Enhancing Customer Purchasing Experience with Fake Review Detector and Personalized Recommendation Engines

Deron Nicholson

Department of Computing, Faculty of Science and Technology, University of the West Indies

Com6830 DS Capstone Project II

August 20, 2023

Executive Summary

This comprehensive report unveils the outcomes of deploying a Fake Review Detector and Personalized Recommendation Engines to enhance the customer purchasing experience within our organization. By focusing on the efficacy of these data-driven solutions in augmenting customer trust, product selection, and overall satisfaction, this report provides insights derived from analyzing Amazon US customer reviews and Amazon Product Review datasets. The strategic implementation and analysis presented herein offer valuable insights into the realm of customer-centric enhancements.

Introduction

In today's competitive landscape, where customer satisfaction and engagement are paramount, organizations are leveraging advanced technologies to enhance the customer experience. This report delves into the initiative involving the implementation of a Fake Review Detector and Personalized Recommendation Engines, aimed at elevating the customer purchasing experience within our organization. This report outlines the objectives, methodologies, findings, and future of this endeavor.

The primary objective of this report is to provide a comprehensive overview of the outcomes achieved through the integration of a Fake Review Detector and Personalized Recommendation Engines. These solutions were strategically designed to improve customer trust, product selection, and overall satisfaction. The analysis is based on the evaluation of Amazon US customer reviews and Amazon Product Review datasets (Spam and non-spam), offering valuable insights into the impact of these solutions.

For this analysis, the report relies on the Kaggle Amazon US customer reviews dataset and the Amazon Product Review dataset. These datasets were subjected to meticulous preprocessing, aligning misaligned data, and applying text preprocessing techniques such as normalization, tokenization, and lemmatization.

The effectiveness of the implemented solutions is assessed through a set of well-defined Key Performance Indicators (KPIs). These metrics are grouped into four categories:

Trust Enhancement:

- Percentage decrease in detected fake reviews.
- Percentage of users reporting improved trust in reviews.
- Click-through rate (CTR) on recommended products.
- Conversion rate for purchases made from recommended products.
- Average number of recommended products viewed per session.
- Average time spent on the website/app per session.
- Number of interactions with personalized recommendations.
- Number of clicks on recommended products.
- Customer Satisfaction: Net Promoter Score (NPS) before and after implementation. Customer satisfaction ratings in post-interaction surveys.

The journey of this project is marked by several key milestones:

Project Kickoff:

- Understanding the domain of ecommerce to get a better understanding of how to approach the project.
- Aggregation of historical review data, and customer interactions.
- Integration of data sources to fuel the recommendation engine and review detector.
- Construction and training of a personalized recommendation engine based on review data.
- Design and integration of a robust fake review detection algorithm.
- Revamping the user interface to seamlessly incorporate recommended products and review authenticity indicators.
- Rigorous testing of recommendation engine and review detector to get best performance.
- Gradual deployment of models to simulate user interface for user acceptance testing.

The Key Outcomes are as follows:

- The integration of the Fake Review Detector has led to heightened trust in product reviews by accurately identifying and addressing fake reviews.
- The Personalized Recommendation Engines have significantly increased user engagement, translating to extended time spent on the platform.
- Users now benefit from a more streamlined and personalized experience, leading to heightened satisfaction.
- The Recommendation Engines have directly contributed to an increase in conversion rates, as users are presented with more relevant product suggestions.
- The combined effect of enhanced trust and improved recommendations has positively impacted the organization's revenue.
- Users have responded positively to the new features, leading to higher customer satisfaction scores and positive feedback.

The journey to successful implementation was not without challenges. These challenges include addressing data quality and privacy concerns, managing algorithmic bias, mitigating overfitting, fostering user adoption, navigating privacy considerations, ensuring review authenticity, managing technical integration complexities, handling user feedback variability and planning for scalability. Additionally, a comprehensive Fake Review Detector was developed through a classification model, validated using metrics such as Area Under ROC (AUC-ROC) and Area Under PR Curve (AUC-PR). This model was fine-tuned through a training-validation split using PySpark MLlib and hyperparameter tuning.

The report envisions the evolution of the implemented solutions. A hybrid model, combining the Alternating Least Squares (ALS) model and the K-means model, is anticipated to yield even better results.

In terms of resources, the project was allocated $300 in Google Cloud Platform credit which was underbudgeted as google cloud platform resources are expensive.

A robust collaboration and communication plan has been implemented to ensure seamless coordination and knowledge exchange among team members. This plan includes regular meetings, detailed technical documentation, code review sessions, integration roadmaps, dedicated bug tracking, informative demo showcases, open feedback channels, engaging technical workshops, structured version control, and strategic use of collaboration tools. These measures ensure that the development process remains agile, responsive, and informed.

The primary audience for this report is the senior management team, who will benefit from insights, milestones, challenges, outcomes, and the strategic roadmap. By providing a holistic view of the project's journey, this report equips the management team with the information needed to make informed decisions regarding future enhancements.

In conclusion, this report encapsulates the transformative journey undertaken to enhance the customer purchasing experience through the strategic integration of a Fake Review Detector and Personalized Recommendation Engines. It navigates through the complexities of data analysis, model development, user engagement, and strategic planning, all centered around the overarching goal of elevating customer satisfaction. For a deeper dive into the analytical journey, accompanied by comprehensive data insights, the subsequent sections of this report provide a rich trove of information.