Customer Experience Analytics for Fintech Apps

Interim Report

by: Yonas Zelalem

June 8, 2025

Introduction

This interim report outlines progress for the Week 2 Challenge of 10 Academy's Artificial Intelligence Mastery program, analyzing Google Play Store reviews for mobile banking apps of three Ethiopian banks: Commercial Bank of Ethiopia (CBE), Bank of Abyssinia (BOA), and Dashen Bank. The goal is to scrape at least 400 reviews per bank, preprocess the data, and conduct sentiment and thematic analyses to identify customer satisfaction drivers and pain points for Omega Consultancy. Tasks 1 and 2 are complete, with results in reviews_with_themes.csv. This report details the methodology, preliminary findings, outputs, and next steps.

Data Collection and Preprocessing

Over 1,200 reviews were scraped using the google_play_scraper library for CBE (com.combanketh.mobilebanking), BOA (com.boa.boaMobileBanking), and Dashen Bank (com.dashen.dashensuperapp), meeting the KPI of 400+ reviews per bank. The scrape.py script saved reviews as CSVs with fields: review_id, review_text, rating, date, bank_name, and source. Preprocessing (preprocess.py) merged CSVs, removed duplicates, filled missing review_text with empty strings, dropped rows without rating or date, and standardized dates to YYYY-MM-DD. The cleaned_reviews.csv file has less than 5% missing data, satisfying the KPI. Code is committed to the task-1 branch with a .gitignore and requirements.txt.

Early Analysis

Sentiment Analysis

Using the distilbert-base-uncased-finetuned-sst-2-english model (sentiment.py), over 90% of reviews received sentiment_label (POSITIVE/NEGATIVE) and sentiment_score (01), meeting the KPI. Findings include:

- CBE: 60% positive reviews, praising reliability, but 40% negative, citing slow transfers.

- BOA: 55% negative, highlighting crashes and slow loading.

- Dashen Bank: 65% positive, lauding UI and budgeting tools, with some transfer speed complaints.

Thematic Analysis

Thematic analysis (thematic.py) used spaCy and TF-IDF for keyword extraction, clustering into themes: Account Access Issues, Transaction Performance, User Interface, Customer Support, and Feature Requests. At least three themes per bank were identified, meeting the KPI:

- CBE: Frequent Transaction Performance issues (e.g., slow transfer) and positive User Interface feedback.

- BOA: Dominant Account Access Issues (e.g., login error) and Transaction Performance complaints.

- Dashen Bank: Strong User Interface praise, but Transaction Performance and Feature Requests (e.g., telebirr transfers) noted.

Results are stored in reviews_with_themes.csv.

Output Review

The pipeline generated cleaned_reviews.csv (1,200+ reviews) and reviews_with_themes.csv, including sentiment_label, sentiment_score, keywords, and themes. CBE's 4.5-star rating aligns with positive sentiment, though transfer delays are a concern. BOA's 4.3-star rating reflects technical issues, while Dashen's 4.4-star rating matches UI strengths. Slow transfers are a shared pain point, relevant to Scenario 1 (user retention).

Next Steps

Task 3 will implement an Oracle database with tables for banks and reviews, inserting over 1,000 entries. Task 4 will derive at least two drivers and pain points per bank, create 3 to 5 visualizations (e.g., sentiment trends, keyword clouds), and produce a final report with recommendations. Scripts will be committed to task-3 and task-4 branches.

Conclusion

Tasks 1 and 2 successfully collected and analyzed over 1,200 reviews (1448 to be exact), revealing sentiment and thematic insights. CBE shows reliability but transfer delays, BOA struggles with technical issues, and Dashen excels in UI with similar transfer concerns. The modular pipeline and Git organization support future tasks, laying a foundation for actionable recommendations for Omega Consultancy.