

Review Analysis and Classification Using Machine Learning Models

Aryan Khaparde
IIT BOMBAY
Email: 21B090010@iitb.ac.in

Abstract—This study analyzes and classifies product reviews based on user ratings and identifies fake reviews provided by non-purchasers. Text-based feature extraction and machine learning models, including LinearSVC and SGDClassifier, were employed. Results indicate a high prevalence of fake reviews and highlight the effectiveness of LinearSVC for sentiment classification. Recommendations for future work are provided.

I. INTRODUCTION

Analyzing product reviews provides valuable insights into user satisfaction and helps businesses identify areas for improvement. However, fake reviews pose a challenge to accurate sentiment analysis. This study focuses on classifying reviews into sentiment classes and identifying fake reviews using machine learning models.

II. DATASET OVERVIEW

A. Data Summary

- **Total Records:** 71,044
- **Features:** 25
- **Missing Values:**
 - Significant null values in columns like *ean*, *reviews.didPurchase*, and *reviews.userCity*.
- **Selected Columns:** *brand*, *manufacturer*, *reviews.didPurchase*, *reviews.rating*, *reviews.text*

III. DATA CLEANING

- Replaced missing values in *reviews.didPurchase* with *Not Available*.
- Dropped rows with null values in other columns.
- Final dataset contains no null values in selected columns.

IV. DATA EXPLORATION

A. Review Ratings Distribution

Ratings were categorized into three sentiment classes:

- **Unhappy:** Ratings 1 and 2
- **Neutral:** Rating 3
- **Happy:** Ratings 4 and 5

A bar plot revealed a significant skew towards *Happy* ratings.

B. Fake Reviews Analysis

A count plot showed a higher number of reviews by non-purchasers, indicating a prevalence of fake reviews.

V. TEXT ANALYSIS

A. Wordcloud

A wordcloud was generated to highlight frequently used terms in reviews, excluding common stopwords.

B. Feature Transformation

- Ratings were mapped to sentiment classes.
- Features were extracted using:
 - **Word-level TF-IDF**
 - **Character-level TF-IDF**
- Combined word and character features into a single dataset.

VI. MODEL DEVELOPMENT

A. Train-Test Split

Data was split into training (70%) and test (30%) sets.

B. LinearSVC

Training Performance:

- Accuracy: 98.87%
- Weighted F1-Score: 99%

Test Performance:

- Accuracy: 90.13%
- Weighted F1-Score: 90%

C. SGDClassifier

Training Performance:

- Accuracy: 93.96%
- Weighted F1-Score: 94%

Test Performance:

- Accuracy: 89.98%
- Weighted F1-Score: 90%

VII. KEY OBSERVATIONS

- Non-purchasers contributed significantly to fake reviews.
- LinearSVC outperformed SGDClassifier.
- Class imbalance was addressed using class weights.

VIII. CONCLUSION

LinearSVC demonstrated superior performance in sentiment classification. The study emphasized the importance of validating reviewer authenticity and proposed future enhancements.