VFAST Transactions on Software Engineering
https://vfast.org/journals/index.php/VTSE@ 2025, ISSN(e): 2309-3978, ISSN(p): 2411-6246
Volume 13, Number 1, January-March 2025 pp: 131-140

Sentiment Analysis of Multilingual Roman Text for E-Commerce Reviews using Machine Learning Approaches

Sana Riaz[1], **Sarfraz Natha**[1, 2*], **Asghar Ali Chandio**[3], **Mehwish Leghari**[4], **Abeer Javed Syed**[5]

[1]Department of Information Technology, Quaid e Awam University of Engineering, Science & Technology, Nawabshah, Pakistan; [2*]Department of Software Engineering, Sir Syed University of Engineering & Technology, Karachi, Pakistan; [3]School of Engineering and Information Technology, University of New South Wales, Canberra, Australia; [3]Department of Artificial Intelligence, Quaid-e-Awam University of Engineering, Science & Technology, Pakistan; [4]Department of Data Science, Quaid-e-Awam University of Engineering, Science & Technology, Pakistan; [5]Department of Computer Science, IQRA University, Pakistan.

Keywords: *Sentiment Analysis, Multilingual Roman Text Reviews, Product Reviews Sentiment Analysis, Sentiment Analysis using Machine Learning.*

Journal Info:
Submitted:
February 2, 2025
Accepted:
March 16, 2025
Published:
March 28, 2025

Abstract Sentiment analysis, a type of natural language processing (NLP) analyzes the text data to extract and identify subjective information including attitudes, opinions, and feelings. Sentiment analysis can be used to examine audience feedback and reviews in the context of multilingual product reviews. In this paper, a sentiment analysis model using machine learning approaches has been developed for multilingual product reviews in Roman Urdu or Sindhi to determine how the public feels about certain posts, products, etc. The importance of sentiment analysis for product context reviews in many languages in Roman is multifaceted. It can offer insightful information on the preferences of the likes and dislikes of the audience. To accomplish multilingual sentiment analysis, a dataset of reviews in Roman Urdu and Sindhi languages from diverse online platforms and social media sources like YouTube, Facebook, TikTok, Daraz, and Instagram was collected. To identify pertinent features essential for categorizing reviews into negative, positive, or neutral sentiments based on polarity, the Term Frequency Inverse Document Frequency (TF-IDF) method was used. For classification, five different machine learning classifiers including Linear Regression (LR), Naive Bayes (NB), Support Vector Machine (SVM), Random Forest (RF), and K-nearest neighbors (KNN) were used. The classification results were measured in terms of precision score, recall score, and F1-score. With TF-IDF, the SVM, and LR outperformed than other classifiers and obtained an F1-score of 0.77%, and 0.78%. To further improve the classification accuracy, the Synthetic Minority Over-sampling TEchnique (SMOTE) was used to manage the class imbalance problem. With SMOTE, the classification accuracy of LR and SVM was improved to 0.79% and 0.80%.

***Correspondence author email address:** sasattar@ssuet.edu.pk
DOI: 10.21015/vtse.v13i1.2067

1 Introduction

Opinion prospecting, another name for Sentiment analysis, has gained more attraction and importance in the last few years [1, 2]. This field stands out for the potential benefits it offers as well as its increasing popularity. This expansion has been made possible largely by technological advancements and the growth of the internet. As a result, there is now much more data that is easily accessible for analysis, which presents both new opportunities and challenges [3]. As social media, online reviews, and other digital communication channels gain popularity, a vast amount of data reflecting the opinions and attitudes of people is generated [4]. This massive volume of data offers vital insights into consumer preferences, industry trends, and societal viewpoints. Businesses and organizations can use sentiment analysis to better understand their customers, create better products and services, and make informed business decisions. The main advantage of sentiment analysis and psychoanalysis is the ability to draw meaningful insights from important unstructured data. Traditional methods of manually assessing moods and emotions are laborious and ineffective. But, with machine learning and NLP models, it is now possible to automate the search process and extract valuable information from large volumes of data [5].

Opinion prospecting and sentiment analysis have numerous applications in various fields. Businesses may use customer reviews and social media posts to gauge the success of their marketing campaigns or products. They can use this data to identify areas for improvement or to modify their marketing strategies as necessary. Sentiment analysis can be applied in the financial sector to measure market sentiment and make informed investment decisions, just as it can be in other industries. In social and political contexts, sentiment analysis and opinion prospecting are also crucial. By looking at what people are saying on social media or in online forums, researchers and decision-makers can find out more about the perceptions of the public on specific issues, spot trends in public opinion, and monitor the effectiveness of initiatives or programs [6]. Precise analysis of sentiment is challenging because language is complex and subjectivity varies among individuals. Context, metaphor, and cultural codes all complicate the analysis.

Over the last few years, there has been a remarkable increase in the number of Roman Urdu or Sindhi users using social media platforms such as Facebook, Twitter, and other online social networking sites [7, 8]. The Roman Urdu and Sindhi languages have become much more prevalent on social media. Users can now communicate with others who share their linguistic background and effortlessly express their thoughts and feelings in the Roman Urdu or Sindhi languages. The increased use of Roman Urdu or Sindhi on social media has many consequences. Roman Urdu or Sindhi users can interact and communicate with each other more easily as a result, which strengthens their sense of linguistic identity and enables them to express themselves more openly on a range of subjects. Furthermore, The widespread use of Roman Urdu or Sindhi in social media has helped spread digital content in that language. Users now have a place to exchange knowledge, narratives, viewpoints, and original works of art in Roman Urdu or Sindhi [9].

Mostly research on sentiment analysis has been done on specific languages, including Chinese, Arabic, Urdu, English, and others [10]. To conduct sentiment analysis with machine learning, this research paper presented a multi-class, multilingual Roman text dataset based on user reviews. The dataset has been compiled from a variety of online sources, including sports, politics, entertainment, Instagram, TikTok and Facebook. The suggested dataset includes reviews that have been manually classified by the expert people into three categories: negative, neutral, and positive. The reviews collected were pre-processed using the natural language toolkit (NLTK), and the Pandas libraries. To extract the useful features from the reviews, a TF-IDF method was applied. Five different machine learning classifiers were used to classify the sentiment into negative, positive, or neutral polarity. To avoid the problem of model overfitting, a SMOTE was applied. The results of both SMOTE and without SMOTE methods are analyzed and compared.

The remainder of the paper is arranged as follows: Section II briefly gives overview on the review from some of the relevant research work, Section III highlights the multilingual Roman dataset collected for sentiment analysis, experimental setup and the results achieved by performing different experiments are discussed in section IV, while section V describes the conclusion of the paper with some future research direction discussed.

2 Related Work

Sentiment analysis focuses to detect, extract, and examine the attitudes, sentiments, and feelings expressed in textual data. Numerous industries, such as market research, social media analysis, and customer feedback analysis, have expressed a great deal of interest in it. Whereas sentiment analysis has been extensively studied for individual languages such as Roman Urdu, English, or Roman Sindhi. Research on the social media data has now become a major focus in sentiment analysis. Sentiment analysis is now more significant in several fields, such as politicians, cultural institutions, filmmakers, advertisers, and other organizations can all benefit from understanding the opinions and feelings of their target audiences.

Zhaoxia Wang et al. [11] proposed MiMuSA, a new interpretable elegant multi-class sentiment analysis method that mimics natural language understanding systems. The proposed approach addressed ambivalence and sentiment competence through a multi-layered modular system that mimicked human speech perception systems. Several knowledge bases were formed to support the sentiment understanding process, such as the sentiment power knowledge base, metaphorical rules and negative knowledge bases, negation and basic knowledge bases, and basic suggestion bases. Selvi and Sreeja [12] applied machine learning techniques for sentiment analysis in Tamil language to identify emotions expressed in a text and classified them as positive, negative, or neutral. They applied and compared SVM, k-NN (k-Nearest Neighbor), and LSTM (Long Short-Term Memory) algorithms. These algorithms were tested using a data set of captions in Tamil from YouTube searches. The research effort collected Tamil text content from YouTube searches for testing purposes. Classifying sensations was made simple by applying the chosen ideal model to new Tamil text data.

Keinan and HaCohen-Kerner [13] created a multilingual dataset of African languages for sentiment analysis. Different models were trained for twelve African languages, and a multilingual dataset consisting of these twelve languages was given a thirteenth model. Four traditional machine learning algorithms, two deep learning techniques, three oversampling techniques based on the TF-IDF feature values, and a range of word and character n-grams were applied to classify the reviews into positive, negative, and neutral categories. The hyper-parameters of each model were tuned. Shah et al., [14] applied multinomial naive Bayes and KNN classifiers for sentiment analysis on the movie reviews dataset. Furthermore, they investigated how two widely used text representation algorithms, the term level counting vectorizer and TF-IDF influence sentiment analysis, especially in movie analysis. Finding the most effective text representation and classifier combination for sentiment analysis of movie reviews was the major contribution of their research. The results shown could impact market research, audience perception studies, and corporate decision-making since they provided important new understandings of the best ways to accurately measure and interpret the emotions depicted in movie reviews. Mahmoud Al-Ayyoub et al., [15] developed a sentiment analysis method for the Arabic language. As the Arabic language is complex in its nature, very limited research work has been performed on it. To fill up the gaps in the literature and set the stage for future research in this area, they reviewed extensive work on Arabic sentiment analysis (ASA), including methods, tools, and techniques that can assist ASA researchers. Whether or not they are publicly available, Currently used sentiment resources for ASA are outlined in this review article. According to their study, most articles provided answers to problems faced by ASA researchers. Both corpus-based and text-based SA methods were also discussed in the review.

Qureshi et al., [16] developed a Roman Urdu Sentiment Analysis for the songs' reviews. Languages with a lot of resources have been the subject of most sentiment analysis research in the past, while languages with few resources have received comparatively little attention. The provision of linguistic resources for Roman Urdu processing has not received significant attention. They conducted sentiment analysis on Urdu, a language with poor source quality written in Roman characters. The dataset used for this research was created by the authors which consisted of song comments collected from different online sources. Three tracks from YouTube were selected by the sub-continent music industry. Following the pre-processing stage, Naïve Bayes, KNN, Decision Tree (ID3), and ANN were used for the analysis of the Roman Urdu reviews. Compared to the other classifiers, Naïve Bayes performed better and achieved an accuracy rate of 82.41%. Jawad et al., [17] identified the implicit opinionated text polarity and extracted subjective statements from Roman Urdu from different online platforms. Google Maps, the manual effort, and the e-commerce portal Daraz were the main sources of the dataset collection. Two datasets of Roman Urdu were the two major contributions of the research. The integrated modules take in user input, recognize textual elements, fix spelling errors, classify emotions, and return the orientation of the input sentences of emotional intensity scores. The proposed system gradually became stronger with every input experience. The findings demonstrated that on a close domain dataset, the language detector provided an accuracy of 97.1%. The overall sensitivity classification accuracy of the proposed system designed was 94.3%.

Hammad and Anwar [18] developed a Sindhi Sentiment Analysis on Twitter data using supervised machine learning techniques. Millions of people use social media sites every day, making it an important platform for exchanging opinions. With so many languages, people are increasingly expressing their opinions in their native language on social media. Opinions of people on the internet are of interest to individuals and companies for their future business and product initiatives. To analyze the Twitter tweets in the Sindhi language, different classifiers were applied in order to analyze the emotions of Sindhi language. A dataset of Sindhi tweets we created. Only two classes with positive and negative labels were used to automatically classify the tweets. After pre-processing, the text was tokenized, and non-Sindhi terms, unnecessary white spaces, and punctuations were removed. The Sindhi tweets dataset was subjected to supervised machine learning algorithms following data cleaning and tagging. For classification, KNN, NB, Decision Tree (DT), and SVM were used. The results obtained showed that SVM and decision trees provided the highest accuracy on the Sindhi Twitter dataset, followed by KNN. Noor et al., [19] developed a Sentiment Analysis method om E-Commerce Applications for Roman Urdu text using SVM. A dataset on the product reviews was collected from the Daraz website. Daraz is one of the most popular and frequently visited e-commerce websites in Pakistan. In all 20286K reviews were annotated by three distinct specialists into three classifications such as positive, negative, and neutral. The bag of words model, also known as the vector space model, was used to extract features, which were then fed to SVM for sentiment classification. A detailed review of the multilingual sentiment analysis, challenges, datasets, tools, applications, and future directions of sentiment analysis is given in [20].

3 Research Methodology

The process of analyzing the sentiment of Roman Urdu and Sindhi language reviews involves many steps. The research methodology applied is illustrated in Figure 1

3.1 Data Collection

The dataset of E-Commerce reviews in Roman Urdu and Sindhi language was created by collecting the sentiments from diverse online platforms and social media sources, including YouTube, Facebook, TikTok, and Daraz. A web crawler or scraper tool was applied to gather reviews from the specified online platforms. Furthermore, some reviews were manually collected as well. Table 1 shows the statistics of the dataset. A total of 2498 reviews were collected from the manual and online sources. The polarity assigned to each review was verified by the Urdu and Sindhi language experts. Figure 2 illustrates some samples of the

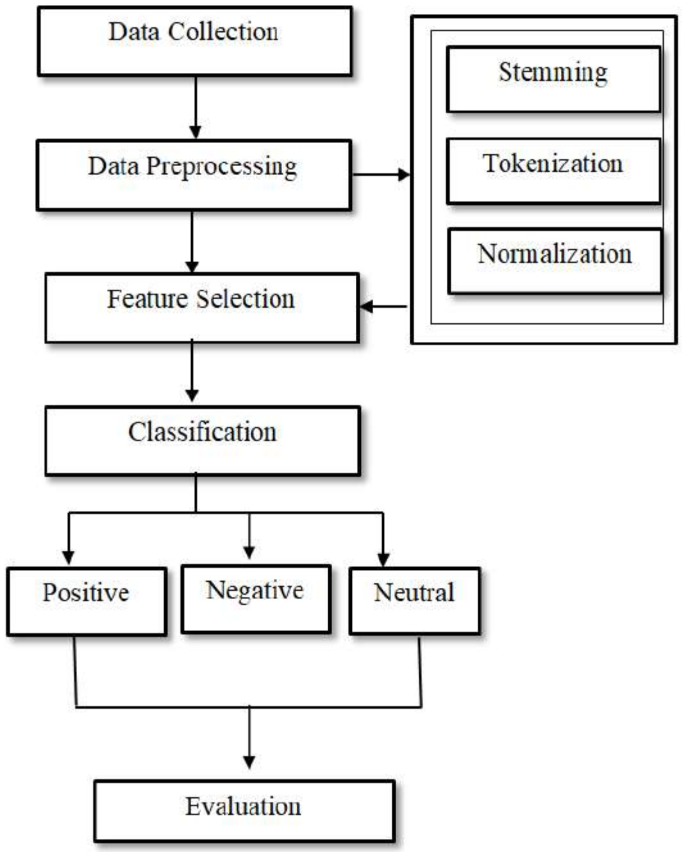

Figure 1. A machine learning-based methodology used for Roman Urdu and Sindhi Sentiment Analysis

Roman Urdu and Sindhi sentiment analysis dataset collected from different online sources.

Table 1. Statistics of Multilingual Roman Urdu and Sindhi Datast

Source	Data Samples
Daraz	1760
YouTube	298
Facebook	180
TikTok	90
Manual	170

3.2 Pre-processing

The Roman Urdu and Sindhi reviews were preprocessed using the appropriate techniques, such as stemming, tokenization, and normalization, to eliminate noise, stop words, and other unnecessary information. The preprocessing was performed using NLTK library in Python. Handling Roman Urdu and Sindhi language characters is challenging, as a variation of a single letter completely changes the context and the meaning of a sentiment. The spelling variations of the same word also make the classification problem more complex. A stemming technique was performed to return the review words to their original or base form. In stemming, for example, "going," "goes," and "go," become the common root word "go." This process helped to reduce the dimensionality of textual data and improved the accuracy and efficiency of language-based tasks such as searching, sentiment analysis, and information retrieval. A tokenization method was applied to break down sentences into smaller discrete units, usually words or phrases. These tokens are the fundamental building blocks for sentiment analysis systems. A normalization technique was used

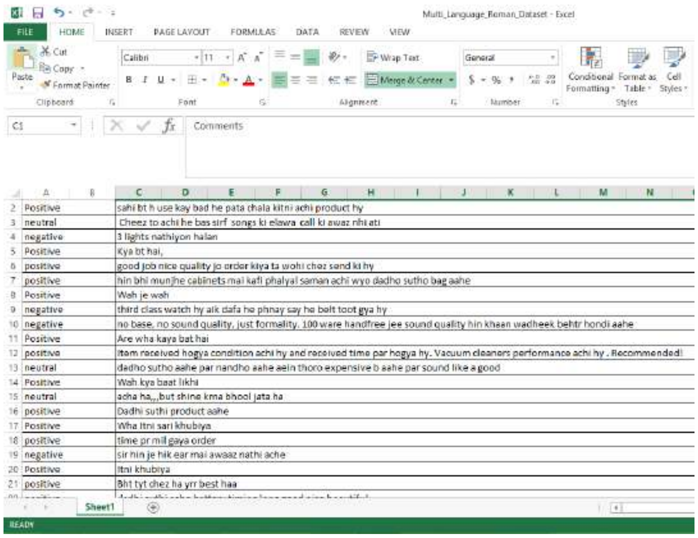

Figure 2. Some samples of Roman Urdu and Sindhi Sentiment Analysis Dataset

to format text consistently. It entailed changing all characters to lowercase, removing punctuation, and possibly addressing other text-specific variations.

3.3 Feature Selection

For appropriate feature selection, a TF-IDF technique was used to assist in categorizing the polarity of the reviews. TF-IDF is a number that indicates the weight of a word in a document or group of documents. It is obtained by measuring the frequency with which a word occurs in a document (Term Frequency) and its rarity in all documents (Inverse Document Frequency). Words that occur frequently in a given document but infrequently in the corpus as a whole are given greater weight by TF-IDF. This technique is widely used to highlight words that are especially relevant to a given paper and minimize terms that are commonly used [21, 22]. The term frequency is computed as:

TF(t) = Number of times term t appears in a d / Total number of terms in a d

where t is the term, and d is the document. While the inverse document frequency is computed as:

$IDF = log\ N / DF_t$

where N is the total documents in the dataset, and DF_t is the number of documents containing the term t and t is any term / word in the document.

The Roman Urdu and Sindhi reviews collected for each sentiment were not equal, which created the problem of class imbalance. To avoid the problem of class imbalance and model overfitting, an oversampling technique was applied. Oversampling and undersampling are the techniques used in data analysis and machine learning, especially in dealing with imbalanced datasets. The oversampling is a technique where synthetic samples are generated for the minority classes. This reduces the problem of class imbalance. The undersampling technique involves reducing the dominant class to balance the dataset. By doing this, the model can avoid biases in favor of the majority classes, respectively.

3.4 Classification

The reviews of Roman Urdu and Sindhi were categorized into positive, negative, and neutral groups using different machine learning techniques including NB, SVM, LR, KNN, and RFC. The results of each classifier were obtained and compared. Furthermore, SMOTE technique was used with each classifier to reduce the class imbalance problem.

3.5 Evaluation

The multilingual Roman Urdu and Sindhi sentiment analysis was evaluated using relevant metrics like F1 score, precision, recall, and confusion matrix. The precision determines the proportion of real positive cases—including real and fake positives—to all cases that were predicted to be positive. A high accuracy level means that the model regularly generates accurate predictions. The precision is calculated as:

$Precision = TP / TP + FP$

where TP corresponds to true positive, and FP corresponds to false positives.

Recall is used to assess how accurate the model will be in making future predictions. It determines the proportion of real positive cases—including real and fake positives—to all cases that were predicted to be positive. It is calculated as:

$Recall = TP / TP + FN$

where FB corresponds to false negative.

F1 score also known as the harmonic mean of recall and precision balances recall and precision, making it suitable in situations where we want to account for both false positives and false negatives. Using the F1 score is very beneficial when dealing with imbalanced datasets. It is computed as:

*F-Score = 2 * Precision * Recall / Precision + Recall*

The confusion matrix is used to evaluate the effectiveness of the classification scheme. It compares the actual class labels to the expected class labels to see true positives, true negatives, false positives, and false negatives.

4 Results and Discussions

To evaluate the performance of each classifier, a dataset of multilingual Roman Urdu and Sindhi sentiments was collected from different online sources including, YouTube, Facebook, Instagram, and Blogs. A total of 2498 sentiments were collected. Figure 3 illustrates the dataset statistics.

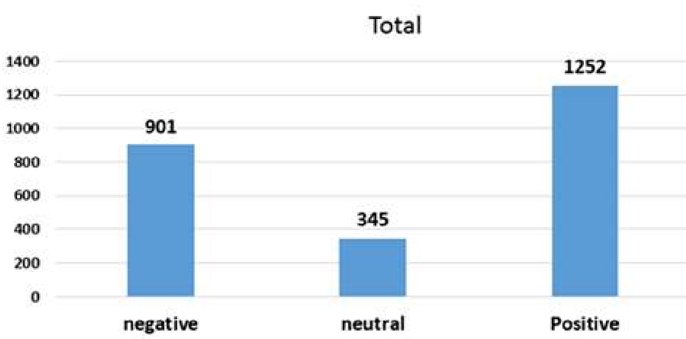

Figure 3. Statistics of multilingual Roman Urdu and Sindhi Sentiment Analysis Dataset.

As illustrated in Figure 3, the dataset of multilingual Roman Urdu and Sindhi sentiment analysis for E-Commerce products contains imbalanced sentiments.

Table 2 presents the classification results of multilingual Roman Urdu and Sindhi Sentiment Analysis in terms of precision, recall, f1 scores, and accuracy. Five different supervised machine learning classifiers including LR, NB, SVM, KNN, and RFC were trained and evaluated on the collected data. The TF-IDF method was used for feature selection.

The precision value obtained with LR classifier is 0.78 while it is 0.71, 0.77, 0.44, and 0.76 with the NB, SVM, KNN, and RFC classifiers. The recall achieved by SVM classifier is 0.79 while the LR, NB, KNN, and RFC classifiers yielded a recall value of 0.78, 0.71, 0.47, and 0.77, respectively. Similarly, the LR and SVM classifiers achieved the f1-score of 0.78 while the NB, KNN, and RFC classifiers achieved the f1-score of 0.72, 0.47, and 0.76, respectively. The overall testing accuracy for the multilingual Roman Urdu and Sindhi sentiment analysis by LR, NB, SVM, KNN, and RFC is 78%, 71%, 77%, 46%, and 76%, respectively.

Table 2. Multilingual Roman Urdu and Sindhi Sentiment Analysis for E-Commerce Results with Different Machine Learning Classifiers

Classifier	Precision	Recall	F-Score	Accuracy
LR	0.78	0.78	0.78	0.78
NB	0.71	0.73	0.72	0.71
SVM	0.77	0.79	0.78	0.77
KNN	0.44	0.47	0.47	0.46
RFC	0.76	0.77	0.76	0.76

The results presented in Table 2 demonstrate that the LR classifier outperformed on the multilingual Roman Urdu and Sindhi sentiment analysis data, while KNN was a poorly performing classifier. The overall test accuracy is not more than 78%. This shows the complexity and challenges of the Roman Urdu and Sindhi sentiment data. The sentiments in the dataset are high dimensional, therefore, the KNN performed poorly. The confusion matrix using all five classifiers for multilingual Roman Urdu and Sindhi sentiment analysis is presented in Figure 4. As illustrated in the confusion matrix of all five classifiers, some samples have been negatively classified in other classes. This is due to very minimum differences in different sentiments.

As illustrated in Figure 3, the dataset contains an unequal number of samples in each class. The negative class has 901 samples, the neutral class contains 345 samples, and the positive class contains 1252 samples. To handle the problem of class imbalance, a SMOTE over-sampling technique was used to synthetically increase the samples of minority classes. All five classifiers were trained on the multilingual Roman Urdu and Sindhi dataset of E-Commerce product reviews after applying SMOTE technique. Table 3

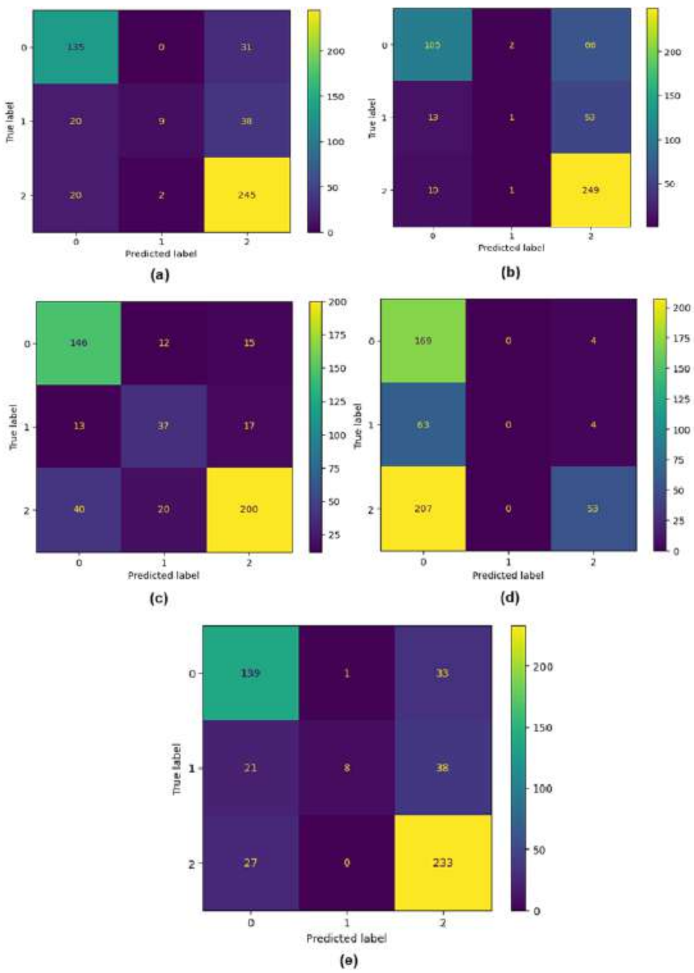

Figure 4. Confusion matrix for multilingual Roman Urdu and Sindhi Sentiment Analysis: (a) Confusion Matrix with LR (b) Confusion Matrix with NB (c) Confusion Matrix with SVM (d) Confusion Matrix with KNN, and (e) Confusion Matrix with RFC classifiers

demonstrates the classification results obtained after applying SMOTE technique.

Table 3. Multilingual Roman Urdu and Sindhi Sentiment Analysis for E-Commerce Results after SMOTE technique

Classifier	Precision	Recall	F-Score	Accuracy
LR	0.79	0.79	0.79	0.79
NB	0.75	0.77	0.76	0.76
SVM	0.78	0.79	0.78	0.78
KNN	0.52	0.55	0.54	0.54
RFC	0.79	0.79	0.79	0.79

Table 3 demonstrates that all the precision, recall, f1-score, and overall accuracy of each classifier were slightly improved when using SMOTE technique to avoid the problem of class imbalance.

4.1 Comparison with E-Commerce Sentiment Analysis Methods

The multilingual Roman Urdu and Sindhi sentiment analysis performance on the reviews of E-Commerce data using machine learning classifiers was compared with the existing multilanguage sentiment analysis systems. The results of multilingual Roman sentiments have been compared in terms of commonly used classification metrics. Table 4 compares the machine learning classifiers applied to the multilingual Roman Urdu and Sindhi E-Commerce sentiment analysis dataset with existing machine learning-based methods.

Table 4. Comparison of Machine Learning Classifiers on the Multilingual Roman Urdu and Sindhi E-Commerce Dataset with Existing Machine Learning Methods

Model	Precision	Recall	F-Score	Accuracy
Das [23]	0.74	0.74	0.74	0.77
Das [23]	0.71	0.71	0.71	0.74
Savci [24]	0.79	0.85	0.82	0.77
Chen [25]	–	–	0.74	0.75
Chandio [26]	0.68	0.71	0.69	0.68
Bilal [27]	0.70	0.37	0.48	0.47
Shah [28]	0.67	0.67	0.67	0.67
Fang [29]	0.63	0.61	0.62	0.5
Proposed	0.78	0.78	0.78	0.78

5 Conclusions and Future Work

This paper examines the practical implications of the research findings, focusing on how sentiment analysis in the Roman Urdu and Sindhi languages can provide a platform for consumer feedback research, market research, and social media management. It focuses on the potential value of applying sentiment analysis to content, marketing, and audience engagement decision-making processes. A dataset of multilingual Roman Urdu and Sindhi E-Commerce sentiments was collected from different online sources, including

YouTube, Facebook, and other blogs. Preprocessing techniques using NLTK library were applied to normalize the data and remove the stop words, punctuations, and other symbols. TF-IDF was used to select the most appropriate features. Five different supervised machine learning classifiers were trained and tested on the collected dataset. To reduce the class imbalance problem, a SMOTE technique was utilized. Among all classifiers, the LR classifier obtained the highest accuracy.

In the future, more multilingual data can be collected from diverse sources. Deep learning methods, such as recurrent neural networks and their variants, can be trained and evaluated on the Roman Urdu and Sindhi E-Commerce datasets. Furthermore, the efficiency of the transformer models, such as BERT can be analyzed on the Roman Urdu and Sindhi E-Commerce datasets.

Author Contributions

Sana Riaz: Conceptualisation, data collection, and Writing the original draft **Sarfraz Natha**: Data labelling, methodology design, and proofreading **Asghar Ali Chandio**: Data collection, experimental work, and writing the original draft. **Mehwish Leghari**: data collection, Algorithm testing, labelling, and paper draft writing. **Abeer Javed Syed**: Algorithm testing, validation, supervision and editing

Compliance with Ethical Standards

The authors declare no conflict of interest. This research study does not involve the collection of data from human participants.

Funding Information

This research has not received funds from any institution.

References

[1] U. Singh, K. Abhishek, & H. K. Azad. ”A Survey of Cutting-edge Multimodal Sentiment Analysis.” *ACM Computing Surveys,* vol. 56, no. 9, pp. 1-38, 2024.

[2] K. Du, F. Xing, R. Mao, & E. Cambria. ”Financial sentiment analysis: Techniques and applications.” *ACM Computing Surveys,* vol. 56, no. 9, pp. 1-42, 2024.

[3] G. Popoola, K. K. Abdullah, G. S. Fuhnwi & J. Agbaje. ”Sentiment Analysis of Financial News Data using TF-IDF and Machine Learning Algorithms.” *In 2024 IEEE 3rd International Conference on AI in Cybersecurity (ICAIC)*, pp. 1-6, 2024.

[4] M. Arif, M. Hasan, S. A. Al Shiam, M. P. Ahmed, et.al,. ”Predicting Customer Sentiment in Social Media Interactions: Analyzing Amazon Help Twitter Conversations Using Machine Learning.” *International Journal of Advanced Science Computing and Engineering,* vol. 6, no. 2, pp. 52-56, 2024.

[5] P. S. Ghatora, S. E. Hosseini, S. Pervez, M. J. Iqbal, & N. Shaukat. 'Sentiment Analysis of Product Reviews Using Machine Learning and Pre-Trained LLM.' *Big Data and Cognitive Computing*, vol. 8, no. 12, pp. 199, 2024.

[6] N. A. Sharma, A. S. Ali, & M. A. Kabir. 'A review of sentiment analysis: tasks, applications, and deep learning techniques'. *International journal of data science and analytics*, pp. 1-38, 2024.

[7] M. A. Soomro, R. N. Memon, A. A. Chandio, M. Leghari, & M. H. Soomro. 'A dataset of Roman Urdu text with spelling variations for sentence level sentiment analysis'. *Data in Brief*, vol. 57, pp. 111170, 2024.

[8] M. A. Soomro, R. N. Memon, A. A. Chandio, M. Leghari, & M. Khalid 'Spelling Variation of Roman Urdu Using Machine Learning.' *Journal of Computing and Biomedical Informatics*, vol. 7, no. 02, 2024.

[9] K. Jawad, M. Ahmad, M. Alvi, & M. N. Alvi. 'RUSAS: Roman Urdu Sentiment Analysis System'. *Computers, Materials & Continua,* vol. 79, no. 1, 2024.

[10] G. Manias, A. Mavrogiorgou, A. Kiourtis, C. Symvoulidis, & D. Kyriazis. 'Multilingual text categorization and sentiment analysis: a comparative analysis of the utilization of multilingual approaches for classifying twitter data'. *Neural Computing and Applications*, vol. 35, no. 29, pp. 21415-21431, 2023.

[11] Z. Wang, Z. Hu, S.-B. Ho, E. Cambria, and A.-H. Tan. 'MiMuSA—mimicking human language understanding for fine-grained multi-class sentiment analysis'. *Neural Comput. Appl.*, pp. 1–15, 2023.

[12] S. Maruvur Selvi and P. S. Sreeja. 'Sentimental Analysis of Movie Reviews in Tamil Text'. *In 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 1157–1162, May 2023.

[13] R. Keinan and Y. HaCohen-Kerner. 'JCT at SemEval-2023 Tasks 12 A and 12B: Sentiment Analysis for Tweets Written in Low-resource African Languages using Various Machine Learning and Deep Learning Methods, Resampling, and HyperParameter Tuning'. *Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023),* pp. 365–378, 2023.

[14] P. Shah, P. Swaminarayan, and M. Patel. 'Sentiment analysis on film review in Gujarati language using machine learning'. *Int. J. Electr. Comput. Eng. IJECE,* vol. 12, no. 1, pp. 1030, 2022.

[15] M. Al-Ayyoub, A. A. Khamaiseh, Y. Jararweh, and M. N. Al-Kabi. 'Comprehensive survey of arabic sentiment analysis'. *Adv. Arab. Nat. Lang. Process. ANLP Its Appl.*, vol. 56, no. 2, pp. 320–342, 2019.

[16] M. A. Qureshi, M. Asif, M. F. Khan, A. Kamal, and B. Shahid. 'Roman Urdu Sentiment Analysis of Songs 'Reviews.' *VFAST Trans. Softw. Eng.*, vol. 11, no. 1, pp. 101-108, 2023.

[17] J. Jawad, A. Kazim, M. Ahmad, M. Alvi, and M. B. Alvi. 'RUSAS: Roman Urdu Sentiment Analysis System'. *Comput. Mater. Continua,* vol. 79, no. 1, 2024.

[18] M. Hammad and H. Anwar. 'Sentiment analysis of Sindhi tweets dataset using supervised machine learning techniques'. *In 2019 22nd International Multitopic Conference (INMIC),* pp. 1-6, 2019.

[19] F. Noor, M. Bakhtyar, and J. Baber. 'Sentiment analysis in e-commerce using SVM on Roman Urdu text'. *In Second International Conference on Emerging Technologies in Computing, iCETiC*, London, UK, August 19–20, 2019,

[20] A. Gandhi, K. Adhvaryu, S. Poria, E. Cambria, & A. Hussain. 'Multimodal sentiment analysis: A systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions.' *Information Fusion,* vol. 91, pp. 424-444, 2023.

[21] M. Das, & P. J. A. Alphonse. 'A comparative study on tf-idf feature weighting method and its analysis using unstructured dataset'. *arXiv preprint arXiv:2308.04037*, 2023.

[22] S. Akuma, T. Lubem, & I. T. Adom. 'Comparing Bag of Words and TF-IDF with different models for hate speech detection from live tweets'. *International Journal of Information Technology,* vol. 14, no. 7, pp. 3629-3635, 2022.

[23] R. K. Das, M. Islam, M. M. Hasan, S. Razia, M. Hassan, & S. A. Khushbu. 'Sentiment analysis in multilingual context: Comparative analysis of machine learning and hybrid deep learning models'. *Heliyon*, vol. 9, no. 9, 2023.

[24] P. Savci, & B. Das. 'Prediction of the customers' interests using sentiment analysis in e-commerce data for comparison of Arabic, English, and Turkish languages.' *Journal of King Saud University-Computer and Information Sciences,* vol. 35, no. 3, pp. 227-237, 2023.

[25] J. Chen, & L. Pan. 'An AI based cross-language aspect-level sentiment analysis model using English corpus'. *Engineering Reports*, vol. 6, no. 12, pp. 2969, 2024.

[26] B. Chandio, A. Shaikh, M. Bakhtyar, M. Alrizq, J. Baber, A. Sulaiman, & W. Noor. 'Sentiment analysis of roman Urdu on e-commerce reviews using machine learning.' *CMES-Comput. Model. Eng. Sci,* vol. 131, no. 3, pp. 1263-1287, 2022.

[27] M. Bilal, M. Israr, M. Shahid, & A. Khan. 'Sentiment classification of Roman-Urdu opinions using Naïve Bayesian, Decision Tree and KNN classification techniques.' *Journal of King Saud University-Computer and Information Sciences,* vol. 28, no. 3, pp. 330-344, 2016.

[28] K. Shah, H. Patel, D. Sanghvi, & M. Shah. 'A comparative analysis of logistic regression, random forest and KNN models for the text classification.' *Augment Hum Res*, vol. 5, no. 1, pp. 12, 2020.

[29] X. Fang, & J. Zhan. 'Sentiment analysis using product review data.' *Journal of Big data,* vol. 2, pp. 1-14, 2015.