Feedback and Review Intelligence System

Phase-wise and Step-wise Methodology

Kunjan Bhatt

Overview

This document outlines a structured, phase-wise methodology to design and implement a feedback and review intelligence system for a digital wallet application. The system focuses on extracting user reviews, performing sentiment analysis, identifying genuine feedback, highlighting critical issues, and enabling manual resolution tracking.

Phase 0: Problem Definition and Scope

Objective: Clearly define the problem and system boundaries.

- Identify the target wallet application.
- Define review sources:
  - Google Play Store (primary source)
  - Google Reviews (optional or manually uploaded)
- Define system outputs:
  - Sentiment classification: Good, Neutral, Bad
  - Genuine vs Non-valid review detection
  - Identification of major negative issues
  - Manual resolution status tracking

Phase 1: Data Collection

Objective: Obtain review data in a usable format.

Phase 1A: Static Data Loading

- Load reviews from CSV, JSON, or Excel files.
- Ensure mandatory fields:
  - Review ID
  - Review text

  - Rating (1–5)
  - Date
- Validate data integrity and remove duplicates.

Phase 1B: Live Data Extraction (Extension)

- Extract reviews from Google Play Store using scrapers.
- Store extracted reviews in CSV or database.
- Enable manual or scheduled data refresh.

Phase 2: Data Cleaning and Preprocessing

Objective: Prepare raw review text for analysis.

- Convert text to lowercase.
- Remove URLs, emojis, numbers, and special characters.
- Perform tokenization.
- Remove stopwords.
- Apply lemmatization.
- Handle empty, null, or extremely short reviews.

Phase 3: Baseline Sentiment Classification

Objective: Establish a simple working sentiment classifier.

- Classify sentiment using star ratings:
  - 4–5 stars: Good
  - 3 stars: Neutral
  - 1–2 stars: Bad
- Store sentiment labels alongside reviews.
- Analyze sentiment distribution.

Phase 4: Machine Learning Based Sentiment Classification

Objective: Improve sentiment accuracy using machine learning.

- Convert text into numerical representations using TF-IDF.
- Train classification models such as Logistic Regression or SVM.
- Evaluate model performance using accuracy and confusion matrix.
- Optionally fine-tune transformer-based models for higher accuracy.

Phase 5: Genuine vs Non-valid Review Detection

Objective: Filter meaningful feedback from noise.

- Define non-valid reviews:
  - Very short texts
  - Generic praise (e.g., "Good app")
  - Emoji-only responses
- Apply rule-based filtering techniques.
- Optionally train a binary classifier to distinguish genuine feedback.

Phase 6: Negative Issue Extraction

Objective: Identify actionable issues from negative reviews.

- Filter reviews classified as Negative and Genuine.
- Extract keywords and phrases using NLP techniques.
- Group similar complaints using clustering or topic modeling.
- Rank issues based on frequency and severity.

Phase 7: Manual Review and Resolution Tracking

Objective: Introduce human-in-the-loop feedback management.

- Define resolution status labels:
  - Not Valid
  - Working
  - Solved
- Store manual review decisions in a database.
- Link resolution status with each review.

Phase 8: Dashboard and Visualization

Objective: Present insights in an interactive format.

- Develop backend APIs for data access.
- Build a frontend dashboard.
- Visualize:
  - Sentiment distribution
  - Top negative issues
  - Genuine vs non-valid reviews
  - Resolution status tracking

Conclusion

This phased methodology ensures a modular, scalable, and practical approach to building a feedback intelligence system. Each phase is independently testable while contributing toward a comprehensive end-to-end solution.