Capstone Project Report

Project Title

Customer Churn Prediction in the Telecom Industry

Prepared by

Om Raju Sangle

1. Introduction

Customer churn is a significant challenge in the telecom industry, affecting profitability and growth. This project aimed to predict which high-value customers are at risk of churn based on usage data from the Indian and Southeast Asian prepaid markets.

2. Objectives

- To identify high-value customers based on their usage and recharge data.
- To predict customer churn using machine learning models.
- To provide actionable insights and recommendations to mitigate churn.

3. Methodology

3.1 Data Collection

The project utilized a dataset containing customer usage data over four months, with a focus on identifying churn in September based on previous months' data.

3.2 Data Preprocessing

- Feature Selection: Correlated features were removed to enhance model performance.

- Missing Values Handling: Missing values were imputed to ensure data integrity.

- Feature Scaling: Standardization was applied to bring all features to a comparable scale.

3.3 Model Development

A logistic regression model was selected for churn prediction. The model was trained and validated using a train-test split approach. Hyperparameter tuning was performed to optimize model performance.

4. Results

4.1 Model Performance

- Confusion Matrix:

	Predicted No Churn	Predicted Churn
Actual No Churn	18,138	70
Actual Churn	16	1,776

- Classification Report:
  - Precision for Churn: 0.96
  - Recall for Churn: 0.99
  - F1-Score for Churn: 0.98
  - Overall Accuracy: 100%

4.2 Best Model Parameters

- Penalty: L2

- C (Regularization Strength): 100

5. Findings

- The model achieved outstanding accuracy and F1-scores, indicating its effectiveness in predicting customer churn.
- The significant imbalance in the dataset was effectively managed, resulting in reliable predictions.

6. Recommendations

1. Targeted Retention Strategies: Focus on high-risk customers identified by the model to offer tailored retention strategies, such as personalized promotions or enhanced customer support.

2. Continuous Monitoring: Implement regular updates to the model with new data to maintain accuracy over time.

3. Data-Driven Decisions: Utilize insights from the model to inform marketing strategies and improve customer engagement efforts.

4. Expand Data Sources: Incorporate additional data sources, such as customer feedback and social media sentiment, to enhance predictive capabilities.

7. Conclusion

The successful implementation of a predictive model for customer churn highlights the potential for data-driven strategies in the telecom industry. By leveraging these insights, telecom companies can proactively address customer concerns and reduce churn rates, ultimately improving profitability and customer satisfaction.