monthly 3G usage in August** (*monthly_3G_8*)
9. **Decrease in volume-based data charges** (*vbc_action_decrease*)
10. **Incoming calls from other operators in August** (*ic_others_8*)
11. **Outgoing calls to other operators in July** (*og_others_7*)

Top Recommendations

Personalized Retention Offers: Customers reducing recharges may be churning. Offer them discounts or bonus data for continued engagement.

Recharge Consistency Plans: Encourage customers to maintain regular recharges by offering loyalty points or free add-ons.

Early Warning System: Flag customers whose recharge numbers are dropping sharply for proactive retention efforts.

High-Value Customer Retention: If high-spending customers are reducing recharge amounts, offer exclusive discounts or premium customer service.

Price Sensitivity Analysis: Test new pricing structures or promotions to prevent further decline in spending.

Call Usage Trends: Customers with decreasing incoming calls might be shifting to other networks. Offer better call quality or free incoming call packs.

Enterprise & Business Solutions: If business users are reducing calls to fixed lines, promote VoIP solutions or bundled corporate plans.

Data Upsell Campaigns: Encourage customers with declining 3G usage to switch to 4G with discounts or better plans.

Push Data Packs: If customers are moving away from pay-per-use data, upsell unlimited data packs.

Competitor Influence Check: If more calls are from other networks, competitors might be attracting customers. Counter this with competitive offers.

Inter-Operator Retention Strategy: Offer better call rates for cross-network calls to reduce churn risk.

Top 5 Business Actions

1. **Develop a churn prevention model** based on declining recharge frequency, recharge amount, and data usage patterns.
2. **Launch proactive retention offers** for customers showing early signs of churn.
3. **Analyze competitor influence based on incoming/outgoing call trends** to other operators.
4. **Segment high-value customers (based on recharge amount)** and provide personalized incentives.
5. **Encourage prepaid customers to switch to postpaid plans** for better customer stickiness.

Thank You