Customer Churn & Revenue Analysis

Apurv panbude / 28/01/25

Table of Contents

Introduction

In today's competitive digital landscape, customer churn is a major concern for subscription-based busienesses .Companies must understand the factors affecting customer retention and how different services impact revenue.

The Objective -

- To find Analyse revenue trends and customer spending behaviour.
- Identifying factors influencing customer churn (customers leaving the service)
- Develop predictive models to estimate churn liklihood and forecast revenue.

Business and Analytical Problem

Business Problem – For subscription-based businesses, customer churn represents a critical challenge as it directly impacts revenue and growth. High churn rates typically stem from short subscription periods, service failures, or inadequate contract lengths, as identified in the analysis. Addressing these issues requires companies to not only predict high-risk customers but also design effective strategies to improve customer retention and optimize pricing models. The business must translate analytical insights into actionable solutions, such as enhancing service quality, offering personalized incentives, or reworking contract terms, to sustainably improve customer satisfaction and revenue stability

Business and Analytical Problem

Analytical Problem – The main analytical challenge lies in handling and preparing the data for meaningful insights. Issues such as missing values, skewed revenue distribution, and potential outliers complicate data preprocessing. Additionally, identifying and selecting key features (e.g., subscription age, bill average, remaining contract) to influence churn prediction and revenue forecasting requires robust feature engineering. The choice of models—Random Forest for classification and regression—may limit capturing more complex relationships, especially in datasets with intricate patterns. Moreover, relying solely on metrics like Accuracy Score and Mean Absolute Error (MAE) might not adequately address the nuances of customer behavior or revenue trends.

Data-Preprocessing

Data-Preprocessing - Data preprocessing was a crucial step to prepare the dataset for analysis. The dataset used was from kaggle's "Internet Service churn" dataset . Here preprocessing involved handling missing values, checking the dataset structure, and exploring key features such as bill average, subscribtion age, remainaing contract length, and churn status. These steps ensured the data was clean and ready for further analysis. By addressing missing values and identifying the relevant features, the basee was laid for effective explorative data analysis and predictive modelling,enabling insights for customer churn and revenue trends

Data Overview - 72272 total entries

Data columns - Total 11 columns

Explorative Data Analysis

From the Explorative Data Analysis (EDA) done we get to know that :

1. Revenue Patterns-

1. The revenue distribution shows a right-skewed pattern
2. A small number of high-value customers generate a disproportionately large portion of total revenue
3. This indicates the presence of some very valuable customers that should be prioritized for retention

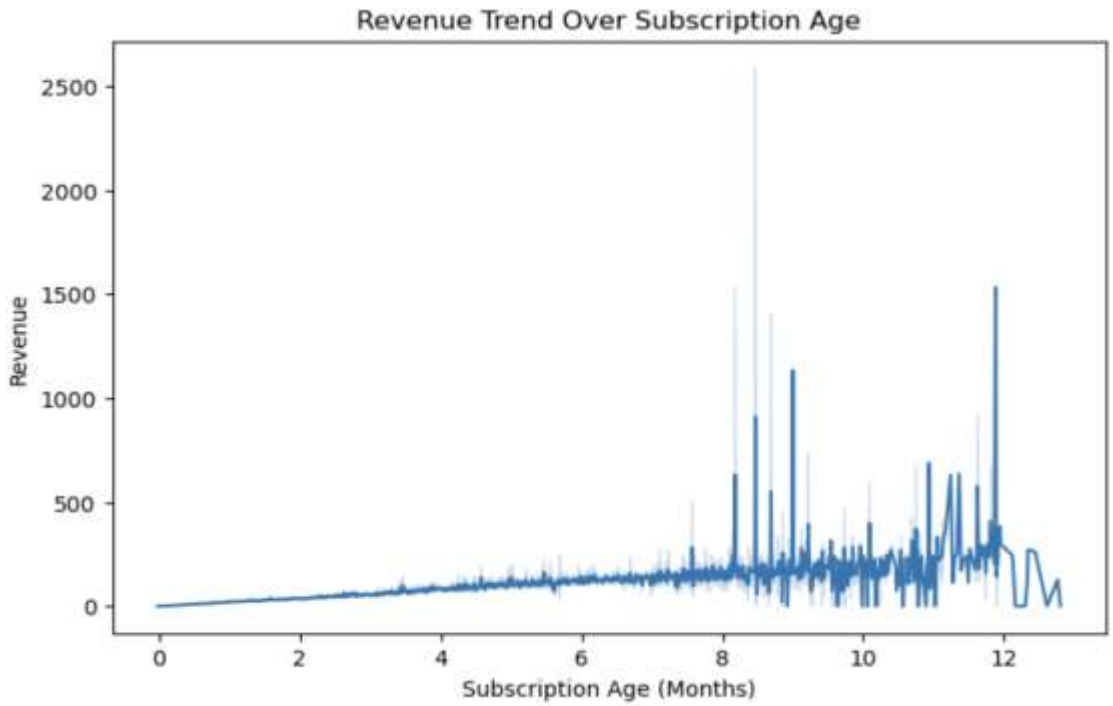
2. Churn Analysis-

1. There is a clear relationship between revenue and customer churn
2. Customers who eventually churned typically had lower average revenue
3. This suggests that revenue levels could be a useful predictor of churn risk

Explorative Data Analysis

3. Customer Lifetime Value-

1. A positive correlation exists between subscription age and revenue
2. Longer-tenured customers tend to generate more revenue
3. This highlights the importance of customer retention for business growth



Predictive model

1. **Model selection and Purpose :** A random forest classifier was implemented with a primary goal to predict customer churn. The model helps identify customers who are likely to cancle their service.

2. The model Performance which was achieved was with a accuracy of 92.8%. The high accuracy suggests that model is quite effective in predicting the customer churn.

3. **Model features and implementation –** The model uses various customer – related features including subscribtion age, Bill average, contract information, usage statistics**.**

4. **Model output -** It is quite strong in interpreting that there are high risk customers and it prioritizes retention of customers.

Recommendations

1. **Customer segmentation and retention strategy**
2. **Proactive Risk management**
3. **Revenue Optimization**
4. **Service Quality Improvements**
5. **Contract and pricing strategies**
6. **Operational improvements**
7. **Data-Driven Decision Making**
8. **Long term Strategic planning**

Bibliograpy

1. Dataset – Internet service churn csv
2. Exploratary data analysis image
3. Business analysis image

Thank you for your attention!