

BUSINESS PROBLEM

E-commerce companies sell products online to customers. Some customers stop purchasing after a few transactions (churn customers), which leads to revenue loss and reduced customer lifetime value.

Introduction

E-commerce has transformed the way customers purchase products and services. Online platforms such as Amazon, Flipkart, Myntra, and others allow customers to browse, compare, and buy products from anywhere at any time. However, one major challenge faced by e-commerce companies is customer churn — where customers stop purchasing from the platform after some time.

Customer retention is more cost-effective than acquiring new customers. Therefore, predicting customer churn and improving repeat purchase behavior is very important for business growth and profitability. Data Science and Machine Learning can help identify customers who are likely to stop shopping and enable companies to take preventive actions in advance.

1. Business Objective

- Identify customers who are likely to stop purchasing (churn).
- Increase customer retention rate.
- Improve personalized marketing strategies.
- Take preventive actions such as:
 - Sending personalized offers and discounts.
 - Providing loyalty rewards.
 - Recommending relevant products.
 - Sending reminder notifications.

Business Success Criteria

- Reduction in customer churn rate (e.g., by 15–20%).
- Increase in repeat purchase rate.
- Higher Customer Lifetime Value (CLV).
- Improved conversion rate.

2. Assess Situation

Inventory of Resources

- Historical customer purchase data.

- Website browsing behavior data.
- Cart abandonment records.
- Customer demographic information.
- Marketing campaign data.
- Data science team and marketing analysts.
- Cloud servers and database infrastructure.

Requirements

- Accurate prediction of churn customers.
- Model should be interpretable for marketing decisions.
- Real-time prediction for personalized recommendations.

Assumptions

- Past purchasing behavior reflects future behavior.
- Customer data is accurate and sufficient.
- Market demand remains relatively stable.

Constraints

- Data privacy regulations (e.g., GDPR compliance).
- Class imbalance (fewer churn customers compared to active customers).
- Limited data for new customers.

Costs:

- Data storage and processing.
- Model development and maintenance.
- Marketing campaign expenses.

Benefits:

- Increased sales revenue.
- Better customer engagement.
- Improved marketing efficiency.
- Higher brand loyalty.

3. Determine Data Science Goals

Data Science Objective (Technical View)

- Build a classification model to predict whether a customer will:
 - Churn (Yes).
 - Continue Shopping (No).

Data Science Tasks

- Data cleaning and preprocessing.

- Feature engineering (e.g., frequency of purchase, time since last purchase, average cart value).
- Handle missing values and class imbalance.
- Train machine learning models such as:
 - Logistic Regression
 - Decision Tree
 - Random Forest
 - XGBoost

Data Science Success Criteria

- High prediction accuracy.
- High Recall for churn customers (to retain risky customers).
- Acceptable Precision (to avoid unnecessary discounts).
- AUC-ROC score above defined threshold (e.g., >0.85).

4. Produce Project Plan

Project Plan Overview

Stages	Activity	Duration
1	Business Understanding	1 week
2	Data Collection and Understanding	2 weeks
3	Data Cleaning and Preparation	2 weeks
4	Model Building	3 weeks
5	Model Evaluation	1 week
6	Deployment	1 week
7	Monitoring and Maintenance	Continue

Resources Needed

- Data Scientists.
- Marketing domain experts.
- IT infrastructure team.
- Database administrators.

Tools & Techniques

- Python (Pandas, NumPy, Scikit-learn).
- SQL for data extraction.
- Power BI / Tableau for visualization.
- Machine learning classification algorithms.

Final Outcome

- A deployed Customer Churn Prediction System for E-Commerce.
- Helps company take proactive retention actions.
- Aligns data science goals with business revenue objectives.

Conclusion

Customer churn prediction in the e-commerce domain plays a critical role in improving customer retention and maximizing profitability. By using machine learning techniques, companies can identify high-risk customers in advance and take preventive measures such as personalized offers and loyalty rewards.

This project aligns business objectives with data science goals and ensures data-driven decision-making. With proper implementation and monitoring, it can significantly improve revenue, customer satisfaction, and long-term growth of the e-commerce platform.