Title Slide

Customer Churn Analysis for Telecom Industry
Predicting and Understanding Customer Churn
Kaldhen Sangay Losel Lepcha

Project Overview

- Objective : Predict Customer Churn and identify key drivers of churn to improve retention.
- Dataset : Information on customer usage for four months (June to September).
- Business Value : Reduce churn by targeting high risk customers and understanding key churn drivers.

Data Preparation and Feature Engineering

- Handling Missing Values : By imputing median for numeric values.
- High Value Customer Identification : Filtered customers based on recharge amounts above 70th percentile.
- Churn Definition : Defined churn based on behaviour in the last month.

Modelling Approach

- Baseline Model : Logistic Regression as the initial model.
- Class Imbalance Handling : Used class_weight = ‘balanced’ to handle class imbalance.
- Tuned Model : Performed hyperparameter tuning using GridSearchCV for better performance.

Key Model Performance Metrics

- Metrics Compared : Precision , Recall , F1 Score , Accuracy.
- Baseline vs Tuned Performance:
- Precision (Churn): Increased from 0.3254 to 0.3261
- Recall (Churn): Maintained at 0.8131
- F1 Score (Churn): Increased from 0.4648 to 0.4655
- Accuracy: Increased from 0.83

Feature Importance Analysis

- Top Feature Influencing Churn:
- total_ic_mou_8(Total Incoming Minutes of Usage in August): More incoming calls are linked to reduced churn risk.
- sep_vbc_3g(Data Volume Charged Beyond Quota): Customers exceeding data limits are more likely to churn.
- total_og_mou_8(Total Outgoing Minutes of Usage in August): Higher outgoing Calls correlate with reduced churn risk.

Business Insights and Recommendations

- Key Insights:
- Engagement coincides with Loyalty : High incoming/outgoing call activity reduces churn risk.
- Data Overages Lead to Churn : Customers exceeding data limits are at higher risk.
- Recommendations:
- Increase Customer Engagement: Offer incentives for high call activity.
- Address Data Overages : Introduce flexible data plans or provide discounts for data overages to reduce dissatisfaction.
- Early Warning System : Develop triggers for declining usage patterns to prompt retention campaigns.

Conclusion and Next Steps

- Conclusion:
- Logistic Regression Model : Slight improvements through hyperparameter tuning.
- Key Indicators Identified : High usage, data overages, and engagement were significant drivers of churn.
- Next Steps :
- Advanced Models : Explore Random Forest or XGBoost to better capture complex behaviour.
- Deployment : Deploy the model to identify churn-prone customers in real-time.
- Actionable Retention : Use identified features to create proactive retention campaigns.