MCA Semester – IV

Research Project – Final Report

Name	Pavan Raj K G
Project	Churn Prediction Model for E-commerce Customer Retention
Date of Submission	22-05-2024

A study on " Churn Prediction Model for E-commerce Customer Retention "

Research Project submitted to Jain Online (Deemed-to-be University)

In partial fulfillment of the requirements for the award of:

Master of Computer Application

Submitted by:

Pavan Raj K G

USN:

221VMTR02085

Under the guidance of:

Mr. Nimesh Marfatia

Mention your Guide's Name

(Faculty-JAIN Online)

Jain Online (Deemed-to-be University)

Bangalore

DECLARATION

I, *Pavan Raj K G,* hereby declare that the Research Project Report titled "Churn Prediction Model for E-commerce Customer Retention*" has been* prepared by me under the guidance of the Mr. Nimesh Marfatia. I declare that this Project work is towards the partial fulfillment of the University Regulations for the award of the degree of Master of Computer Application by Jain University, Bengaluru. I have undergone a project for a period of Eight Weeks. I further declare that this Project is based on the original study undertaken by me and has not been submitted for the award of any degree/diploma from any other University / Institution.

Place: Bangalore

Date: 22-05-2024

Pavan Raj K G
USN: 221VMTR02085

List of Tables		
Table No.	**Table Title**	**Page No.**
1	**Overview on Columns in the Dataset**	12
2	**Overview on data available**	12
3	**Comparison of Models**	17
4	**customer count on account segment**	19
5	**customer count on Payment Method**	20
6	**customer count on Device used**	20
7	**Correlation Insights**	20

List of Graphs		
Graph No.	**Graph Title**	**Page No.**
1	**Pie charts of Payment used for purchasing, Gender and Account segment**	13
2	**Pie chart of Marital Status of user and Login device used for ordering**	13
3	**Distribution of Churn**	14
4	**Bar chart of Various Factors v/s Churn**	15
5	**Corelation of columns with Churn**	21

Table of Contents

1. Abstract

This project tackles customer churn, a major hurdle for e-commerce companies, by developing a predictive model to pinpoint customers at risk of abandoning the platform. Leveraging machine learning techniques, we analyzed a rich dataset encompassing customer tenure, service satisfaction scores, preferred payment methods, and complaint history. This in-depth data analysis fueled the creation of a model that generates actionable insights to inform targeted marketing campaigns and refine customer retention strategies. The evaluation process identified the Random Forest model as the most effective, demonstrating exceptional accuracy and precision in distinguishing between churning and non-churning customers. This report details the methodology employed for data exploration, model building, and evaluation. Additionally, it presents key findings and recommendations for utilizing the model to bolster customer loyalty and drive sustainable business growth for the e-commerce company.

2. Introduction

The e-commerce industry thrives on a constant influx of new customers. However, a hidden threat lurks beneath the surface – customer churn. This phenomenon, where customers cease doing business with a company, acts like a relentless tide eroding profitability. Studies have delivered a sobering truth: acquiring new customers can be five to ten times more expensive than retaining existing ones. This highlights the critical need for e-commerce companies to prioritize customer retention strategies in today's fiercely competitive landscape.

Introduction and Background

The e-commerce landscape is no longer a static marketplace. It's a dynamic battleground where new competitors emerge constantly, and customer expectations are constantly evolving. Today's empowered e-commerce customers are no longer passive consumers. They have access to a vast array of choices, with online retailers competing fiercely on price, product selection, and convenience. Offering a competitive product or service is simply not enough to guarantee customer loyalty. Customers actively seek positive experiences and value their interactions with online retailers. A slow checkout process, a confusing return policy, or a lack of personalized recommendations can all contribute to customer dissatisfaction and ultimately, churn.

Problem Statement

Developing effective customer retention strategies hinges on understanding the root causes of churn. These causes are multifaceted, weaving a complex tapestry that can include:

- **Negative Customer Service Experiences:** Long wait times, unhelpful interactions, or a lack of resolution to customer concerns can significantly impact customer satisfaction and loyalty.
- **Friction in the Buying Journey:** A cumbersome checkout process with hidden fees, complex navigation, or limited payment options can frustrate customers and lead to cart abandonment.
- **Lack of Perceived Value:** Customers are value-conscious. If they perceive a lack of product variety, competitive pricing, or personalized recommendations, they may be more likely to explore other options.
- **Limited Interaction with the Platform:** Infrequent purchases or a lack of engagement with the platform (e.g., browsing product pages, leaving reviews) can indicate a waning interest and a higher risk of churn.

Identifying these red flags allows for early intervention. By recognizing the warning signs, e-commerce companies can take proactive steps to address customer concerns before they reach a tipping point and churn.

Objective of Study

This project aims to empower e-commerce companies with a data-driven weapon against customer churn. We will leverage the power of machine learning, a form of artificial intelligence, to transform vast amounts of customer data into actionable insights. By analysing customer data points such as:

- Purchase history (frequency, product categories, average order value)
- Demographics (age, location, income)
- Customer service interactions (number of interactions, nature of issues raised)
- Platform engagement (frequency of logins, browsing behavior)

We can identify key patterns and trends associated with customer churn. This knowledge will be used to develop a predictive model that can identify customers at risk of leaving the platform. This proactive approach allows the e-commerce company to implement targeted retention efforts before customers churn. Imagine being able to identify customers who haven't purchased in a while and offering them a personalized discount or suggesting relevant products based on their past purchases. This not only minimizes customer loss but also fosters long-term customer loyalty, ultimately translating into increased customer lifetime value and sustainable business growth.

3. Literature Review

Company and Industry Overview

The e-commerce industry has experienced phenomenal growth in recent years, driven by factors like increasing internet penetration, rising smartphone usage, and a growing preference for online shopping. This dynamic landscape is characterized by fierce competition, with established players like Amazon and Walmart vying for market share alongside a vast array of specialized online retailers.

E-commerce businesses offer a wide range of products and services, catering to diverse customer needs and preferences. They constantly strive to provide competitive pricing, convenient delivery options, and a seamless customer experience to differentiate themselves in a crowded market. However, customer acquisition is only half the battle. Retaining existing customers and fostering loyalty is crucial for sustainable business growth in the e-commerce industry.

Overview of Theoretical Concepts

- **Customer Churn:** Customer churn refers to the phenomenon where customers cease doing business with a company. It's a significant concern for e-commerce businesses, as studies have shown that acquiring new customers can be five to ten times more expensive than retaining existing ones. Understanding and predicting churn is crucial for developing effective retention strategies.
- **Customer Lifetime Value (CLTV):** Customer Lifetime Value (CLTV) represents the total revenue a customer is expected to generate throughout their relationship with the company. Churn directly reduces CLTV, as lost customers no longer contribute future revenue. Predicting churn allows e-commerce companies to identify at-risk customers and implement targeted retention efforts, ultimately maximizing CLTV.
- **Machine Learning for Churn Prediction:** Machine learning is a powerful tool that can be leveraged to predict customer churn in e-commerce. These algorithms can analyse vast amounts of customer data to identify patterns and trends associated with churn behaviour. Common machine learning algorithms used for churn prediction include Logistic Regression, Decision Trees, Random Forest, and Gradient Boosting Machines. These algorithms can be trained on historical customer data to learn the characteristics of customers who churn and then use this knowledge to predict which existing customers are at risk of leaving the platform.

Survey of Existing Models

Understanding Why Customers Leave

In the cutthroat world of e-commerce, keeping customers happy is crucial. This section dives into how online stores predict when a customer might stop shopping with them (churn). We'll explore the data they use, the clever machines that analyze it, and how they measure success.

Data Detectives: Uncovering Clues

Imagine a detective board filled with information about customers. Here's what online stores consider:

- **Buying Habits:** How often you buy, how much you spend, and what you typically purchase. Did you ever leave items in your cart but not buy them?
- **Your Background:** Your age, location (if available), and income (if provided).
- **Customer Service:** How often you contact them and why (complaints, returns, etc.).
- **Website Activity:** How often you log in, what you look at, if you write reviews, and how long you stay browsing.

The Superpower of Machine Learning

These data points are fed into special machines called algorithms. Think of them as super smart guessers that learn patterns to predict which customers are more likely to churn. Here are a few popular types:

- **Simple Choice Machine (Logistic Regression):** Good at yes or no questions (churn or not churn?) and can explain its reasoning.
- **Decision Tree Machine:** Works like a flowchart, asking questions about your data to decide if you're likely to churn. Easy to understand why it makes a guess.
- **Forest of Decision Trees (Random Forest):** A powerful team of decision tree machines, good at finding complex patterns, but harder to understand why it thinks you might churn.
- **Super Learner Machine (Gradient Boosting Machine):** Combines the best of all the machines for even better guesses, great for tricky patterns, but again, not the best at explaining its reasoning.

Measuring Success:

Just like detectives checking their clues, online stores use fancy scoring systems to see how accurate their guesses are:

- **Right On The Money (Accuracy):** How many people did the machine guess correctly (churn or not churn)?
- **Good at Spotting Churners (Precision):** How many identified churners were actually churners (not mistaken for someone who will stay)?
- **Not Missing Churners (Recall):** How many actual churners did the machine catch?

By continuously improving churn prediction, online stores can become better at keeping you, the customer, happy and shopping with them!

4. Methodology

Exploratory Data Analysis (EDA) and Business Implications

Following the data collection process outlined earlier, we can begin visually inspecting the data to gain initial insights and prepare it for model development.

About the Data:

Type	**Name**	**Desctiption**
Target Variable	Churn	Indicates if the account has churned
Account-Level Features	AccountID	Unique identifier for each account
	Tenure	Duration of the account in months
	City_Tier	Tier of the city where the account is located
	Payment	Type of payment method used
	Account_user_count	Number of users associated with the account
	account_segment	Segment to which the account belongs
	rev_per_month	Revenue per month
	rev_growth_yoy	Year-over-year revenue growth
	coupon_used_l12m	Coupons used in the last 12 months
	cashback_l12m	Cashback received in the last 12 months
Customer-Level Features	Gender	Gender of the customer
	Marital_Status	Marital status of the customer
	Login_device	Device used to log in
Customer Service Interaction Features	CC_Contacted_L1_2m	Contact with customer service in the last 1-2 months
	Day_Since_CC_connect	Days since last contact with customer service
	CC_Agent_Score	Customer service agent score
	Complain_l12m	Complaints lodged in the last 12 months

Table 1: Overview on Columns in the Dataset

Description	Value
Total number of rows	11,260
Total number of columns	19
Number of Numerical columns	14
Number of object (categorical) columns	5
Number of Numerical rows with missing values	785
Number of object rows with missing values	1,891

Table 2: Overview on data available

The data exploration process began with univariate analysis, then dove deeper into relationships between variables and churn using bivariate analysis, and culminated in multivariate analysis to examine the interplay of multiple factors.

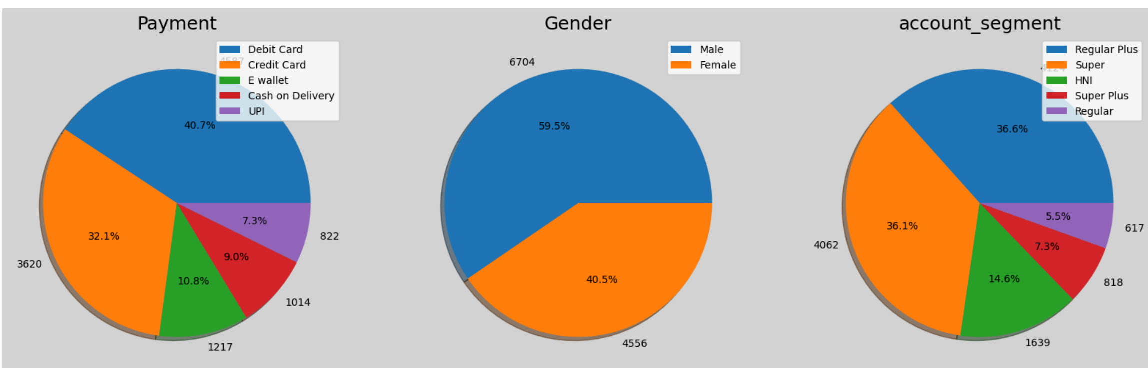

Fig 1: Pie charts of Payment used for purchasing, Gender and Account segment

Analysing customer data revealed interesting trends in payment methods and account types. Credit cards emerged as the clear winner, with over 40% of customers utilizing them for transactions. This preference suggests that customers might value the flexibility and potential rewards offered by credit cards compared to debit cards or e-wallets. Perhaps the ability to manage cash flow better or earn points for future purchases are driving factors.

On the account side, Regular Plus accounts hold the dominant position, accounting for a significant 36.6% of the customer base. This finding highlights a potential sweet spot for attracting new customers. By understanding the characteristics and preferences of this prevalent segment, businesses can tailor their offerings and marketing strategies to resonate with this key demographic. Analysing what attracts customers to the Regular Plus account could provide valuable insights into crafting similar, attractive options for new users.

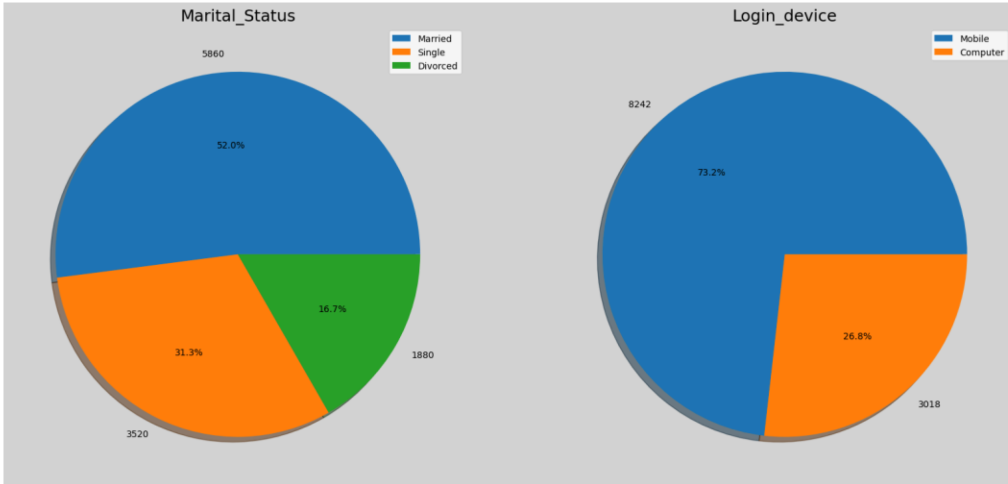

Fig 2: Pie chart of Marital Status of user and Login device used for ordering

Our survey shows marriage is the dominant marital status (52%), followed by single (31.3%). Interestingly, mobile devices reign supreme for logins (73.2%) compared to computers (26.8%).

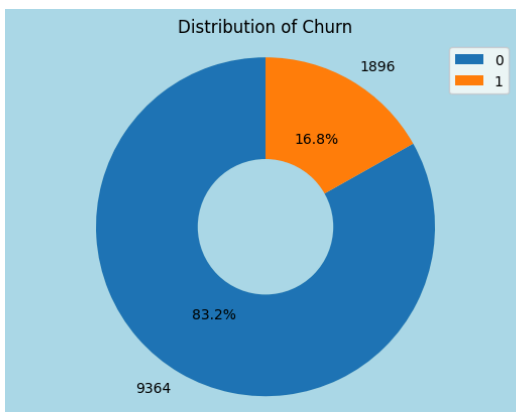

Fig 3: Distribution of Churn

This chart depicts customer churn rate, a key metric for business health. The positive news is that customer churn is relatively low at 16.8%. This indicates a strong and healthy customer base, with the vast majority (83.2%) of customers staying on board within the measured timeframe. A low churn rate suggests customer satisfaction with the service and a strong likelihood of continued business.

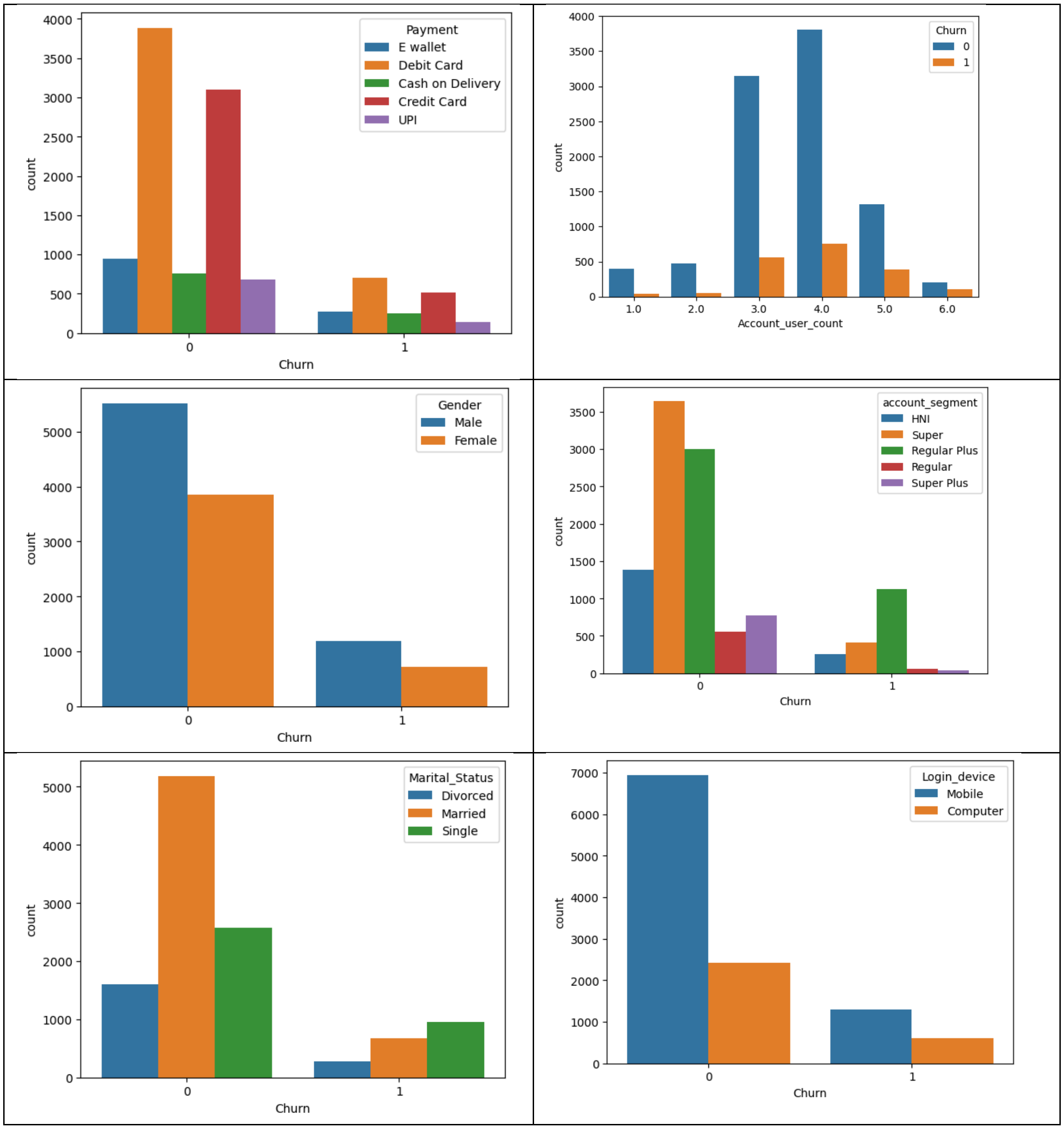

fig 4: Bar chart of Various Factors v/s Churn

The report examines customer churn through the lens of several factors:

- **Payment Methods:** While the provided charts doesn't directly relate to churn, it highlights popular payment methods (e-wallet, debit card, credit card, cash on delivery, UPI). Analysing churn alongside payment methods might reveal if specific payment methods correlate with higher churn rates.
- **Account Users:** The bar chart indicates a positive correlation between the number of users per account and customer retention. Accounts with a single user churn more frequently compared to those with multiple users. This could be due to factors

like shared accounts, subscription value for multiple users, or free accounts for single users.

- **Customer Gender:** The bar chart suggests a higher churn rate among male customers compared to females. However, without additional context, it's difficult to determine if gender directly causes churn. Further investigation is needed to understand the underlying reasons.
- **Account Segment:** The bar chart reveals a higher churn rate for "Super Plus" accounts compared to other segments ("Regular" having the lowest). This might be due to price sensitivity in the "Super Plus" segment or a mismatch between offered features and their needs. Segments like "HNI" (High Net Worth Individual) might have different churn metrics due to customer lifetime value considerations.
- **Device Usage:** The bar chart indicates a higher churn rate among users who primarily log in via mobile devices compared to computer users. This could be due to a less optimal mobile user experience, lower engagement on mobile, or demographic factors associated with mobile usage.

Data Cleaning and Preprocessing

- **Missing Values:**
  - We have 2676 missing data points across 19 features.
  - We'll fill in the gaps (imputation) or remove them strategically based on data type (numerical vs categorical).
- **Data Outliers:**
  - We'll identify outliers that skew our data's view of reality.
  - We'll either bring them in line or remove them entirely based on their impact.
- **Data Transformation:**
  - Some features need adjustments for our models to understand them better.
    - Numerical features might be reshaped.
    - Categorical features will be converted into a format our models can use, like one-hot encoding. This creates unique binary features for each category allowing the model to understand their significance.

By cleaning and prepping the data, we'll have a well-trained force ready to predict churn and retain valuable customers.

Model Building and Validation

The landscape of machine learning models offers a variety of options, each with its own strengths and weaknesses. Here are some prominent contenders we'll consider for churn prediction:

- **Logistic Regression:** A well-established algorithm suitable for binary classification tasks like churn prediction (churn or not churn).

- **Decision Trees:** These flexible and adaptable models can handle both numerical and categorical data, offering clear interpretations of their decision-making process.
- **Random Forest:** This powerful ensemble method combines the strengths of multiple decision trees, resulting in improved accuracy and reduced overfitting (overly relying on training data).
- **Gradient Boosting:** Another ensemble technique, where models are built iteratively, each one learning from the errors of the previous one, leading to continuous improvement.

Model Training: Equipping the Model

Similar to training any professional, our model requires training data. We'll divide our data into two sets:

- **Training Set:** This larger set will be used to train the model, allowing it to learn the patterns that differentiate churning customers from loyal ones.
- **Testing Set:** This unseen set will be used to evaluate the model's performance on new data, reflecting its effectiveness in real-world scenarios.

Model Evaluation: Assessing Performance

Once trained, we'll evaluate the model's performance using various metrics:

- **Accuracy:** The overall proportion of correct predictions (correctly identifying churned and non-churned customers).
- **F1 Score:** A balance between precision (predicting true churners) and recall (not missing actual churners).
- **Recall:** The proportion of actual churned customers who were correctly identified by the model.
- **Precision:** The proportion of predicted churned customers who actually churned (avoiding unnecessary outreach to non-churning customers).

Model Name	Accuracy	F1 Score	Recall Score	Precision Score
Logistic Regression	0.8956	0.601	0.4797	0.8045
Decision Tree	0.9347	0.8073	0.8347	0.7817
Random Forest	0.9729	0.9127	0.8645	0.9667
XGBoost Classifier	0.8832	0.4971	0.3523	0.8442

Table 3: Comparison of Models

Key Findings

- Random Forest emerged as the leading performer with the highest Accuracy (0.9729) and Precision Score (0.9667). This indicates its ability to accurately predict both churned and non-churned customers, minimizing false positives (unnecessary outreach).

- Decision Tree achieve-d a good balance between Recall and Precision, effectively identifying churned customers. However, its overall Accuracy was lower than Random Forest.
- Logistic Regression offered a lower F1 score compared to Decision Tree, suggesting it might miss a higher proportion of churned customers despite having a similar Accuracy.
- XGBoost Classifier exhibited the lowest performance among the evaluated models based on the provided metrics.

5. Results and Discussion

Findings Based on Observations:

1. **Dataset Size and Features**:
    - The dataset contains information on 11200+ customers.
    - It includes 19 features, both numerical and categorical.
    - The target variable, 'Churn,' indicates whether a customer has churned.
2. **Imbalanced Data**:
    - The data is imbalanced, with only 16.5% of customers having churned.

Findings Based on Analysis of Data:

1. **Customer Tenure**:
    - Average tenure of customers: 11 months.
2. **Users Per Account**:
    - Average number of users per account: 3.6.
3. **Revenue Metrics**:
    - Average revenue per month: 6.3.
    - Total revenue per month: 71356.0
    - Average revenue growth year-over-year: 16.2
    - Total revenue growth year-over-year: 182337.0
4. **Payment and Usage**:
    - Average number of coupons used for payment: 1.7
    - Total number of coupons used for payment: 20160.0
    - Average number of days since the customer connected: 4.6 days.
    - Average cashback amount: 196.23.
    - Total cashback amount: 2209621.52
5. **Demographic and Account Information**:
    - Most common city tier: Tier 1, followed by Tier 3 and Tier 2.
    - Most common account segment: Regular, followed by Super Plus and Regular Plus.

Category	Count
Regular Plus	4124
Super	4062
HNI	1639
Super Plus	818
Regular	617

Table 4: customer count on account segment

- Most common payment method: Debit Card, followed by Credit Card and E wallet.

Payment Method	Count
Debit Card	4696
Credit Card	3511
E wallet	1217
Cash on Delivery	1014
UPI	822

Table 5: customer count on Payment Method

- Most common login device: Web, followed by Mobile.

Device	Count
Mobile	8021
Computer	3239

Table 6: customer count on Device used

General Findings:

Features	Correlation with Churn
account_segment_Regular Plus	0.213095
Marital_Status_Single	0.181387
Account_user_count	0.10584
CC_Agent_Score	0.105165
City_Tier	0.083853
Tenure	-0.230327
Marital_Status_Married	-0.151135
Day_Since_CC_connect	-0.145553
account_segment_Super	-0.132436
account_segment_Super Plus	-0.089368

Table 7: Correlation Insights

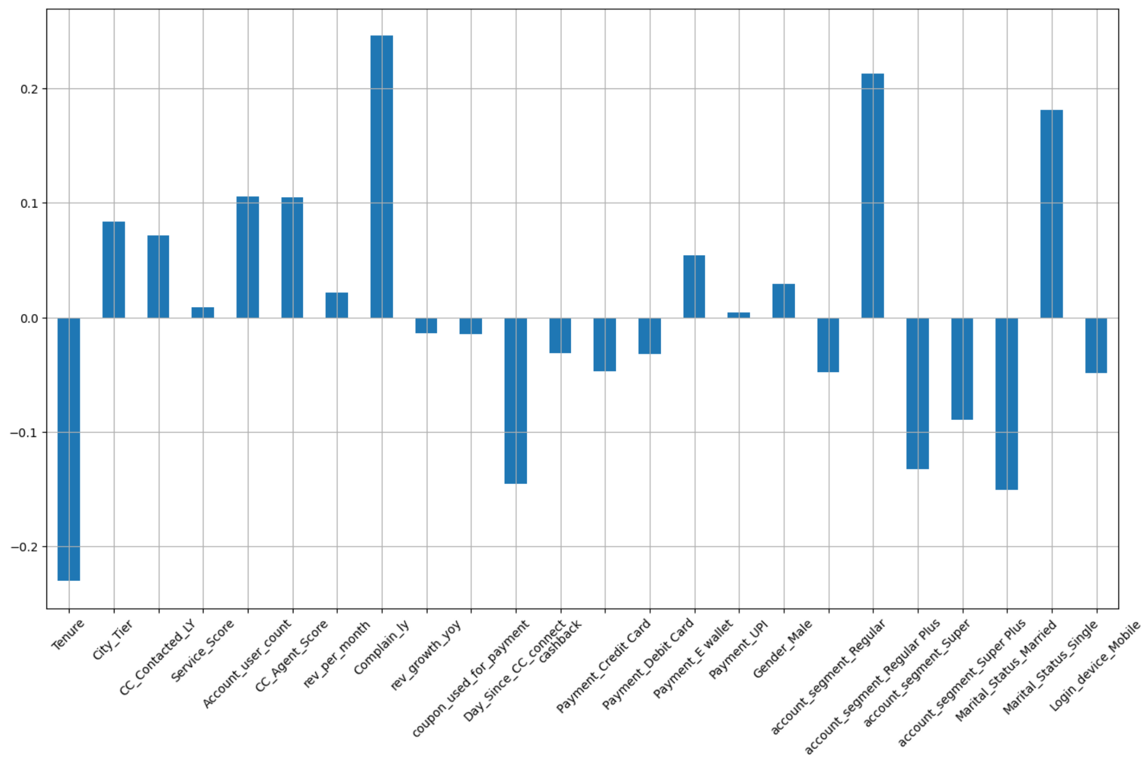

Fig 5: Corelation of columns with Churn

Positive Correlations with Churn:

1. **Regular Plus Account Segment:** Customers in this segment are more likely to churn, suggesting issues specific to this group.
2. **Single Marital Status:** Single individuals show a higher likelihood of churn, potentially due to different needs or expectations.
3. **Account User Count:** Higher user counts are linked to increased churn, indicating challenges with larger account structures.
4. **CC Agent Score:** Surprisingly, higher scores for customer service interactions correlate with churn, indicating potential dissatisfaction despite good service.
5. **City Tier:** Customers in higher city tiers are slightly more likely to churn, possibly due to market dynamics or competition.
6. **Complaint:** Customers who are raising complaints regularly are about to churn higher than other reasons

Negative Correlations with Churn:

1. **Tenure:** Longer-tenured customers are less likely to churn, reflecting stronger loyalty and satisfaction.
2. **Married Marital Status:** Married individuals exhibit lower churn rates, suggesting greater stability and satisfaction.
3. **Days Since CC Connection:** Longer duration since connecting the credit card correlates with lower churn, indicating stronger loyalty.

4. **Super and Super Plus Account Segments:** High-value segments show lower churn rates, highlighting the importance of customer value.

These insights help prioritize retention efforts by addressing factors most strongly associated with churn.

Model Performance:

- The Random Forest Classifier is the best model for predicting churn, with:
    - Accuracy: 96.8%
    - F1 Score: 89.59%
    - Recall Score: 84.01%
    - Precision Score: 95.97%

6. Recommendations and Conclusion

Recommendations

for a churn prediction model for e-commerce customer retention are comprehensive and strategic. Let's break down each point:

1. **Churn Prevention Strategies:** Identifying key customer segments with indicators of potential churn, such as low tenure, low revenue per month, and stagnant revenue growth, is crucial. By tailoring specific churn prevention strategies to these segments, the company can effectively address their needs and concerns before they decide to leave.
2. **Incentives for Loyalty:** Offering incentives like loyalty programs or discounts can foster a sense of loyalty among customers, encouraging them to continue shopping with the company. These incentives provide tangible benefits that can increase customer tenure and overall satisfaction.
3. **Customer Engagement:** Enhancing customer engagement through personalized recommendations and relevant content is essential for retaining customers. By leveraging data analytics and machine learning algorithms, the company can deliver targeted recommendations that resonate with individual preferences, thereby keeping customers engaged and interested in the platform.
4. **Monitoring and Identification:** Continuous monitoring of customer behaviour allows the company to stay vigilant for signs of potential churn. Utilizing a churn prediction model enables proactive identification of at-risk customers, empowering the company to intervene and implement retention strategies before it's too late.
5. **Proactive Steps:** Taking proactive measures such as offering personalized offers, improving customer service, and addressing pain points identified through feedback demonstrates the company's commitment to customer satisfaction. By actively listening to customer feedback and responding promptly to their needs, the company can build stronger relationships with its customers and reduce churn rates.

Implementing these recommendations can indeed lead to a significant reduction in churn rates, an improvement in customer satisfaction, and ultimately, enhanced revenue growth for the e-commerce platform.

Suggestions for Areas of Improvement

1. **Data Quality and Completeness**:
   - **Address Missing Data**: Improve data collection methods to reduce missing values. Implement better tracking systems to ensure comprehensive data capture for all customer interactions.
   - **Enhance Data Integration**: Integrate data from various sources (e.g., social media, customer support interactions) to get a more holistic view of customer behavior.

2. **Model Refinement**:
    - **Feature Engineering**: Develop new features that may improve the predictive power of the model, such as social media engagement metrics or detailed browsing behaviour patterns.
    - **Algorithm Tuning**: Continuously experiment with different algorithms and hyperparameter tuning to improve model accuracy and reliability.
3. **Customer Segmentation**:
    - **Granular Segmentation**: Develop more granular customer segments to tailor retention strategies more precisely. Use clustering techniques to identify distinct customer personas and their unique needs.
4. **User Experience**:
    - **Mobile Optimization**: Given the higher churn rate among mobile users, focus on optimizing the mobile shopping experience. Ensure the mobile app is user-friendly, fast, and offers a seamless shopping experience.
    - **Simplify Checkout**: Reduce friction in the checkout process by simplifying steps, offering multiple payment options, and providing clear information about fees and policies.

Scope for Future Research

1. **Advanced Machine Learning Techniques**:
    - Explore the use of advanced machine learning techniques such as deep learning, neural networks, and ensemble methods to improve churn prediction accuracy.
    - Investigate the potential of real-time predictive analytics to identify churn risks as they emerge.
2. **Incorporating Additional Data Sources**:
    - Integrate additional data sources such as social media activity, customer sentiment analysis, and external market data to enrich the predictive model.
    - Explore the impact of external factors (e.g., economic conditions, competitive actions) on customer churn.
3. **Longitudinal Studies**:
    - Conduct longitudinal studies to track customer behaviour over time and understand the long-term effectiveness of retention strategies.
    - Analyse the impact of different retention interventions on customer lifetime value (CLTV) and overall business growth.
4. **Customer Behaviour Analysis**:
    - Dive deeper into understanding the behavioural triggers that lead to churn. Use qualitative research methods such as customer interviews and focus groups to gain deeper insights.

Conclusion:

The development of a Churn Prediction Model for E-commerce Customer Retention represents a significant step forward in addressing one of the most pressing challenges faced by online retailers. Through this project, we have delved deep into the complexities of customer churn, exploring its underlying causes and implications for business sustainability. Leveraging the power of machine learning and data analytics, we have crafted a predictive model capable of identifying customers at risk of abandoning the platform, thus enabling proactive intervention and targeted retention efforts.

The journey through this project has been illuminating, revealing insights into the diverse factors influencing customer churn within the e-commerce landscape. From analyzing customer demographics and purchase behaviors to evaluating the impact of customer service interactions and payment methods, every aspect has contributed to a comprehensive understanding of churn dynamics. Our findings underscore the importance of personalized strategies, customer engagement, and proactive measures in fostering long-term customer loyalty and driving sustainable business growth.

Among the various models evaluated, the Random Forest algorithm emerged as the most effective in predicting churn, demonstrating exceptional accuracy and precision. Its ability to discern subtle patterns and make reliable predictions empowers e-commerce companies to make informed decisions and implement targeted retention strategies with confidence.

As we look to the future, the insights gleaned from this project pave the way for continued innovation and refinement in customer retention strategies. By embracing data-driven approaches and leveraging advanced analytics techniques, e-commerce companies can stay ahead of the curve, anticipating customer needs, and delivering unparalleled experiences that foster lasting relationships.

In essence, the Churn Prediction Model presented in this project serves as a valuable tool for e-commerce companies striving to navigate the dynamic landscape of customer retention. By harnessing the power of data and machine learning, we can turn the tide against churn, cultivate customer loyalty, and chart a course towards sustained success in the competitive realm of online retail.