Final Project:

Topic: Predictive Modeling for Customer Retention in Mobile Food Delivery

McKenzie Payne

DSC 550

Introduction:

In today's competitive landscape, customer churn poses a significant challenge for businesses, particularly in the mobile app-based food delivery sector. Customer churn refers to the phenomenon where customers discontinue their engagement with a service, leading to a loss in revenue and market share. For the food delivery service, understanding and predicting customer churn is vital for maintaining sustainable growth and profitability. The primary objective of this project is to develop a predictive model aimed at addressing the challenge of customer churn within a mobile app-based food delivery service. The goal of the model is to predict whether a customer is likely to churn based on their historical interactions with the food delivery app, enabling the business to implement targeted retention strategies.

Importance and Justification:

Solving the problem of customer churn is crucial for several reasons. Firstly, churn directly impacts the bottom line by reducing revenue and increasing customer acquisition costs. Retaining existing customers is often more cost-effective than acquiring new ones, as loyal customers tend to spend more over time and provide free word-of-mouth marketing. Additionally, by understanding the underlying reasons for churn, businesses can implement targeted retention strategies to address customer dissatisfaction and improve overall satisfaction. This enhanced customer loyalty leads to repeat business and strengthens the brand's reputation in the market, providing a competitive edge.

Pitch to Stakeholders:

When pitching this problem to stakeholders, it is essential to emphasize the potential financial implications of customer churn and the benefits of addressing it proactively. Presenting data-backed evidence of the impact of churn on revenue and profitability can help stakeholders recognize the urgency and importance of investing resources into churn prediction and mitigation efforts. Highlighting the competitive landscape and the need to differentiate through superior customer service and retention strategies can further drive buy-in from stakeholders.

Official Pitch: “*Addressing customer churn is critical for our mobile food delivery service, as it directly impacts our revenue and profitability. Our recent project has successfully developed a predictive model that identifies customers at risk of churn with 85% accuracy and a 0.90 ROC-AUC score. By leveraging this model, we can proactively target these customers with personalized retention strategies, such as optimized delivery times, enhanced food quality, and tailored marketing campaigns. These insights not only enhance customer satisfaction and loyalty but also reduce customer acquisition costs, ultimately boosting our bottom line. Investing in the deployment and continuous refinement of this model will position us ahead of competitors and ensure sustainable growth in a competitive market."*

Data Source:

The data for this project was obtained from Kaggle. It comprises a comprehensive dataset containing various customer attributes, transactional data, and feedback mechanisms within the food delivery app. This dataset provides a robust foundation for building predictive models and gaining insights into customer behavior.

Milestones 1-3 Summary

EDA:

In Milestone 1, exploratory data analysis (EDA) was conducted to gain insights into customer demographics, preferences, and behaviors. Visualizations such as histograms, count plots, and correlation matrices were utilized to understand the distribution and relationships among different variables. Key insights from EDA included that most users were between the ages of 25-34, customers who ordered more frequently were less likely to churn, and higher average order values were associated with lower churn rates. These insights informed subsequent data preparation and modeling steps, ensuring a targeted approach in addressing customer churn.

Data Preparation:

Milestone 2 focused on preparing the data for predictive modeling. This involved several steps, including dropping unnecessary features, selecting relevant columns, transforming variables, handling missing data, and encoding categorical variables. Feature engineering techniques were also applied to derive new features, such as customer tenure and discount usage, which could enhance the predictive power of the models. These steps ensured that the dataset was clean, relevant, and optimized for building accurate predictive models.

Model Building and Evaluation:

In Milestone 3, predictive models were selected, built, and evaluated for customer churn prediction. Two models, Logistic Regression and Random Forest Classifier, were chosen for their suitability and performance in binary classification tasks. The models were trained on the prepared dataset and evaluated using various performance metrics. The Logistic Regression model achieved an accuracy of 78%, precision of 0.76, recall of 0.79, and an ROC-AUC score of 0.82. The Random Forest Classifier outperformed it, achieving an accuracy of 85%, precision of 0.83, recall of 0.86, and an ROC-AUC score of 0.90. These results indicate the effectiveness of the models, particularly the Random Forest Classifier, in identifying customers at risk of churn.

Conclusion:

Analysis Insights:

The analysis provided valuable insights into customer behavior and factors influencing churn within the food delivery service. Both models demonstrated strong predictive performance, with the Random Forest Classifier achieving an accuracy of 85% and an ROC-AUC score of 0.90, indicating their effectiveness in identifying customers at risk of churn. Key factors influencing churn included customer tenure, order frequency, average order value, and discount usage. These insights can inform strategic decision-making and retention efforts to improve customer satisfaction and loyalty, ultimately reducing churn rates and boosting profitability.

Model Deployment Readiness:

While the models show promising performance, further validation and testing are necessary before deployment into production. Continuous monitoring and refinement of the models based on real-world feedback will be crucial to ensure their effectiveness and reliability over time. This iterative process will help maintain the models' accuracy and relevance in a dynamic market environment.

Recommendations:

Based on the analysis, several recommendations for improving customer retention strategies are proposed. These include optimizing delivery times to ensure timely deliveries, enhancing food quality to maintain high standards, personalizing offerings based on customer preferences, implementing targeted marketing campaigns aimed at high-risk customers, and developing loyalty programs to reward repeat customers. These strategies can help address the factors contributing to churn and improve overall customer satisfaction and loyalty.

Challenges and Opportunities:

Challenges in deploying the models include the dynamic nature of customer preferences, data privacy concerns, and the need for ongoing model maintenance. However, there are also significant opportunities for further exploration, such as incorporating real-time data streams, leveraging advanced machine learning techniques, and exploring customer segmentation strategies to enhance model performance and predictive accuracy. These advancements can provide deeper insights and more effective retention strategies.

Great Results and Future Prospects:

The project successfully developed a robust predictive model for customer retention in a mobile food delivery service, demonstrating strong predictive capabilities. These results highlight the potential for substantial improvements in customer retention strategies, providing a clear path to enhance customer satisfaction and profitability. With further refinement and real-world testing, the models can be fully integrated into the business processes, driving significant value for stakeholders. The insights and methodologies developed in this project can also be adapted for similar challenges in other sectors, showcasing the broader applicability and impact of advanced predictive analytics.