TELECOM CHURN ANALYSIS

BY : STUTI SINGH

CONTENTS

01

Introduction

Project Overview

Purpose of the Project

The primary goal is to analyze customer churn rates, understand contributing factors, and develop strategies to improve retention.

Importance of Study on Churn

Understanding why customers leave is crucial to reducing churn, improving customer satisfaction, and maintaining a competitive edge in the telecom industry.

Telecom Industry Background

Industry Trends

The telecom industry is rapidly evolving with advances in technology, increasing competition, and changing consumer behavior driving significant shifts.

Challenges & Opportunities

Telecom companies face challenges like high competition and regulatory issues but can capitalize on opportunities in emerging technologies and new market segments.

02

Data Collection and Preparation

Data Sources

01

Internal Data

Internal data includes information collected from within the organization, such as sales records, employee information, and internal communications.

02

External Data

External data comes from outside the organization, including market research, industry reports, and data from external partners or public sources.

03

Data Integration

Data integration involves combining internal and external data to create a comprehensive dataset for analysis and insight generation.

Data Cleaning

Handling Missing Values

This process addresses gaps in datasets, employing techniques like mean imputation, deletion, or the use of machine learning models to estimate missing values.

Outlier Detection

Identifying and managing outliers, which are data points significantly different from others, usually involves statistical methods or machine learning techniques.

Data Normalization

Data normalization adjusts values measured on different scales to a common scale, often essential for accurate data analysis and machine learning.

03

Exploratory Data Analysis

Initial Insights

Descriptive Statistics

Summarizes data using measures like mean, median, mode, standard deviation, and range to understand patterns and trends.

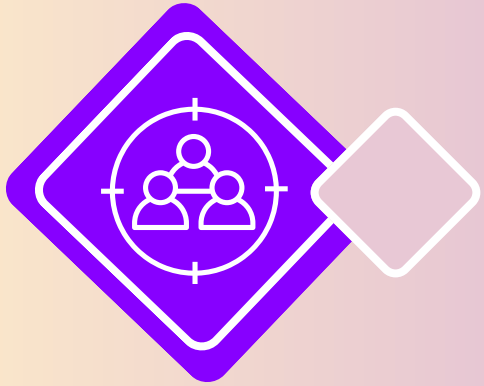
Demographic Analysis

Examines population characteristics such as age, gender, income, and education to identify market segments and trends.

Churn Rates

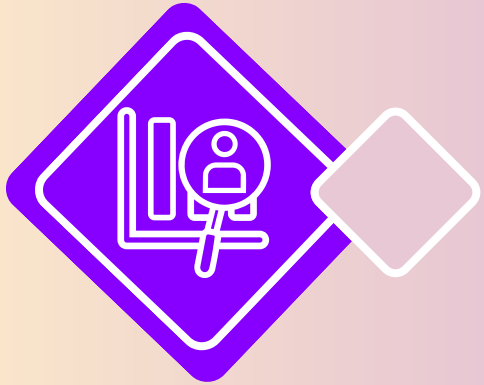
Measures the percentage of customers lost over a specific period, indicating customer retention and business sustainability.

Visualization



Customer Segmentation

Divides a customer base into distinct groups based on factors like demographics, behavior, and purchasing patterns. This helps businesses target specific segments more effectively with tailored marketing strategies and improve customer satisfaction and loyalty.



Churn Drivers

Identifies the key factors contributing to customer attrition, such as poor service, high prices, or better competitor offerings. Understanding these drivers helps businesses reduce churn by addressing the root causes and improving customer retention.

04

Predictive Modeling

Model Selection

01

Logistic Regression

Logistic Regression is a statistical method for analyzing datasets in which there are one or more independent variables that determine an outcome, used for binary classification problems.

02

Decision Tree

A Decision Tree is a flowchart- like structure that is used for decision making, where each internal node represents a test on an attribute, and each leaf node represents a class label, used for classification and regression tasks.

Model Evaluation

Accuracy Metrics

Accuracy Metrics are quantitative measures used to evaluate the performance of a model, including precision, recall, F1- score, and overall accuracy, vital for assessing classification models.

ROC Curve

The ROC Curve is a graphical representation of a classifier's performance, plotted with the True Positive Rate against the False Positive Rate, helping in understanding the trade-offs between sensitivity and specificity.

Confusion Matrix

A Confusion Matrix is a table used to describe the performance of a classification model, laying out the true positives, true negatives, false positives, and false negatives, crucial for visualizing model accuracy.

05

Results and Recommendations

Findings

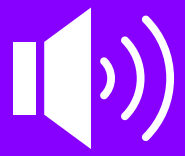
Key Predictors of Churn

Key predictors of customer churn include low engagement levels, frequent service complaints, and dissatisfaction with the value received. Identifying these can help target interventions effectively.

Customer Profiles at Risk

Customer profiles at risk of churn typically include those with low usage rates, recent negative experiences, and those who have not engaged with loyalty programs. Tailored engagement strategies are crucial.

Strategic Recommendations



Retention Strategies

To improve retention, consider personalized communication, enhanced loyalty rewards, and proactive customer service outreach. Regularly assess customer satisfaction and make adjustments as needed.

Policy Changes

Review and revise policies that commonly dissatisfy customers, such as return/exchange processes and subscription cancellations. Simplify procedures to enhance customer experience and trust.

Future Work

Future work should focus on continuously monitoring churn predictors, refining customer profiles, and developing innovative engagement strategies. Periodic reassessment and adjustment are key to long- term success.

06

Conclusion

Summary of Project

We can see that the logistic model with no PCA has good sensitivity and accuracy, which are comparable to the models with PCA. So, we can go for the more simplistic model such as logistic regression with PCA as it expliains the important predictor variables as well as the significance of each variable.

The model also hels us to identify the variables which should be act upon for making the decision of the to be churned customers. Hence, the model is more relevant in terms of explaining to the business.