Transforming User Reviews into App Strategy

A Data-Driven Customer Experience Analytics Report on Ethiopia's Mobile Banking Apps

Executive Summary

In an increasingly digital financial ecosystem, **user experience (UX)** in mobile banking is critical. Through this analysis, I extract **insights from app store reviews** of three major Ethiopian banks **CBE**, **BOA**, and **Dashen** to uncover **drivers of satisfaction** and **key pain points**.

By applying **natural language processing**, **thematic clustering**, and **enterprise-grade storage via Oracle**, this project bridges the gap between customer feedback and **business strategy**.

Project Understanding & Business Alignment

Problem Statement

Omega Consultancy, a fintech advisory firm, is supporting Ethiopian banks in their **digital transformation**. The mission:

> "Extract strategic insights from user reviews to inform mobile banking app improvements."

Objectives

1. Scrape and clean >1,200 reviews across 3 banks such as C**ommercial bank of ethiopia (CBE), Bank of Abyssinia(BOA), Dashn Bank**
2. Identifying user satisfaction drivers and pain points
3. Analyze user sentiment and themes
4. Store data using scalable, enterprise-compliant DB (Oracle)
5. Generate actionable recommendations

Scope Justification

- Focused on Google Play Store as it captures active app users
- Covers both qualitative (themes) and quantitative (ratings/sentiment) signals
- Selected models balance **accuracy, interpretability, and speed**

Key Business Challenges

Challenge	Response
Biased or vague reviews	Applied rule-based and transformer sentiment filtering
Enterprise integration requirement	Used Oracle XE with modular Python insertion
Identifying actionable patterns	Created theme mapping + keyword clustering

Methodology & Technical Execution

Tools & Libraries

- **Scraping**: `google-play-scraper`
- **NLP**: `TextBlob`, `VADER`, `DistilBERT`, `TF-IDF`, `spaCy`
- **DB**: `Oracle XE` (via Docker) with full Python-based CRUD
- **Visualization**: `Seaborn`, `Matplotlib`, `WordCloud`

Workflow

1. **Data Collection**: Reviews were scraped using the google-play-scraper Python library and official app IDs, targeting:
    - 400+ reviews per bank (Total: ~1,400)
    - Fields collected: review, rating, date, bank, source
2. **Cleaning**: deduplication, missing value handling, date normalization
3. **Sentiment Analysis**:
    - Applied 3 techniques
        - Simple: `TextBlob`, `VADER`
        - Advanced: `DistilBERT`
    - `Labels: Positive, Negative, Neutral`
    - `Scores: Normalized polarity`
4. **Keyword Extraction**: TF-IDF across cleaned review tokens
5. **Theme Mapping**: Rule-based logic into 6 core business-relevant themes:
    - Account Access, Transaction Issues, App Functionality, Customer Support, Performance, Account Info
6. **Database Storage**: Normalized Oracle schema (`banks`, `reviews`),Database Integration
    - Setup Oracle XE using Docker
    - Inserted 1,200+ records via `oracledb` Python driver
7. **Insight Extraction**: Visual & tabular summaries

Technical Considerations

- `DistilBERT` was chosen for **accuracy** and **neutral bias detection**
- Oracle was used over SQLite/PostgreSQL for **enterprise relevance**
- Theme extraction preferred **interpretability over unsupervised LDA**, aligning better with business decision-making

Results & Analysis

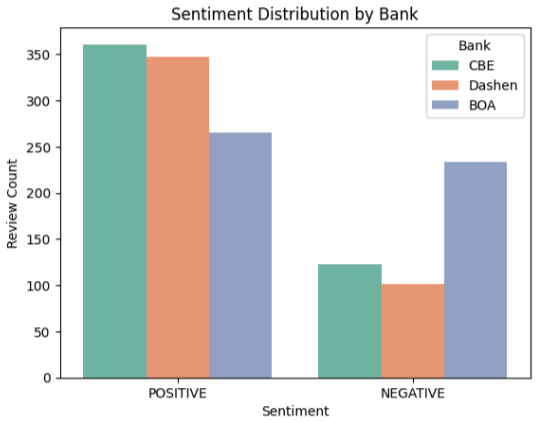
Summary Table of Theme Mentions

Theme	BOA	CBE	Dashen
Account Access	9	0	4
App Functionality	27	12	8
Transaction Issues	9	17	15
Performance	14	2	9
Customer Service	1	2	4
Other (Generic)	438	451	408

> 📌 Most reviews were generic praise (e.g., “good app”), but i identified clear pain points in ~10% of reviews.

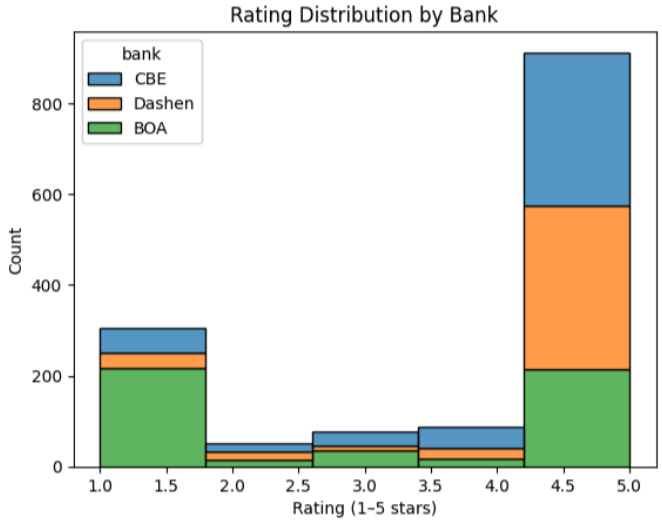
📈 Visual 1: Sentiment Distribution by Bank

CBE leads in positive reviews; BOA shows more negative signals.

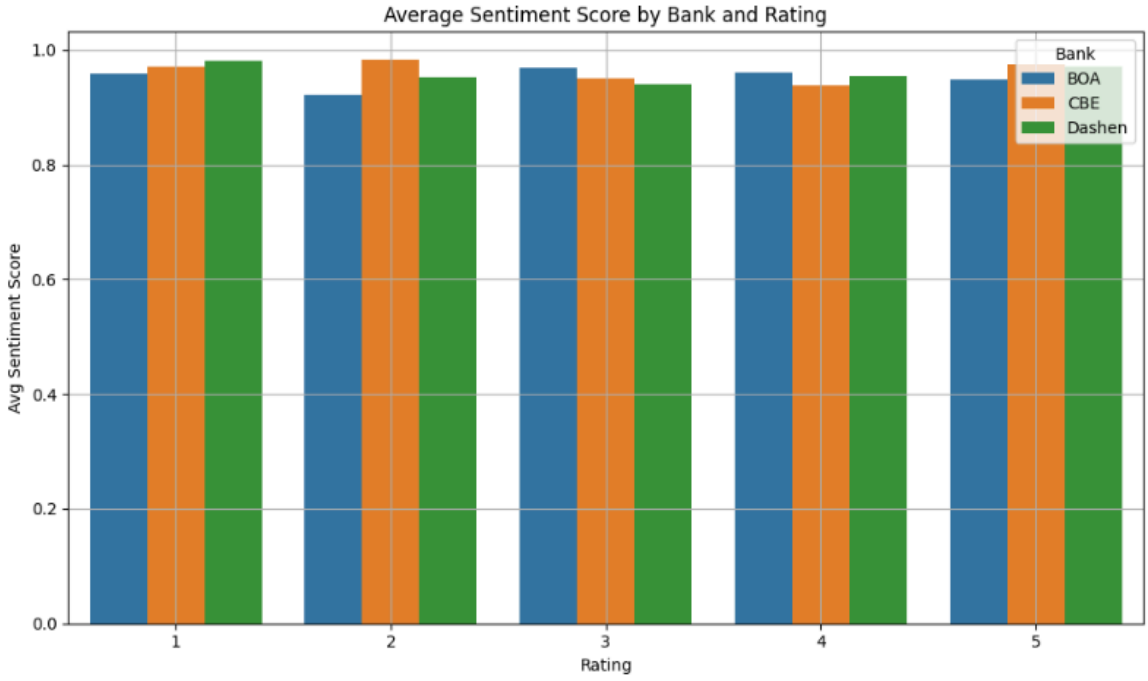


📈 Visual 2: Rating Histogram

BOA has a lower proportion of 4–5 star reviews compared to CBE/Dashen.

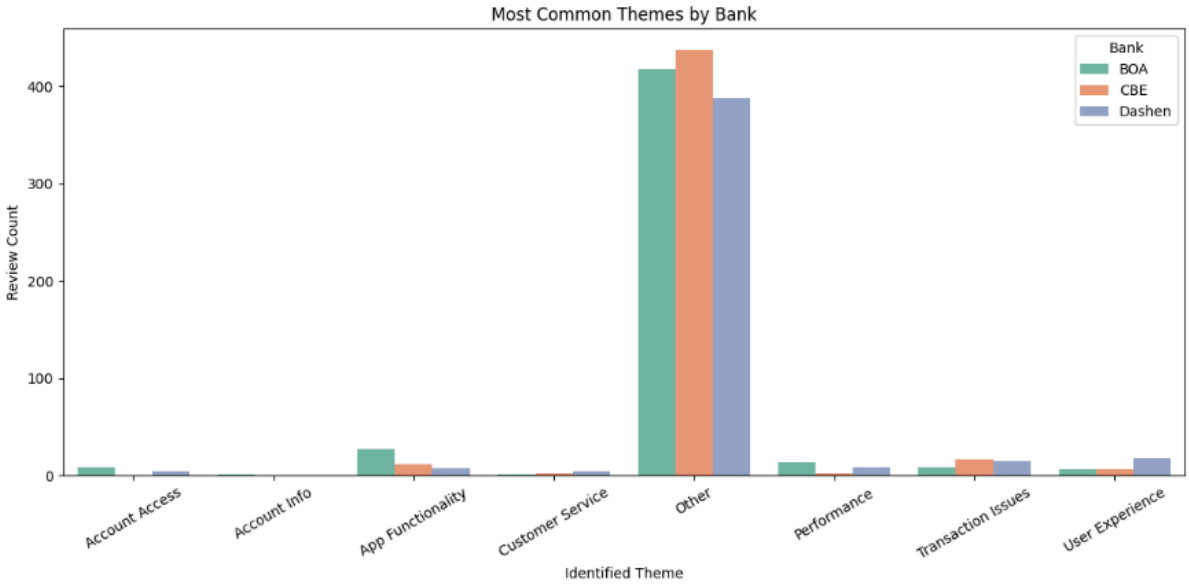


📈 Visual 3: Average Sentiment Score by Bank and Rating



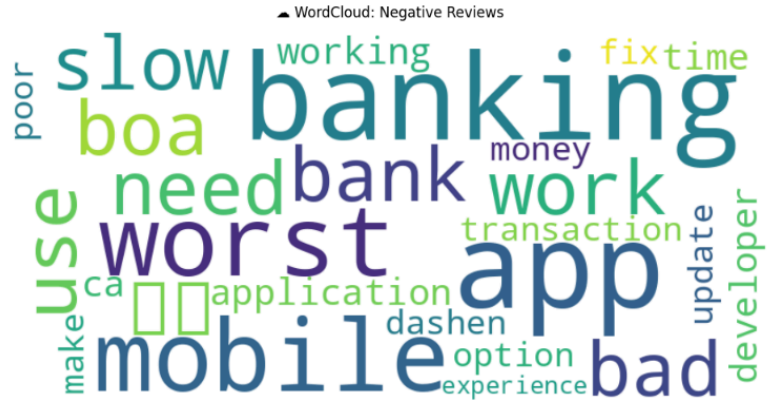
📈 Visual 4: Theme Distribution

Each bank's pain points are distinct and actionable. But most themes belong to other



📈 Visual 5: Word Cloud for Positive Reviews

📈 Visual 6: Word Cloud for Negative Reviews



Top Keywords (TF-IDF Results)

Bank	Top Keywords
CBE	good, app, nice, best app, transfer
BOA	login, worst, crash, password
Dashen	UI, crash, slow, wow, amazing

Key Insights

Bank	Drivers	Pain Points
CBE	Positive mention of features	**Transaction reliability, Failed transfers** (17 mentions)
BOA	Some layout/UI praise	**Crashes**, login/password issues, app crashes (27 functionality issues)
Dashen	Loved for visuals (UI/UX)	**Slow performance**, crashes, transaction lag (9 mentions of lag)

Recommendations

Bank	Actionable Fixes
CBE	Improve backend for transaction reliability; add loading indicators
BOA	Overhaul login logic; fix crash-prone features;improve error feedback
Dashen	Optimize performance; maintain strong UI design momentum

Ethical Considerations & Limitations

Bias Type	Observed?	Mitigation Strategy
Review Selection Bias	✅	Used multi-source keywords, not just star ratings
Language Bias	✅	Limited to English reviews, noted as limitation
Skewed Sentiment	✅	Balanced with BERT + simple classifiers

🔎 Note: Reviews often reflect frustration bias. Real-world testing or A/B experiments should supplement insights.

Final Conclusion

This report bridges **data science and business strategy** by converting noisy customer reviews into a roadmap for better mobile banking UX. Through a robust pipeline **from NLP to Oracle to visualization**, I've delivered targeted, data-driven insights that Ethiopian banks can use to **retain users, reduce churn**, and ultimately **enhance trust**.

This is not just a project, it's a **blueprint for customer-centric product thinking**.

Attachments:

- `review_sentiments_themes.csv` (cleaned)
- `oracle_db_handler.py` (DB insert & CRUD)
- `review_handler.ipynb` (interactive notebook)
- Plots: [Sentiment by Bank], [Rating Distribution], [Theme Distribution]