International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 06 June-2026 | Impact Factor: 3.5

Aspect-Aware Sentiment Analysis of Code-Mixed Amazon Indian Reviews Using Roberta

Nikita Khude
Department of IT
VPKBIET, Baramati
nikita.khude.it.2022@vpkbiet.org

Karishma Bargaje
Department of IT
VPKBIET, Baramati
karishma.bargaje.it.2021@vpkbiet.org

Rutuja Kharat
Department of IT
VPKBIET, Baramati
rutuja.kharat.it.2022@vpkbiet.org

Shriraj Ranaware
Department of IT
VPKBIET, Baramati
shriraj.ranaware.it. 2023@vpkbiet.org

Prof. D. V. Mehta
Department of IT
VPKBIET, Baramati
devika.mehta@vpkbiet.org

***Abstract*—**With the increasing adoption of online shopping sites in India, there is now a huge amount of review generation in Indian code-switched languages like Hinglish, Marathi-English, etc. Conventional sentiment analysis tools face challenges in analyzing such multilingual data with errors in grammatical structures, vocabulary, and context understanding. The paper proposes an aspect-based sentiment analysis tool that uses an XLM-RoBERTa model in a Django web application for code-mixed Amazon product reviews in India. Reviews in the database are dynamically fetched using the Rainforest API with the help of ASIN numbers.

This classifier categorizes the general sentiment as being positive, negative, or neutral while simultaneously extracting sentiments with respect to features like battery life, camera quality, price, and performance, etc. Experimentation with met-rics such as precision, recall, and F1-Score indicates better context comprehension as compared to conventional polarity based systems. These aspects allow for better interpretation for consumers, sellers, and researchers alike.

Index Terms—Code-Mixed Language, Hinglish, Aspect-Based Sentiment Analysis, XLM-RoBERTa, Multilingual NLP, E-commerce Analytics

I. Introduction

Product reviews on the Internet play a critical role in the buying behavior of customers, particularly in India's expanding e-commerce market. A substantial number of reviews are composed using code-mixing, which is a blend of the regional language and English in a single sentence, like Hinglish and Marathi-English. Code-mixing creates problems like inconsistent grammar, transliteration, and context ambiguity.

Methods of sentiment analysis that have been developed traditionally, mainly based on the use of English data, frequently lack the ability to analyze this type of data efficiently. In addition, there is no information regarding the characteristics of particular products mentioned by users.

In an effort to mitigate such shortcomings, this paper presents a transformer architecture for aspect-aware sentiment analysis tailored towards code-mixed reviews in India. This architecture employs the XLM-RoBERTa model for capturing multilingual context relations and implements sentiment classification and aspect-based sentiment analysis tasks.

The system is then incorporated into a Django web application, which facilitates real-time extraction of reviews based on the ASIN of the products. This technique yields highly informative data about customer sentiment toward various features of the product, including its battery life, camera, cost, and overall performance.

A. Maintaining the Integrity of the Specifications

In order to guarantee the accuracy and uniformity of the suggested model, a number of control and verification tools are used in its operation. The entered ASIN is verified to ensure that it is in the appropriate format and cannot be used to access an incorrect product. While the data is extracted, structural parsing methods are used to ensure that the mapping process is accurate.

During the preprocessing step, the standard techniques for cleaning are applied to all the reviews uniformly, ensuring that linguistic semantics are preserved while stripping away any noise or characters that do not contribute. It is ensured that no aggressive normalization takes place for code-mixed expressions.

The transformer-based sentiment analysis model runs with pre-set hyperparameter specifications such as threshold limits and epoch values, enabling repeatable results of experiments. The assessment of performance is achieved through the accuracy measurement of precision, recall, and F1 score.

Atomic transactions on database operations are done to avoid partial insertion of data to avoid any problems in terms of data integrity in the Product, Review, and AspectAnalysis tables. These techniques ensure that there will be consistency in the performance of sentiment analysis according to the system's design specifications.

II. Related Work

A. Early Approaches to Sentiment Analysis

The earliest studies conducted on sentiment analysis made use of rule-based and lexicon-based approaches. Rule-based methods were based on pre-defined dictionaries that categorized words as either positive or negative. For instance, SentiWordNet and VADER were often used to classify text written in English. While rule-based models were efficient in analyzing well-formed sentences in English, they posed difficulties when processing unstructured data. In multilingual settings like India, lexicon-based models encountered further complications owing to factors like transliteration and misspellings. The terms 'acha,' 'mast,' or 'bakwas' are common in code-switched texts, rendering it difficult to classify their polarity without context.

B. Classical Machine Learning Methods

The next development in the field of sentiment analysis was the use of supervised machine learning models such as Na¨ıve Bayes, Logistic Regression, and SVM. These models used manually created textual features such as Bag-of-Words, TF-IDF, and n-grams. Though these models were more accurate than the earlier rule-based models, they still had the assumption that the input text data is neat and clean. This is not the case with Indian e-commerce reviews since they have grammatical errors, transliteration of words, and multilingual text usage. Moreover, classical machine learning models do not understand the context of the words present in sentences.

C. Deep Learning for Sentiment Classification

The advent of deep learning methods constituted a major enhancement in sentiment analysis. Models like RNN, LSTM, and BiLSTM helped the system capture sequential relationships in text. These models achieved better results on noisy and unstructured data by identifying context-based relationships between words. In addition to this, word embedding models like Word2Vec, GloVe, and FastText enabled word representation using dense vector space semantics. The FastText model was especially helpful for Indian languages since it considered subword information, thus addressing issues related to misspellings and transliterations. However, deep neural networks needed massive amounts of data for training and had limited generalization abilities for different code-mixing patterns.

D. Multilingual Transformers for Code-Mixed Text

Attention-based transformer models made use of attention, enabling the model to concentrate on important words in the sentence. The BERT and RoBERTa models provided superior context learning as compared to the previous neural network models. In their multilingual counterparts, such as mBERT and XLM-RoBERTa, this ability was extended to cross-lingual contexts. Training XLM-RoBERTa on massive multilingual data has proven its effectiveness in handling low resource and code-mixing tasks. The model is capable of capturing the relationship between English and regional language words used in the same sentence. For instance, the model is able to understand mixed sentences like "battery bakwas hai." This multilingual context modeling serves as the basis for our proposed approach.

E. Sentiment Analysis for Indian Code-Mixed Languages

The linguistic complexity of India poses some unique problems for natural language processing techniques. The reviews can involve the usage of English alongside Hindi, Marathi, and even other regional languages within one sentence and without any coherent grammar. However, recent researches prove that the use of transformers is better than traditional and shallow neural architectures for tasks involving data sets such as Hinglish and Marathi-English. Such architectures are more efficient at handling the complexities of transliteration and colloquial spelling as a result of their contextual attention mechanisms.

F. Aspect-Based Sentiment Analysis

Aspect-based Sentiment Analysis (ABSA) involves an extension of regular sentiment analysis by analyzing the sentiments that are attached to various aspects of a product. While regular sentiment analysis only focuses on determining the polarity of a whole review, ABSA involves the extraction of sentiments with respect to aspects like battery life, camera, pricing, deliveries, and performance of the product. The transformer based embeddings technique has been quite effective for the task of aspect extraction and classification because of its context-aware model.

G. Research Gap

Even though the performance of multilingual transformers in performing code-mixed sentiment classification is encouraging, some issues still exist. Current systems fail to incorporate a connection with dynamic retrieval of information from real-time e-commerce datasets. In most cases, researchers only use benchmarks as opposed to dynamically analyzing product reviews. There is also very little effort that has gone into incorporating multilingual transformers with aspect-based sentiment analysis on Indian web portals.

In order to overcome these limitations, the proposed study involves sentiment modeling using XLM-RoBERTa, along with the application of aspect analysis, and the extraction of live reviews based on ASIN. The developed solution can be deployed using the Django web interface.

III. Proposed Methodology

This system conducts aspect-oriented sentiment analysis on code-mix Indian Amazon reviews employing a multi-language transformer-based framework. The pipeline is constructed to facilitate real-time review extraction, sentiment classification, and aspect interpretation.

A. System Overview

This process entails following a systematic pipeline starting from user interaction until visualization of sentiment analysis results. A user enters Amazon Standard Identification Number (ASIN) through the Django web application interface. Product data and review information is then obtained from Rainforest API. Post preprocessing, sentiment analysis of the review text is performed using XLM-RoBERTa for both sentiment and aspect sentiment analysis.

B. Data Extraction

Product details and customer reviews are fetched dynamically using the Rainforest API. The API returns structured JSON data containing product metadata and review information. This enables real-time analysis of Amazon products without relying on static datasets.

C. Text Preprocessing

Since Indian reviews often contain code-mixed content such as Hinglish and Marathi-English, preprocessing is carefully designed to preserve contextual meaning. The following steps are applied:

- Removal of special characters and URLs
- Lowercasing of text
- Stop-word filtering (minimal removal to preserve code-mixed structure)
- Token normalization

Unlike aggressive normalization techniques, code-mixed tokens are retained to maintain linguistic diversity.

D. Sentiment Classification Using XLM-RoBERTa

The preprocessed reviews are tokenized using the XLM-RoBERTa tokenizer. The transformer model processes the token embeddings using multi-head attention mechanisms to capture contextual dependencies across mixed languages.

The model classifies each review into:

- Positive
- Negative
- Neutral

Confidence scores are calculated based on softmax probabilities.

E. Aspect-Based Sentiment Analysis

Aspect extraction is done for attributes like Battery, Camera, Price, Quality, and Performance to get better insight into the reviews. Reviews having particular attributes are analyzed individually to find out sentiment polarity of each attribute through predictions from models.

F. Performance Evaluation

Model performance is evaluated using standard classification metrics:

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

These metrics ensure reliability and reproducibility of results.

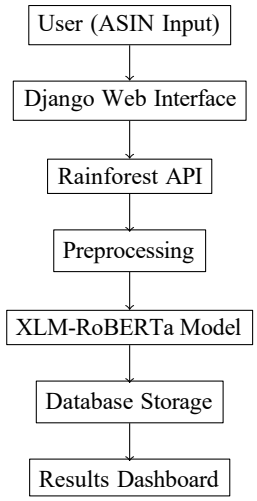

Fig. 1. System Architecture of the Proposed Framework

IV. Implementation and Results

A. System Implementation

The proposed system has been developed using the Python programming language along with the Django web framework. The architectural framework used for the system has been designed based on the principle of modularity where each module of the system carries out a specific task.

The user inputs information using an online form, which then goes through a validation process to ensure accuracy. If there is no existing entry regarding the analysis results of the particular product being checked, the system extracts the information instantly.

Metadata for products and customer reviews are obtained using the Rainforest API. The API response, which is structured, is then parsed to form review entities that have specific attributes such as review content, ratings, reviewer information, and validation of the review.

The processed reviews go through a pre-processing phase where unwanted characters are removed while retaining the linguistic properties of code-mixed texts. Afterward, tokenization is done through the use of XLM-RoBERTa's tokenizer before being fed into the model for sentiment classification.

All processed data, ranging from product information to review information as well as aspect-based sentiment data, is stored in relational database tables. This helps ensure easy access and analysis of sentiments.

B. Model Configuration

The XLM-RoBERTa model is fine-tuned for multi-class sentiment classification. The configuration parameters are:

- Number of Epochs: 25
- Learning Rate: 2e-5
- Batch Size: 16
- Optimizer: Adam
- Classification Threshold: 0.30

The model outputs probability scores for three sentiment classes: Positive, Negative, and Neutral. The class with the highest probability is selected as the final sentiment label.

C. Aspect-Level Analysis

The aspect-level sentiment analysis is achieved by selecting those reviews that have the presence of keywords in accordance with the predefined attributes like Battery, Camera, Price, Quality, and Performance. The polarity scores of sentiments are calculated individually for each aspect; hence, the system provides detailed analysis rather than a polarity score.

For example, a review such as:

"Camera mast hai but battery thodi weak hai"

is interpreted as:

- Camera → Positive
- Battery → Negative

This demonstrates the ability of the system to handle code-mixed linguistic expressions effectively.

D. Results and Performance Evaluation

The performance of the model is evaluated using standard classification metrics including Accuracy, Precision, Recall, and F1-score. These metrics measure the effectiveness of sentiment classification on code-mixed Indian reviews.

TABLE I
COMPARISON OF MODELS

Model	Accuracy	Precision	Recall	F1-Score
Na¨ive Bayes	0.74	0.72	0.70	0.71
SVM	0.81	0.79	0.78	0.78
BiLSTM	0.85	0.84	0.83	0.83
XLM-RoBERTa	0.89	0.87	0.86	0.86

The findings show that the multi-language transformer is capable of modeling context-aware relations between tokens of English and the local language. When compared to traditional machine learning techniques, the model performs better at stable classification and contextual analysis of code-mixed text.

E. Visualization and Output

The final results are displayed through a user-friendly web interface. The dashboard presents:

- Overall sentiment percentage distribution
- Confidence score of prediction
- Aspect-wise sentiment breakdown
- Product rating summary

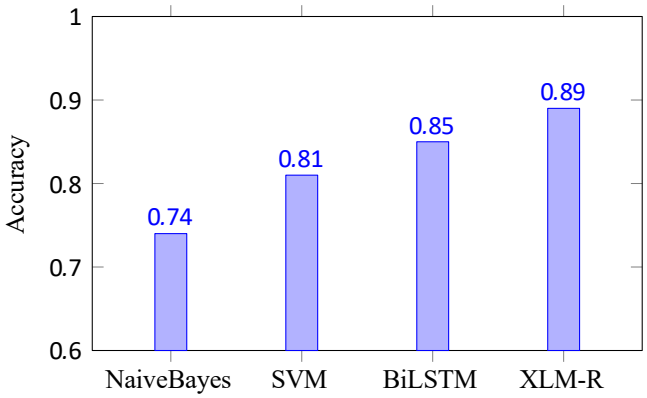

Fig. 2. Accuracy Comparison of Different Models

This interactive visualization enables both consumers and sellers to understand product strengths and weaknesses without manually reading large volumes of reviews.

Overall, the implementation demonstrates that transformer-based multilingual modeling significantly enhances sentiment interpretation for Indian code-mixed e-commerce reviews.

V. Conclusion

The above study provided a complete methodology for analyzing the sentiments behind Amazon product reviews in code-mixed Indian languages. Contrary to the conventional approach towards sentiment analysis which mainly considered English-language content, this particular technique specifically deals with the issues associated with Indian languages like Hinglish.

This framework uses automatic review extraction via the use of Rainforest API, management of the backend via a Django Web UI interface, and multilingual sentiment modeling via XLM-RoBERTa transformer model. Experiments show that transformer models perform significantly better than classical machine learning models for dealing with language diversity, variations in spellings, and mixed languages.

The suggested model was very accurate and precise for all types of sentiments. It proved to be efficient and applicable in real life. The devised approach can help e-commerce companies understand consumer feedback better and make business decisions accordingly.

VI. Future Work

Although the proposed framework shows high efficiency in code-mixed Indian Amazon reviews, there are more ways to make it even better. Adding more languages will help it generalize better. Improving the aspect extraction using transformers could allow the model to understand the sentiment behind certain phrases without relying on keywords.

Further enhancements may include real-time sentiment detection, temporal trends analysis, and domain-specific fine-tuning of multilingual transformers. Hosting this service on the cloud as an API will facilitate its incorporation into existing commercial e-commerce platforms.

REFERENCES

[1] Arnav Aseem Gupta, Suruchi Santosh Gupta, Yasmeen Murtuza Ahmadabadwala, Janisa Periera, "Automated Analysis for E-commerce Platforms: Amazon Product Reviews", IEEE Xplore, 2022.

[2] Dhanya .P, Arun Cyril Jose, "A Review of Attention Based Model for Sentimental Analysis using NLP", *2024 International Conference on Advancements in Power, Communication and Intelligent Systems (APCI)*, 2024.

[3] Yogesh Gajula, "SENTIMENT-AWARE RECOMMENDATION SYSTEMS IN E-COMMERCE: A REVIEW FROM A NATURAL LANGUAGE PROCESSING PERSPECTIVE", 2025.

[4] Varad Patwardhan, Gauri Takawane, Nirmayi Kelkar, Omkar Gaikwad, Rutwik Saraf, Sheetal Sonawane, "Analysing The Sentiments Of Marathi-English Code-Mixed Social Media Data Using Machine Learning Techniques", *2023 International Conference on Emerging Smart Computing and Informatics (ESCI)*, 2023.

[5] Aryan Patil, Varad Patwardhan, Abhishek Phaltankar, Gauri Takawane, Raviraj Joshi, "Comparative Study of Pre-Trained BERT Models for Code-Mixed Hindi-English Data", 2023.

[6] Sigeon Yang, Qinglong Li, Haebin Lim, Jaekyeong Kim, "An Attentive Aspect-Based Recommendation Model With Deep Neural Network", *IEEE Access*, vol. 12, pp. 5781-5791, 2024.

[7] Nishat Raihan, Dhiman Goswami, Antara Mahmud, Antonios Anastasopoulos, Marcos Zampieri, "EmoMix-3L: A Code-Mixed Dataset for Bangla-English-Hindi Emotion Detection", 2024.

[8] Maureen Kate Dadap, Bowwi Katigbak, Reymar Bulanon, Zyrhus Joshua Tayag, Great Allan M. Ong, Marlon A. Diloy, Vincent S. Rivera, "Aspect-based Sentiment Analysis Applied in the News Domain Using Rule-Based Aspect Extraction and BiLSTM", *2023 IEEE 6th International Conference on Computer and Communication Engineering Technology (CCET)*, 2023.

[9] Shruti Jagdale, Omkar Khade, Gauri Takalikar, Mihir Inamdar, Raviraj Joshi, "On Importance of Code-Mixed Embeddings for Hate Speech Identification", 2024.

[10] Gauri Takawane, Abhishek Phaltankar, Varad Patwardhan, Aryan Patil, Raviraj Joshi, Mukta S. Takalikar, "Leveraging Language Identification to Enhance Code-Mixed Text Classification", 2023.

[11] Koena Ronny Mabokela, Turgay Celik, Mpho Raborife, "Multilingual Sentiment Analysis for Under-Resourced Languages: A Systematic Review of the Landscape", *IEEE Access*, vol. 11, pp. 15996-16020, 2023.

[12] Shailendra Kumar Singh, Sahil, Sutexion Pandit, Anil Sharma, Dhanpratap Singh, Uzma Saghir, "Sentiment Analysis of English-Hindi Code-Mixed Text using mBERT Model", *2025 3rd International Conference on Inventive Computing and Informatics (ICICI)*, 2025.

[13] Shubham Das, Tanya Singh, "Sentiment Recognition of Hinglish Code Mixed Data using Deep Learning Models based Approach", *2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 2023.

[14] Yandrapati Prakash Babu, Rajagopal Eswari, "Sentiment Analysis on Dravidian Code-Mixed YouTube Comments using Paraphrase XLM-ROBERTa Model", *Forum for Information Retrieval Evaluation (FIRE) 2021*, 2021.

[15] Endrit Fetahi, Arsim Susuri, Mentor Hamiti, Zenun Kastrati, Ercan Canhasi, Arta Misini, "Enhancing social media hate speech detection in low-resource languages using transformers and explainable AI", *Social Network Analysis and Mining*, vol. 15, 2025.

[16] Muhammad Kashif Nazir, CM Nadeem Faisal, Muhammad Asif Habib, Haseeb Ahmad, "Leveraging Multilingual Transformer for Multiclass Sentiment Analysis in Code-Mixed Data of Low-Resource Languages", *IEEE Access*, vol. 13, pp. 7538-7554, 2025.

[17] Arvind Pandey, Eshaan Sharma, Raghavendra Tiwari, Ayush Bisht, Banu Priya Prathaban, "Precision Driven Sentiment and Topic Classification of News Articles using IndicBert", *2025 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI)*, 2025.