

Data Analysis Project with NLP

Project Title:

Customer Feedback Analysis for a personal financial management software product using NLP

Project Overview:

Applied Natural Language Processing (NLP) techniques to analyze user reviews for a financial management software product, extracting meaningful business insights to enhance user experience.

Tools and Libraries Used:

Python, Pandas, Matplotlib, Seaborn, NLTK, spaCy, Jupyter Notebook

Key Contributions and Achievements:

- Leveraged Python for data cleaning, exploratory data analysis (EDA), and advanced Natural Language Processing.
- Extracted and preprocessed user reviews from trustpilot's [website](#), spanning April 2022 & December 2023.
- Employed descriptive statistics to understand the dataset, comprising **6620** user feedback entries.
- Identified and categorized major pain points causing user dissatisfaction, using NLP-based thematic analysis.

Techniques Utilized:

- Data cleaning and feature engineering
- Exploratory data analysis
- Natural Language Processing
 - NLP techniques used:
 - Tokenization and text preprocessing for effective analysis.
 - Sentiment analysis to gauge user sentiments from reviews.
 - Clustering to group similar reviews and identify recurring themes.

KEY FINDINGS

Most Active Days

The most active days were Mondays, Tuesdays, Fridays & Sundays.

Mondays & Tuesdays accounted for 33.37% while **Fridays & Sundays accounted for 28.26%** of user reviews posted within the past 16 months; together, these four days accounted for approximately **62%** of all user reviews.

User Review Distribution

Our sentiment analysis model distributed user reviews into three(3) categories - positive, negative and neural. Below are the distributions across these categories:

Positive: 53.3%

Negative: 25.1%

Neutral: 21.5%

User Pain Points

To extract themes and user pain points, the negative statements were analyzed, extracting ten(10) reviews that were most representative of all review comments and the following themes and pain points were identified;

Critical Pain Points

Data Integrity Concerns:

- **Missing Transactions:** Users are losing track of vital financial data due to missing transactions, creating confusion and jeopardizing trust.
- **Download Discrepancies:** Inaccurate bank account downloads sow further doubt, making reconciliation a frustrating chore.
- **Unwanted Automation:** The automatic entry system, introduced without proper consent, feels intrusive and undermines user control.

Navigation problems:

- **Clunky Interface:** Navigating the platform feels cumbersome and unintuitive, hindering user efficiency and satisfaction.

Connectivity issues:

- **Syncing Hiccups:** Frequent errors and disruptions during synchronization disrupt workflow and erode user confidence.

- **Unreliable Updates:** Automatic transaction updates, meant to save time, instead require manual intervention, adding burden and frustration.

Other pain points

- **Bill Tracking Gaps:** Limitations in bill tracking functionality hamper users' ability to manage finances effectively.
- **Check Pay Shortcomings:** The check pay feature lacks the flexibility and control users need.
- **Desktop Performance Lag:** Sluggish desktop app performance further slows users down and impacts productivity.
- **Mac Management Misfire:** The Mac version falls short in managing stock efficiently, leaving users dissatisfied.