

Customer Churn Prediction and Risk Segmentation – Telecom Industry

Introduction

Customer churn is a major challenge in the telecom industry. Losing customers directly affects revenue and profitability. This project aims to predict customer churn using machine learning and identify high-risk customers for retention strategies.

Objective

To build a classification model that predicts whether a customer will churn and segment customers into risk levels (Low, Medium, High) for proactive retention strategies.

Tools Used

- Python
- Pandas
- Scikit-learn
- Matplotlib
- Logistic Regression
- VS Code

Dataset Overview

- Total records: 7043
- Target variable: Churn
- Features: Contract type, Monthly charges, Internet service, Payment method, etc.

Class Imbalance Handling

The dataset contained fewer churn customers compared to non-churn customers, resulting in class imbalance. Since the primary goal of this project is to correctly identify customers who are likely to churn, recall for the churn class was prioritized. To address this imbalance and improve churn detection performance, Logistic Regression with `class_weight="balanced"` was used.

Model Used

Balanced Logistic Regression was implemented to handle class imbalance and improve churn detection performance. The model was trained using an 80–20 train-test split. Recall was prioritized over accuracy since identifying churn customers is critical for business retention strategies.

Model Performance:

- Accuracy: 77%
- Recall (Churn class): 78%
- Precision (Churn class): 57%

Key Insights

Factors Increasing Churn:

- Month-to-month contracts
- Electronic check payment method
- Fiber optic internet service
- High monthly charges
- Senior citizens

Factors Reducing Churn:

- Two-year contracts
- One-year contracts
- Tech support services
- Online security
- Automatic credit card payment

Risk Segmentation

Customers were segmented into Low, Medium, and High risk groups based on predicted churn probability. High-risk customers require immediate retention strategies such as discounts, loyalty benefits, and service improvements.

Business Recommendations

- Encourage customers to switch to long-term contracts
- Promote automatic payment methods
- Improve fiber optic service support
- Offer retention discounts to high-risk customers
- Provide free tech support trials to vulnerable customers

Conclusion

The model successfully identified high-risk customers with strong recall performance. By leveraging predictive analytics and customer segmentation, telecom companies can proactively reduce churn and improve long-term customer retention.