

Sentiment Analysis of Roman Text: Challenges, Opportunities, and Future Directions

Mubasher H. Malik^{*1}, Hamid Ghous², Muhammad Imran Ali³, Maryem Ismail⁴, Zeeshan Haider Ali⁵,
Hafiz Muiz Amin⁶

¹Vision, Linguistics & Machine Intelligence Research Lab Multan, Pakistan

²Australian Scientific & Engineering Solutions, Sydney, New South Wales, Australia

^{3,4,5,6}Institute of Southern Punjab, Multan, Pakistan

Corresponding Author: Mubasher H. Malik (mubasher@isp.edu.pk)

ABSTRACT

The surge of social media platforms has generated significant interest among researchers in Natural Language Processing (NLP) regarding Sentiment Analysis in recent years. Numerous studies have been dedicated to analyzing sentiment in different languages, including Urdu, which is widely spoken in South Asia and serves as the national language of Pakistan. While considerable research has been conducted on sentiment analysis using Machine Learning (ML), Deep Learning (DL), and other methodologies, the importance of sentiment analysis for Roman Urdu text is equally noteworthy. This paper provides a comprehensive examination of existing literature on sentiment analysis in various languages, with a specific emphasis on Roman Urdu text and approaches such as ML, DL, and others. Moreover, the paper proposes new directions for future researchers in this field.

Keywords: Roman text, Machine Learning, Deep Learning, Sentiment Analysis, NLP.

1. INTRODUCTION

NLP is a branch of Artificial intelligence (AI) that is still in its infancy and assesses, comprehends, examines, and even generates human language. It is used in the development of human-machine interaction interfaces. Currently, the internet and technology are both at their peak. Since the last decade, social media platforms have been widely used by a large number of people. People utilize social media platforms to share their thoughts, ideas, experiences, and concerns with family and friends. The number of social media networks, users, social media usage time, and data have all expanded dramatically. People connect in languages they speak on social media networks, and no language filter is used. In terms of NLP, this is a challenging scenario [1], [2].

The fast-expanding area of NLP is sentiment analysis (SA). Utilizing this method, the text is analyzed to reveal its underlying meaning. SA allows an application to identify whether a text's sentiment is positive, negative, or neutral. Urdu is the native language of Pakistan, the world's fifth-largest country, with a population of 216 million people. In South Asia, Roman Urdu text is frequently used for social media conversation. Roman Text SA is a difficult problem for academics because of its unconventional text, unstructured form, and lack of resources [3], [4]. Roman Text has un-structured and non-standard language Example;

“Ali bohat zaheen ha.”

“Ali both zahen hy.”,

“Ali bht Zahin hai.”

All of the above mentioned sentences are understandable and proper in Roman Urdu language, however, they are difficult to comprehend in ML. Because there is no sentiment dataset for Roman text, the data is processed using ML algorithms. Machines follow specific rules and regulations based on criteria. The most difficult step in processing Roman Text Datasets is preprocessing. This is how we prepare data for ML. The Roman text expresses emotions through a variety of emotive words that are used to denote such emotions. This is a distinct language with sophisticated grammar and a diverse morphology. Because the tools available for other languages are insufficient for Urdu, unique tools, particularly for SA, are required [5], [6].

The human interest in others' feelings and opinions is as old as verbal communication [7]. In the first decades of the twentieth century, attempts were made to calculate and measure public sentiment through questionnaires [8]. In 1937, a scientific journal on public opinion was published. We discovered that the first scholarly study of public

sentiments was conducted during and after WWII, and its objective was purely political [9]. Modern SA becomes much more popular in the mid-2000s and it is mainly based on online product reviews [10]. Since then, SA has been applied to a variety of other fields, including financial market forecasting and terrorist attack reactions [11]. Moreover, attempts to distinguish negative emotions such as anger and grief are moving from basic polarity recognition to more complicated variations of emotions and separating negative emotions such as anger and grief [12].

2. LITERATURE REVIEW

Many ML and DL methods have been applied for text-based SA in the past.

2.1. Machine Learning Methods

A ML model was developed to detect sentiments from Roman text. The model was tested using Roman Urdu song reviews collected from YouTube. Preprocessed comments were analyzed using Naïve Bayes (NB), K-nearest Neighbor (KNN), Decision Tree (DT), and Artificial Neural Network (ANN) algorithms. The Naïve Bayes algorithm achieved the highest accuracy of 82.41% [13]. A study on dual language sentiment analysis (SA) was conducted using a Roman Urdu dataset from Kaggle. The dataset consisted of comments from the social media platform YouTube in both English and Roman Urdu languages. Preprocessing steps involved removing null values, stopwords, punctuations, special characters, spaces, numbers, unnecessary text, and duplicate values. Three feature selection techniques, namely Term Frequency - Inverse Document Frequency (TF-IDF), count vectorization, and word2vec, were utilized. Logistic Regression with count vectorization achieved a model accuracy of 87%. The study suggested the potential use of Google Translate for multi-language SA and video ranking based on qualitative content on YouTube [14].

In this study, the Twitter dataset from GitHub was utilized for sentiment analysis (SA) of English text. The dataset contained 40,000 tweets, which were used for training purposes. The aim was to classify tweets into happy, sad, and other emotions. Preprocessing steps, including tokenization, lowercase conversion, filtration, and POS tagging, were applied. Various features extracted from the text were categorized into morphological features and word N-gram features. Classification was performed using Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF) algorithms, with DT achieving an accuracy of 99.3%. For future research, the evaluation of feature extraction methods for Twitter data can be further explored [15].

TERUN, a transliteration-based encoding method, was used to normalize Roman Urdu/Hindi text. The RUSAD dataset contained 11,000 social media reviews, enabling the detection of positive and negative emotions. Preprocessing involved removing repeated digits and zeros. Word-level n-gram features were employed. The studies demonstrated the importance of correct normalization for improved classification accuracies. It was emphasized that phonetic algorithms need to be tailored to specific languages. Both intrinsic and extrinsic evaluations were conducted, and future work involves exploring deep learning approaches for standardizing RUSA [16]. A study was conducted to analyze multilingual textual data from social media and identify the level of extremism. The dataset included 20,000 rows annotated for moderate, low extreme, high extreme, and neutral categories, covering Urdu, Roman Urdu, and English languages. Data cleaning steps involved removing URLs, converting text to lowercase, and eliminating blank spaces and emojis. TF-IDF was used to assign weights to important words for content classification. Two machine learning algorithms, Linear Support Vector Classifier and Multinomial Naive Bayes, were employed for data classification. The Linear Support Vector Classifier achieved the highest accuracy of 82% [17].

Developed a model for Aspect-based SA of the Urdu language. For experiments dataset was prepared using Twitter APIs, 10,000 tweets 3000 from football, and 7000 from the cricket domain. Preprocessing removed special characters, hashtags, and punctuation using Sci-kit Learn. TF-IDF vectorizer and N-gram models were used for feature extraction. In the proposed methodology, the main steps were data collection, preprocessing, ML algorithm, and evaluation. A baseline evaluation was performed It presented a benchmark corpus for aspect-based SA in Urdu that will cover many areas of research. Improvement of work by applying different techniques for feature extraction was the future goal of this work [18]. Here, presented the online sentiments analysis using NLP and human validation to score net sentiment scores and online conversion topic categorization. The data from 1720810 reviews were collected from five banks in South Africa with the help of social media APIs. Positive and negative emotions were detected from Reviews that are in the English Language. The combined process of the ML algorithm and manual validation created a confidence level of 95% and an overall 0.1% margin error [19].

The research introduced the first offensive language dataset for Roman Urdu, comprising 147,000 YouTube comments. Preprocessing techniques like normalization, stop word removal, and character-level tri-grams were employed. Machine learning algorithms were applied, with Simple Logistic Regression achieving a remarkable 95.8% F-measure on the Urdu dataset. Future work includes creating a multilingual text dataset for Urdu and Roman Urdu from different social platforms [20]. A model was developed to measure hate speech and the content of hate speech. A publically available dataset from GitHub has 43,871 English tweets annotated against positive, negative, and neutral classes used in this work. In Preprocessing stage, conversion into lowercase, deletion of punctuation marks, duplicate texts, and digits took place. ML libraries were used and results showed 89% of tweets were negative, 7% positive, and 4% neutral. Legal research about hate speech must be done in the future [21].

Analyzed that SA performed social media comments in Roman Urdu and prepared a dataset of 806 comments from different social media sites and deal only with positive and negative sentiments. Stop words, punctuation, and numeric characters were removed in the preprocessing phase. Uni-gram, bi-gram, tri-gram, and union with TF-IDF feature extractions were applied. After the data collection, preprocessing, and feature extraction ML techniques like NB, LR, LRSKD, and SVM to predict sentiments applied in model construction. Predicted results were compared with manual targeted results and evaluated using the WEKA tool. SVM with Unigram+Bigram-TF-IDF gave the highest accuracy 87.22%. In the future, Will expand the dataset and use DL approaches [22] Composed a methodology to perform sentiment classification on Urdu news tweets. An Urdu dataset was collected through Twitter APIs with 600 tweets in the Urdu Language. URLs, stop words, and Twitter-specific emotions were removed during preprocessing. Feature vector composed through the identification of positive and negative word count, presence of negation, and POS tags. The DT algorithm was used for classification and gave 90.00% accuracy and evaluation was done through cross-fold validation as well as precision, accuracy, and recall [23].

Analyzed Twitter live data, as well as past, feeds for getting emotions. Twitter and 20 newsgroups were used as datasets annotated against enjoyment, disgust, and surprise emotion classes. The Preprocessing involved replacing HTTP links with URLs, eliminating new lines, tab and hash characters, and conversion into lowercase. Different ML algorithms like k-means, NBand SVM are applied in this model. Further, more words describing emotions can be added to the bags [24] . In this study, emotions expressed on Twitter were detected and analyzed using a benchmark dataset called ISEAR. The dataset consisted of 7,246 English tweets and was used for experimental work. Positive, negative, and neutral emotions were identified from the tweets. Preprocessing steps involved removing emojis, redundant characters, non-alphanumeric characters, tokenization, stop words removal, Part of Speech tagging, and NAVA text processing. The feature set used for Naive Bayes (NB) included words from each tweet and reply. The model consisted of data collection, preprocessing, sentiment detection, and influential user detection stages. Naive Bayes yielded the best results for both full text and NAVA words in terms of sentiment and emotion analysis. As a future direction, it is suggested to consider complete posts rather than individual words to provide insights on user emotions for any social network or topic [25].

Aspect Based SA was performed for Urdu political tweets that were collected through Twitter APIs. 8760 Urdu tweets were collected then special characters, duplicate tweets, URLs, and hyperlinks are removed in preprocessing and segmentation also took place in the preprocessing stage. The N-gram feature set and SVM algorithm are used for sentiment purposes. The developed corpus has been tagged at aspect and sentiment levels with an IAA score of 75%. The sentiment classifier also achieved an f1 score of more than 70% in many cases. A baseline system has been developed using the corpus for analyzing its reliability in future use [26].

The development of the Urdu Sentiment Lexicon involved using textual data sourced from various Urdu blogs. The aim was to detect three emotions: positive, negative, and neutral, within Urdu text. Instead of using a supervised machine learning model, a lexicon-based system was developed. The model took into consideration positive and negative words, as well as handling negation, intensifiers, and content-dependent words. The system achieved an accuracy of 89.03% with a precision of 0.86%, recall of 0.90, and an F-measure of 0.88. These results surpass previous work in sentiment analysis (SA), and the evaluation was performed using Kappa statistics to validate the findings. It is important to note that this research did not address the detection of sarcasm, a common occurrence in Urdu text. Thus, a potential future direction could focus on handling sarcasm within the lexicon-based system to further enhance its effectiveness and coverage [27].

Roman Urdu dataset was used to perform SA on 779 reviews from different social media platforms. Textual data were converted into a numeric form in the preprocessing stage. N-gram features worked as feature sets and ML

algorithms NB, LR, SVM, KNN, and DT were used in the model. The model was constructed by following these steps: Data gathering, data annotation, data preparation, feature identification, and Supervised ML algorithms and results. F-measure is higher than all previous work and to analyze the impact of stop words list and automatically normalize the Roman Urdu text suggested as future work [28].

The Waikato Environment for Knowledge Analysis was utilized as a platform for executing classification models in the context of text classification for Roman Urdu. The dataset was collected by scraping various automobile websites, resulting in a dataset containing 1000 positive and 1000 negative reviews, which were then saved in WEKA. The data was divided into training and testing sets, with 80% used for training purposes and the remaining portion for testing. Different models were applied during the testing phase, and the results were analyzed accordingly. The experiments revealed that the Multinomial Naive Bayes (NB) classifier outperformed other models such as deep neural networks, decision trees (DTs), random forests (RF), bagging, K-nearest neighbors (KNN), AdaBoost, and support vector machines (SVM) in terms of precision, recall, accuracy, and F-measure. The preprocessing steps involved data extraction, removal of stop words, conversion to lowercase, and the development of a corpus. The Multinomial NB classifier achieved an accuracy of 89.75% in the classification task. In future work, there is a plan to train the machine to identify the specific areas of interest within the automobile domain, such as individual parts of a vehicle [29].

A dataset consisting of 11,000 reviews in Roman Urdu was collected from six different domains for sentiment analysis (SA). The dataset was annotated with three classes: positive, negative, and neutral sentiments. Unnecessary information was removed from the reviews, and the cleaned data was stored in an Excel file. In terms of analysis, the ensemble technique called wVoting outperformed machine learning algorithms from seven different categories in most cases. The best results showed a significant reduction in the error rate by 12% compared to the baseline, which was initially at 80.07%. The evaluation of the SA was performed using deep learning techniques. For future work, it would be interesting to further analyze and improve the results obtained from these studies. Additionally, exploring feature reduction techniques could be beneficial in enhancing the accuracy and efficiency of the sentiment analysis model [30]. In order to enhance the accuracy of Urdu sentiment analysis (SA) across multiple domains, special attention was given to dealing with negations. Negation plays a crucial role in SA, so it was a primary focus of the research. To address this, a corpus of 6025 Urdu sentences was utilized, where negations were identified and observed. To tackle negation, the first important step involved handling morphological negation. This negation was incorporated into the word negative file for SA, ensuring its consideration in the analysis. The second step employed a rule-based approach to address both implicit and explicit negations. Specifically designed rules were implemented, resulting in a significant improvement in accuracy, from 73.88% to 78.32%. The precision, recall, and F-measure also demonstrated statistically good improvement, with precision at 0.745, recall at 0.788, and F-measure at 0.745. In summary, by effectively dealing with negations, particularly considering both local and distance negations, the accuracy of Urdu SA in various domains can be significantly enhanced. The rule-based approach applied in this research proved successful, leading to notable improvements in accuracy and performance metrics [31].

A Discriminative Feature Spamming Technique (DFST) was introduced, utilizing the Term Utility Criteria (TUC) to identify distinctive terms. This technique was employed on a Roman Urdu dataset consisting of 11,000 reviews, which were collected and annotated specifically for this research. To enhance the classification accuracy, a customized tokenizer was developed. This tokenizer played a crucial role in improving the performance of the model. A cross-scheme comparison was conducted, comparing the current results with previous ones. The findings demonstrated that the model's performance had been improved in comparison to the previous results. In summary, the research presented the DFST approach, incorporating TUC to identify distinctive terms. The customized tokenizer contributed to improved classification accuracy. The comparison with previous results showcased the enhanced performance of the model [32].

Targeted the students' sentiments posts on Facebook groups to express their feelings. The dataset was based on the issues of the students. This system provided detection to students' posts in the English language, related issues were Positive, negative, and neural. Long text broke into small chunks after preprocessing chunks were merged and tokenization performed. Feature selection is performed based on unigram, bigram, and trigram. After data collection and preprocessing, the scoring phase took place. Then there was semantic score calculation and classification. This model provided the best outcomes to improve related issues and institution policies. Moreover, this work can be performed in different languages in the form of formal and informal writing styles, emotional symbols, and images that will be used for emotional facial extraction [33].

A novel approach for sentiment analysis (SA) on data from multiple languages was proposed. The experiments involved two publicly available datasets in English and German languages, which were annotated with positive, negative, and neutral sentiments. Notably, no preprocessing steps were applied in this particular model. The proposed approach utilized a hybrid model combining bidirectional long short-term memory networks (LSTM) and support vector machines (SVM). This hybrid model achieved an accuracy of 74.5%. The F1-score without oversampling was 49%. It is worth mentioning that the designed model, featuring a 300-dimensional embedding vector, outperformed the base model. The model demonstrated its applicability to mixed languages, although there was a notable decline in predictive performance. Nonetheless, the approach proved successful in conducting sentiment analysis across multiple languages, albeit with some reduction in accuracy compared to single-language analysis [34].

Sentiment mining performed on Youtube Roman Urdu/Hindi reviews. A Roman Urdu/Hindi dataset was presented that contains 600 reviews annotated against positive, negative, and neutral reviews. Data integration, filtration, and labeling were performed as preprocessing. The reviews of three selected songs were scrapped and preprocessed and analyzed using NBand ID3. Both algorithms gave 75% accuracy. An increase in dataset size by using ML techniques is suggested as a future domain [35]. Proposed a fake product review system that was an intelligent interface and took URLs related to products from different platforms and analyzed the product reviews and provide original ratings. The system analyzed reviews in Urdu, Roman Urdu, and English language. Deceptive Opinion Spam Corpus had 1600 reviews in the English language. Stop words and punctuation removal, lemmatize and bigram steps were related to preprocessing. Count-Vectorizer converted each review into a bag of words and then TF-IDF transformed applied. The model was based on preprocessing, feature extraction, classifiers, and fake review detection steps. SVM Classifier gave 87% accuracy. SVM Classifier used with Bigram feature used for evaluation. In the Future, the system will process and filter the reviews in multiple languages [36].

To determine the best classifier for Urdu sentiment analysis (SA), evaluation measures were effectively utilized. A dataset of 6025 Urdu sentences was collected from various online sources. In the data cleaning process, stop words were removed to improve data quality. After data collection and annotation, five classifiers (SVM, DT J48, KNN-IBK, PART, and NB multinomial Test) were employed using the WEKA tool. To select the optimal classifier, a thorough evaluation was conducted. A 10-fold cross-validation technique was applied, and the top three classifiers with high accuracy, precision, recall, and F-measure were chosen. These classifiers were SVM, J48 (decision tree), and IBK (k-nearest neighbors). IBK exhibited superior performance compared to other classifiers when evaluated using additional measures such as the Kappa statistic and root mean squared error [37]. Stated, Used the Twitter handle sample data, and analyzed and check the status of Urdu language communication on Twitter. Using different hashtags 2234 tweets were collected in Urdu and Roman Urdu. Roman Urdu and Urdu Tweets were annotated against positive, negative, and neutral emotions. Applied Combine ML and lexicon-based approaches. Python is used for data analysis, Tweepy API, and maps tweets converted from latitude to longitude to an address. Results showed that Urdu was not a strange language to Twitter. Urdu had issues with lexicon-based SA which were suggested as a future domain [38].

SA of Sindhi language using supervised ML techniques. 4236 Sindhi tweets downloaded using Twitter API annotated against positive and negative emotions. Preprocessing involved the removal of non-Sindhi words, extra white spaces, and punctuation and tokenization. SVM, NB, DT, and KNN Algorithms were used. SVM gave accuracy= 64.6%. The dataset was evaluated based on F-score, Accuracy, Precision, and recall. Future work can focus on more classes and some other tokenization and preprocessing techniques can be opted for better results [39].

Classify the emotions of tweets using the ML approach. Three datasets were used named Stanford sentiment140, Polarity Dataset, and University of Michigan Having records 1,600,000, 2000, and 7086 respectively. Two types of emotions were detected positive and negative. The Preprocessing stage included lowercase conversion, conversion into URLs, stemming, trimming tweets, and tuning parameters. SVM and DT algorithms were merged to get better results in terms of f-measure and accuracy. DT: F-measure84% Accuracy=84% [40].

The main focus was Urdu SA using the Supervised ML Approach. Data was collected from 151 different blogs which were 6025 sentences. Data annotated against positive, negative, and neutral classes. Three classifiers SVM, DT, and KNN were tested and compared. Different steps included stops words removal, feature extraction, identification, and extraction of important features improved results. K-NN performed better than SVM and DT in terms of precision, recall, accuracy, and f-measure. KNN (IBK) Accuracy= 100%. In the future, more data can be added to blogs to get better results [41].

2.2. Deep Learning Methods

A DL model was developed to evaluate word embeddings specifically for Roman Urdu text. The CNN and LSTM models were employed to extract features from the data. Subsequently, different machine learning (ML) classifiers were utilized for classification purposes. The proposed DL model achieved impressive accuracy results. It obtained accuracies of 0.904, 0.841, 0.740, and 0.748 on the MDPI, RUSA, RUSA-19, and UCL datasets, respectively [42]. A DL model was introduced to conduct data mining on the emotions expressed in Roman Urdu text. The study utilized a Roman Urdu Corpus, specifically the RUSA-19 dataset, for conducting experiments. Various models, including FRCNN, RCNN, rule-based, and N-Gram-based models, were employed for the analysis. The models were utilized for both binary classification and tertiary classification of emotions [43]. A DL model was developed for SA in Roman Urdu. A dataset named RUSA-19 was used which contains 10,021 Roman Urdu sentences and UCL Roman Urdu Dataset was used for comparison. Positive, Negative, and Neutral emotions were detected from different social media platforms. Illegal characters and URLs were removed in a preprocessing step. N-gram feature extraction was performed for binary classification. RCNN model outperformed the Rule-based model concerning evaluation measures that were used. For future work, Improvement in results was suggested [44]. NLP and DL algorithms are applied for automatic content monitoring from tweets in English and Hindi Language. In the preprocessing phase, conversion into lowercase, negation removal, URL and non-alphanumeric removal, and tokenization is performed. Pre-trained language models such as DistilBERT and BERT achieved the best performance for AOB detection. The problem of Noisy labels to correct wrongly assigned classes is represented as future work [45].

A specialized model was created to detect hate speech and offensive language on social media platforms, with a particular focus on Twitter. For the experiments, a publicly available dataset called RUHSOLD, containing 10,012 tweets in Roman Urdu, was utilized. A novel deep learning model called CNN-gram was developed specifically for offensive and hate speech detection. The study concluded that transfer learning, as opposed to training embedding from scratch, yielded more favorable results. The CNN-gram model demonstrated greater robustness compared to the baseline models. However, the analysis also revealed that there were semantically challenging cases at a fine-grained level due to the complexities of language. As a result, further research and development are needed to address these linguistic complexities and improve the model's performance in handling such cases [46]. The analysis focused on Twitter data related to COVID-19, specifically targeting tweets in the English language with positive and negative emotions. Data collection was done using Twitter APIs, and an additional cleaning process was applied using the "clean tweet" method. To classify emotions in the collected tweets, a Recurrent Neural Network (RNN) model was developed. The RNN model demonstrated good performance and accuracy in text classification tasks. To delve deeper into the analysis, a sample of 50 tweets containing the keyword 'COVID' was analyzed using both TextBlob and the RNN model. Additionally, classification and clustering techniques were employed to explore the data from various perspectives and areas. Overall, the study successfully employed a combination of data collection, cleaning, and RNN-based sentiment analysis to gain insights into COVID-19-related Twitter data. Further analysis using classification and clustering techniques expanded the understanding of the data from diverse angles [47].

A novel approach called the Self-Attention Bidirectional LSTM (SA-BiLSTM) network was introduced to address the issue of inconsistent sentence structure in text representation. The model was evaluated on a Roman Urdu Dataset comprising 20,000 sentences sourced from social media. These sentences were annotated with positive, negative, and neutral labels. As part of the preprocessing phase, non-textual symbols and characters were eliminated. The proposed SA-BiLSTM model leveraged self-attention and a Bidirectional LSTM network to achieve improved results. The SA-BiLSTM model attained high accuracy rates of 68.4% and 69.3% for the preprocessed and normalized datasets, respectively [48].

DL methods applied to perform Roman Urdu SA. A balanced dataset with 8k records was used for experimental work. The proposed DL model gave 0.97, 0.92, 0.94, 0.95 precision, recall, F1 score, and accuracy respectively. LSTM is used to train the model and it will help to ensure further exploration [49]. A novel model called SLCABG (Sentiment Lexicon and CNN-Attention-based Bidirectional Gated Recurrent Unit) was developed to tackle the limitations of existing models used for analyzing product reviews. This model combines the advantages of sentiment lexicons and deep learning techniques to overcome these shortcomings. It leverages a combination of CNN and Bidirectional Gated Recurrent Unit (GRU) networks to extract essential sentiment and contextual features from the reviews. The dataset was extensive, reaching a magnitude of 100,000 orders, thus providing wide applicability in the field of Chinese sentiment analysis. Experimental results demonstrated the model's effectiveness in improving text sentiment analysis performance [50]. Examined that A Model based on a CNN and gating structure developed that

was efficient and accurate. SemEval Datasets contained restaurant and laptop English reviews annotated against positive, negative, neutral, and conflict classes. Novel Gated Tanh-ReLU units selectively output sentiment features according to aspect and unity. The convolutional Layer did not have time dependency as in the LSTM layer that's why the computations of the proposed model could be easily parallelized. Leveraging large-scale sentiment lexicon in neural networks would be a future task [51].

2.3. Hybrid Methods

The three neural word embeddings—word2vec, FastText, and Glove—were provided. A data set in Roman Urdu with three thousand two hundred forty one sentiments from Twitter and "what mobile" platforms annotated for positivity, negativity, and neutrality. The performance of DL and ML methods was compared using NB and SVM Logistic Regression and Hybrid approaches, as well as neutral word embeddings. The proposed model had accuracies of 0.8417, precision of 0.8168, recall of 0.8284, and F1-core of 0.8221. The future work will be look into the impact of developed neural word embeddings on cyberbullying detection and machine translation [52].

A hybrid approach of DL called CNN and LSTM is used for sentiment classification of reviews. Four different English language datasets were prepared: Movie Reviews, Airline reviews, self-driving car data, and US presidential election statistics are all available. These datasets are categorised either as positive, negative, or neutral. Removal of special characters, removal of emoticons, and removal of hyperlinks were performed in the preprocessing stage. Feature Vectorization was performed and CNN, RNN, and LSTM algorithms were applied in this work. In the Model Development phase, Input reviews, preprocessing, Vectorization, feature matrix, model training using DL, classification, and results are performed and evaluated through accuracy, precision, recall, and f-measure parameters. The results showed that performance in several areas is significantly improved when the key aspects are studied sequentially. In the future, the social media-specific word embeddings model will be able to improvise to deal with misspellings and other linguistic irregularities that can be challenging to deal with now [53].

2.4. Other Methods

A SA technique based on Markov Chains was presented for sentiment mining of Urdu tweets. The primary focus was on three Urdu categorization models: positive, negative, and neutral. For data cleansing, stop words, non-Urdu words, user names, special characters, emotions, and diacritics were eliminated. A classifier trained using NB Algorithm and feature matrix was computed using the TF-IDF metric. Corpus construction was performed after input tweets and then preprocessing, partitioning, modeling, and evaluation. The dataset for positive tweets increased to improve results in the future [54].

The problem of antisocial behavior detection from Twitter and Facebook Urdu text was discussed. Hashtags, HTML tags punctuation, and extra spaces were removed in preprocessing. The processes used to create the dimensional model for emotion recognition included input text preprocessing, polarity identification, segmentation, and ASB detection. In terms of accuracy, this approach exceeded the state of the art method by 77%. In the future, a multilingual or bilingual strategy may be used [55]. The segmentation approach was applied to Urdu SA. The bag of words model fails to classify a subset of complex emotions. A sentiment is classified into three steps, First segment sentiment into two fragments, the second orientation score of fragments is calculated independently, and the last polarity of sentiment is estimated. The computational model evaluated this method. Increase in Precision= 8.46%, Recall=37.25%, Accuracy=24.75% which was a big improvement over a bag of word model. Future work should focus on developing a complex interaction between sub-opinion and rigor theory to manage sentiments with multiple sub-opinions [56].

An emotion corpus for Roman Urdu was analyzed and generated for emotion evaluation. To help the Roman Urdu emotion recognition system, a corpus of ten thousand sentences were gathered from several sources. Noise removal, URL removal, punctuation marks removal, spelling consistency, and lowercase conversion steps were performed in preprocessing. The model is based on data collection, preprocessing, data annotation, and corpus characteristics. Future research may yield more corpus-based instances of emotions and findings from various classifiers [57]. Sentiment orientation emoji-based components and textual features studied. Data was collected from Twitter using different hashtags from Arabic tweets. In the preprocessing phase, unnecessary data was removed and comments were compiled in One Microsoft Word file. The proposed model analyzed positive, Negative, and neutral sentiments after the cleaning and compiling process. Emojis serve as a supplement to text or other emoji to express emotion; they do not, however, offer accurate information or convey what they are meant to convey. This work combines manual and computational techniques. In the future, one text or emoji lexicon with a broad list of neutral terms will

be employed in data analysis. Sarcasm and politeness will also be interesting subjects to investigate [58],[59],[60]. ML, DL, Hybrid and Other Methods details are shown below in Table 1.

Table 1. Urdu Roman Text Literature Survey using ML, DL, Hybrid and Other Methods

Author/Year	Dataset	Language	Preprocessing Steps	Methodology	Description
Qureshi et al. (2023)	Roman Urdu Songs Review	Urdu Songs Reviews	Remove Null, Spaces, Punctuation, and Special Characters	NB, KNN, DT and ANN	Navies Bayes Accuracy 82.42%
Nagra et al. (2022)	Roman Urdu Dataset using RUSA-19	Roman Urdu	Text Normalization	RCNN, FCNN, Binary and Tertiary Classification	RCNN Binary 91.73% Tertiary 89.94%
Khan L. et Al. (2022)	MDPI, RUSA, RUSA-19, and UCL datasets	Roman Urdu Text and English Dialects	Remove Null, Spaces, Text Normalization	CNN, LSTM, SVM, DT	SVM produced an accuracy of 90%
SaminaYasin et al (2020)	Roman-Urdu dataset	English & Roman Urdu	Remove: Null values, Stop 1 words, punctuation, Special characters, Spaces, Numbers, Unnecessary text, Duplicate values	Logistic Regression Based Model	Logistic Regression with count Vectorization: accuracy=87%
ZainabMahmood et al (2020)	RUSA-19,10021 sentences	Urdu	Remove: Illegal characters, URLs	Recurrent CNN model	RCNN model outperformed with the highest accuracy
Shikha Tiwari et al (2020)	Twitter dataset from Github,40,000 tweets	English	Tokenization, Converting into lowercase, Filtration, POS tagging	SVM, DT, RF	DT Accuracy= 99.3%
KhawarMehmood et al (2020)	Roman Urdu dataset RUSAD	Roman Hindi and Roman Urdu	Remove: Repeated digits, All zeros	KNN, DT, RF, NB, ANN, SVM, ADA boost	NB, DT Produced an Accuracy of 90%.
Abdulrahman Alrumaih et al (2020)	data was collected using different hashtags,1000 tweets	Arabic	Remove: Undesirable data, compiling the comments in one MS Word file	A predictive model	Accuracy above 90\$
Faiza Memood et al. (2020)	Roman Urdu Dataset, 3241 sentiments	Roman Urdu	Null	CNN	84% Accuracy

Muhammad Asif et al (2020)	Multilingual dataset, 20,000 rows, and 4300 lexicon count	Urdu, Roman Urdu, English	Remove: URL from posts, Text in URL, Blank Posts, Blank comments, Emojis, and Other characters, Convert into lowercase.	Multinomial NB, Linear support Vector Classifier	Linear SVC 82%. Accuracy
Zarmeen Nasim & Sayeed Ghani (2020)	Urdu tweets	Urdu	Remove: Stop words, Non-Urdu words, URL, User name, Special character, Emojis, Diacritics, Replacement of selected characters	Markov Chain	Accuracy above 80%
Elizaveta Zinov'yeva et al (2020)	1-“Toxic Comment Classification” from Kaggle	English, Hindi	Convert all data into lowercase, Negation removed, Substitute emojis with their semantic representation, Remove stop words, Remove URLs, Remove all non-alphanumeric characters, Tokenization	Deep learning Model	DL Methods average accuracy above 90%
Muhammad Sohail et al (2020)	—	Urdu	Remove: Hashtag HTML tags, Punctuation marks, Diacritics, Extra space	Dimensional model	Accuracy 77% achieved
Sadaf Rani & Muhammad Waqas Anwar (2020)	The dataset prepared using Twitter APIs, 10,000 tweets	Urdu	special characters, hashtags, links, and punctuation marks removal	NB, RF, KNN	Accuracy above 90%
Hammad Rizwan et al (2020)	RUHSOLD dataset, 10012 tweets	Roman Urdu	Null	CNN	Accuracy above 85%
László Nemes & Attila Kiss (2020)	collected through Twitter APIs	English	use the 'clean tweet' method to do the 'additional' cleaning.	Recurrent Neural Network (RNN)	RNN gives good performance
James Lappeman et al (2020)	Banks reviews from social media APIs, 1720810 posts	English	Null	Multinomial NB	A 95% confidence level and a margin of error of 0.1% .
Muhammad Arslan Manzor et al (2020)	Roman Urdu Dataset, 20,000 sentences	Roman Urdu	Remove: Non-textual symbols, Characters	CNN, LSTM, BiLSTM	High accuracy of 68.4% and 69.3%

MUHAMMAD PERVEZ AKHTER et al (2020)	Roman Urdu Dataset, 1,47,000 comments	Urdu and Roman Urdu	Normalize, Stop words removal, Remove digits, Eliminate punctuation marks, Remove nonlanguage characters	NB, KNN, DT, SVM, RF, logistic regression, Rule based	Logistic Regression achieved a 95.8% F-measure value.
Ranjan Kumar Behera et al (2021)	1-Movie Review,25000 2-Airline Review Dataset 3-Self Driving Car dataset 4-US Presidential Election Dataset, 13871 reviews	English	Remove: Special characters, Emoticons, Hyperlinks	CNN, RNN, LSTM	LSTM achieved higher accuracy as compared to CNN and RNN
F. H. A. Shibly et al (2021)	Twitter dataset 43871 tweets	English	Convert into lowercase, Remove all punctuation marks, Remove digits, Remove duplicate texts	ML libraries	There are 89% negative (446 out of 500), 4% neutral, and only 7% good tweets.
Ayesha Rafique et al (2019)	Roman Urdu Dataset, 806 comments	Roman Urdu	Remove: Stop words, Punctuation, Numerical characters	NB, Logistic regression with stochastic Gradient Descent(LRSGD), SVM	SVM gave the highest accuracy at 87.22% with Unigram+Bigram_TF-IDF
Hussain Ghulam et al (2019)	balanced Roman Urdu dataset, 8k records	Roman Urdu	Nil	NB, RF, SVM	RF produced Accuracy= 0.95
Raheela Bibi et al (2019)	Urdu dataset through Twitter APIs, 600 tweets	Urdu	Remove URL, Urdu stop words, Twitter-specific emotions removal	DT C45 algorithm used for classification	90.00% accuracy
Shaikh Abdul Salam & Rajkumar Gupta(2018)	Twitter and 20newsgroups were used as datasets, 6000 tweets	English	Replacing the HTTP links by URL Eliminating: new lines, tabs, Hash characters and converting	K-means, NB, SVM	NB 90% accuracy

KashfiaSailun aza&RedaAlh ajja (2019)	Benchmark ISEAR dataset, 7246 tweets	English	Remove: Emojis, URLs, redundant letters, non- alphanumeric characters, etc. Clean text, stop words tokenization, and POS tagging.	NB classifier	For full-text and NAVA phrases, NB produces the greatest results.
EhsanulHaq et al (2019)	Tweets collected using Twitter API, 8760 tweets	Urdu	Remove special characters, Resolve segmentation issues Remove duplicate tweets, Remove URLs and hyperlinks	SVM	classifiers have also achieved an F1 score of more than 70%.
NeelamMukh tar& Mohammad Abid Khan (2019)	Textual data	Urdu	Nil	Lexicon-based system rather than supervised ML	The system's accuracy is 89.03 percent, with 0.86 precision, 0.90 recall, and 0.88 F- measure.
KhawarMeh mood et al (2019)	Roman Urdu Dataset, 779 reviews	Urdu	Convert textual data into numeric form	NB, Logistic Regression, SVM, KNN, DT	F-measure, which is higher than all prior Urdu SA work
Moin Khan & Kamran Malik (2018)	reviews collected from different websites, 2000 reviews	Roman Urdu	Data extraction stop words removal, Conversion into lower case, Development of corpus	DT, Bagging, RF, Multinomial NB, KNN , Adaboost, SVM	Multinomial NBgave the highest accuracy of 89.75%
Muhammad Hassan & Muhammad Shoaib (2018)	product reviews of electronic itemsand reviews of automobiles and cosmetics.,	Urdu	N/A	ML and DL methods	Precision is increased by 8.46%, recall is increased by 37.25%, and accuracy is increased by 24.75%.
KHAWAR MEHMOOD et al (2019)	Collected from six different domains, 11000 reviews	Roman Urdu	Remove unnecessary data and save it as an Excel file	The ML algorithm of seven different categories	ensemble technique performed better
NeelamMukh tar et al (2019)	Data collected from 151 different blogs, 6025 sentences	Urdu	N/A	Supervised ML approaches	With 0.745, 0.788, and 0.745 Precision, Recall, and F-measure, accuracy improved from 73.88% to 78.32%.

KhawarMehmood et al (2019)	Roman Urdu Dataset, 11000 reviews	Roman Urdu	N/A	LR, NB, ANN, VOTING	82.16 was achieved by WVoting
LI YANG et al (2019)	prepared book review dataset, 10000 reviews	Chinese	Remove stop words and non-Chinese characters from the remark using the word segmentation tool in the jilbab package.	NB, CNN	NN produced better accuracy as compared to CNN
AteyaIram (2018)	Facebook Comments of a 500-Post Local College	English	Tokenization, Noisy data removal	semantic classification	delivers the best results to enhance relevant issues, matters, and institution policies.
Muhammad HaroonShakeel et al. (2020)	1-Twitter corpus, 5513 tweets 2-GermEval dataset,	English, German, mixed	N/A	SVM, CNN, NB, DT, biDirectionalL SVM	Model F1-score of 61:24%
Wei Xue and Tao Li (2018)	SemEval datasets	English	N/A	CNN	Extensive trials on SemEval datasets show that our neural model outperforms other neural models.
Muhammad Umair Arshad et al . (2019)	Roman Urdu Emotion Detection Dataset) RUED,10000 sentences	Roman Urdu	Noise removal, URLs removal, Removal of punctuation marks, Spelling consistency, Conversion into lowercase	Data collection Preprocessing, Data annotation, Corpus characteristics	The model produced outperformed results.
Muhammad Asif et al (2019)	Roman Urdu/Hindi dataset,600 reviews	Roman Urdu /Hindi	Data integration, Data filtration Data labeling	supervised ML algorithm, DT, NB	accuracy : NB=75%, ID3(DT) =75%
Ata-Ur-Rehman et al. (2019)	Deceptive Opinion Spam Corpus1, 1600 reviews	English	Remove punctuation, Stop words removal, and Lemmatize words	SVM	Accuracy: SVM=87%

NeelamMukhtar et al (2017)	data collected from online sources, 6025 sentences	Urdu	stop words removal	NB, SVM, DT, J48, KNN, IBK	SVM, J48, and IBK produced high accuracy,
Saqib Muhammad Ghulam& Tariq Rahim Soomro (2018)	tweets collected using Twitter API by hashtags, 2234 tweets	Urdu, Roman Urdu, English	N/A	Combine ML and lexicon-based approach	The model produced a Higher Accuracy of 90%.
Muhammad Hammad&Harris Anwar (2019)	Sindhi tweets Dataset	Sindhi	Tokenization, Remove Punctuations, Remove White Spaces and Non-Sindhi Words	SVM, NB, DT, KNN	SVM accuracy=64.6%
MeghaRathi et al. (2018)	1-Stanford Sentiment, 1600000 tweets 2- Polarity Dataset, 2000 reviews 3-university of Michigan 7086 sentences	English	Lowercase conversion, Converting www. Or HTTP to URLs, Convert user name to _HANDLE, Using hashtags without hash symbol #, Trimming tweets, Repeating words, Emoticons, Stemming, Tuning of parameters	SVM, ADABOOSTED, DT	DT F-measure84% Accuracy=84%
NeelamMukhtar& Mohammad Abid Khan (2018)	151 blogs, 6025	Urdu	Stop words removal	SVM, DT, KNN	KNN IBK Accuracy= 100%

3. CONCLUSION

This work has provided a comprehensive analysis of the existing literature on SA in various languages, with a specific focus on the Roman text. The studies discussed in this paper have highlighted the significance of SA in understanding public opinion, especially in the context of social media platforms. Researchers have utilized various approaches, including machine learning, DL, and other methods, to analyze sentiment in different languages. Roman Urdu SA, despite its importance as a widely spoken language, has received relatively less attention compared to other languages. However, the studies conducted so far have demonstrated the feasibility and effectiveness of applying machine learning and DL techniques to Roman Urdu SA.

4. FUTURE DIRECTIONS

There are several promising directions for future researchers in the field of Roman Urdu SA. First, exploring and developing more advanced machine learning and DL models specifically tailored for Roman Urdu text can enhance the accuracy and robustness of SA results. Additionally, incorporating contextual information and domain-specific knowledge into SA models can improve their performance in understanding the sentiment of Roman Urdu text in specific domains such as politics, entertainment, or healthcare. Furthermore, investigating the impact of linguistic features, including lexical, syntactic, and semantic information, on Roman Urdu SA can provide valuable insights into the language-specific characteristics affecting sentiment. Moreover, developing methods to handle challenges

such as sarcasm, irony, and ambiguity in Roman Urdu SA will be crucial for achieving more accurate sentiment classification.

Lastly, the creation of large-scale annotated datasets specifically for Roman Urdu SA can facilitate the development and evaluation of novel models and approaches in this domain. Overall, the future of Roman Urdu SA lies in advancing the existing methodologies, addressing language-specific challenges, and leveraging emerging techniques to extract sentiment from Roman Urdu text effectively.

REFERENCES

- [1] D. Khurana, A. Koli, K. Khatter, and S. Singh, 'Natural language processing: State of the art, current trends and challenges', *Multimedia tools and applications*, vol. 82, no. 3, pp. 3713–3744, 2023.
- [2] J. Williams, S. M. Fiore, and F. Jentsch, 'Supporting artificial social intelligence with theory of mind', *Frontiers in artificial intelligence*, vol. 5, p. 750763, 2022.
- [3] D. Khurana, A. Koli, K. Khatter, and S. Singh, 'Natural language processing: State of the art, current trends and challenges', *Multimedia tools and applications*, vol. 82, no. 3, pp. 3713–3744, 2023.
- [4] L. Khan, A. Amjad, N. Ashraf, and H.-T. Chang, 'Multi-class sentiment analysis of urdu text using multilingual BERT', *Scientific Reports*, vol. 12, no. 1, p. 5436, 2022.
- [5] M. H. Malik, S. A. Raza, and H. M. S. Asif, 'Context based emotion analyzer for interactive agent', *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 1, 2017.
- [6] B. A. Chandio, A. S. Imran, M. Bakhtyar, S. M. Daudpota, and J. Baber, 'Attention-based RU-BiLSTM sentiment analysis model for Roman Urdu', *Applied Sciences*, vol. 12, no. 7, p. 3641, 2022.
- [7] M. V. Mäntylä, D. Graziotin, and M. Kuutila, 'The evolution of sentiment analysis—A review of research topics, venues, and top cited papers', *Computer Science Review*, vol. 27, pp. 16–32, 2018.
- [8] D. D. Droba, 'Methods used for measuring public opinion', *American Journal of Sociology*, vol. 37, no. 3, pp. 410–423, 1931.
- [9] R. Stagner, 'The cross-out technique as a method in public opinion analysis', *The Journal of Social Psychology*, vol. 11, no. 1, pp. 79–90, 1940.
- [10] K. Dave, S. Lawrence, and D. M. Pennock, 'Mining the peanut gallery: Opinion extraction and semantic classification of product reviews', in *Proceedings of the 12th international conference on World Wide Web*, 2003, pp. 519–528.
- [11] A. K. Nassirtoussi, S. Aghabozorgi, T. Y. Wah, and D. C. L. Ngo, 'Text mining for market prediction: A systematic review', *Expert Systems with Applications*, vol. 41, no. 16, pp. 7653–7670, 2014.
- [12] E. Cambria, P. Gastaldo, F. Bisio, and R. Zunino, 'An ELM-based model for affective analogical reasoning', *Neurocomputing*, vol. 149, pp. 443–455, 2015.
- [13] M. A. Qureshi, M. Asif, M. F. Khan, A. Kamal, and B. Shahid, 'Roman Urdu Sentiment Analysis of Songs 'Reviews'', 2023.
- [14] S. Yasin, K. Ullah, S. Nawaz, M. Rizwan, and Z. Aslam, 'Dual Language Sentiment Analysis Model for YouTube Videos Ranking Based on Machine Learning Techniques', *Pakistan Journal of Engineering and Technology*, vol. 3, no. 2, pp. 213–218, 2020.
- [15] S. Tiwari, A. Verma, P. Garg, and D. Bansal, 'Social media sentiment analysis on Twitter datasets', in *2020 6th international conference on advanced computing and communication systems (ICACCS)*, 2020, pp. 925–927.
- [16] K. Mehmood, D. Essam, K. Shafi, and M. K. Malik, 'An unsupervised lexical normalization for Roman Hindi and Urdu sentiment analysis', *Information Processing & Management*, vol. 57, no. 6, p. 102368, 2020.
- [17] M. Asif, A. Ishtiaq, H. Ahmad, H. Aljuaid, and J. Shah, 'Sentiment analysis of extremism in social media from textual information', *Telematics and Informatics*, vol. 48, p. 101345, 2020.
- [18] S. Rani and W. Anwar, 'Resource Creation and Evaluation of Aspect Based Sentiment Analysis in Urdu', in *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing: Student Research Workshop*, 2020, pp. 79–84.
- [19] J. Lappeman, R. Clark, J. Evans, L. Sierra-Rubia, and P. Gordon, 'Studying social media sentiment using human validated analysis', *MethodsX*, vol. 7, p. 100867, 2020.
- [20] M. P. Akhter, Z. Jiangbin, I. R. Naqvi, M. Abdelmajeed, and M. T. Sadiq, 'Automatic detection of offensive language for urdu and roman urdu', *IEEE Access*, vol. 8, pp. 91213–91226, 2020.

- [21] F. H. A. Shibly, U. Sharma, and H. M. M. Naleer, 'Classifying and measuring hate speech in Twitter using topic classifier of sentiment analysis', in *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2020*, Volume 1, 2021, pp. 671–678.
- [22] A. Rafique, M. K. Malik, Z. Nawaz, F. Bukhari, and A. H. Jalbani, 'Sentiment analysis for roman urdu', *Mehran University Research Journal of Engineering & Technology*, vol. 38, no. 2, pp. 463–470, 2019.
- [23] R. Bibi, U. Qamar, M. Ansar, and A. Shaheen, 'Sentiment analysis for Urdu news tweets using decision tree', in *2019 IEEE 17th international conference on software engineering research, management and applications (SERA)*, 2019, pp. 66–70.
- [24] S. A. Salam and R. Gupta, 'Emotion detection and recognition from text using machine learning', *Int. J. Comput. Sci. Eng.*, vol. 6, no. 6, pp. 341–345, 2018.
- [25] K. Sailunaz and R. Alhajj, 'Emotion and sentiment analysis from Twitter text', *Journal of Computational Science*, vol. 36, p. 101003, 2019.
- [26] U. H. Ehsan, S. Rauf, S. Hussain, and K. Javed, 'Corpus of aspect-based sentiment for Urdu political data', in *Proc. 7th Int. Conf. Lang. Technol.*, 2020.
- [27] N. Mukhtar and M. A. Khan, 'Effective lexicon-based approach for Urdu sentiment analysis', *Artificial Intelligence Review*, vol. 53, no. 4, pp. 2521–2548, 2020.
- [28] K. Mehmood, D. Essam, and K. Shafi, 'Sentiment analysis system for Roman Urdu', in *Intelligent Computing: Proceedings of the 2018 Computing Conference*, Volume 1, 2019, pp. 29–42.
- [29] M. Khan and K. Malik, 'Sentiment classification of customer's reviews about automobiles in roman urdu', in *Advances in Information and Communication Networks: Proceedings of the 2018 Future of Information and Communication Conference (FICC)*, Vol. 2, 2019, pp. 630–640.
- [30] K. Mehmood, D. Essam, K. Shafi, and M. K. Malik, 'Sentiment analysis for a resource poor language—Roman Urdu', *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, vol. 19, no. 1, pp. 1–15, 2019.
- [31] N. Mukhtar, M. A. Khan, N. Chiragh, A. U. Jan, and S. Nazir, 'Recognition and effective handling of negations in enhancing the accuracy of Urdu sentiment analyzer', *Mehran University Research Journal Of Engineering & Technology*, vol. 39, no. 4, pp. 759–771, 2020.
- [32] K. Mehmood, D. Essam, K. Shafi, and M. K. Malik, 'Discriminative feature spamming technique for roman urdu sentiment analysis', *IEEE Access*, vol. 7, pp. 47991–48002, 2019.
- [33] A. Iram, 'Sentiment Analysis of Student's Facebook Posts', in *Intelligent Technologies and Applications: First International Conference, INTAP 2018, Bahawalpur, Pakistan, October 23-25, 2018, Revised Selected Papers 1*, 2019, pp. 86–97.
- [34] M. H. Shakeel, S. Faizullah, T. Alghamidi, and I. Khan, 'Language independent sentiment analysis', in *2019 International Conference on Advances in the Emerging Computing Technologies (AECT)*, 2020, pp. 1–5.
- [35] M. Asif, M. A. Qureshi, A. Abid, and A. Kamal, 'A dataset for the sentiment analysis of indo-pak music industry', in *2019 International Conference on Innovative Computing (ICIC)*, 2019, pp. 1–6.
- [36] N. M. Danish et al., 'Intelligent interface for fake product review monitoring and removal', in *2019 16th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*, 2019, pp. 1–6.
- [37] N. Mukhtar, M. A. Khan, and N. Chiragh, 'Effective use of evaluation measures for the validation of best classifier in Urdu sentiment analysis', *Cognitive Computation*, vol. 9, pp. 446–456, 2017.
- [38] S. M. Ghulam and T. R. Soomro, 'Twitter and Urdu', in *2018 international conference on computing, mathematics and engineering technologies (ICOMET)*, 2018, pp. 1–6.
- [39] M. Hammad and H. Anwar, 'Sentiment analysis of sindhi tweets dataset using supervised machine learning techniques', in *2019 22nd International Multitopic Conference (INMIC)*, 2019, pp. 1–6.
- [40] M. Rathi, A. Malik, D. Varshney, R. Sharma, and S. Mendiratta, 'Sentiment analysis of tweets using machine learning approach', in *2018 Eleventh international conference on contemporary computing (IC3)*, 2018, pp. 1–3.
- [41] N. Mukhtar and M. A. Khan, 'Urdu sentiment analysis using supervised machine learning approach', *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 32, no. 02, p. 1851001, 2018.
- [42] L. Khan, A. Amjad, K. M. Afaq, and H.-T. Chang, 'Deep sentiment analysis using CNN-LSTM architecture of English and Roman Urdu text shared in social media', *Applied Sciences*, vol. 12, no. 5, p. 2694, 2022.

- [43] A. A. Nagra, K. Alissa, T. M. Ghazal, S. Kukunuru, M. M. Asif, and M. Fawad, 'Deep sentiments analysis for roman urdu dataset using faster recurrent convolutional neural network model', *Applied Artificial Intelligence*, vol. 36, no. 1, p. 2123094, 2022.
- [44] Z. Mahmood et al., 'Deep sentiments in roman urdu text using recurrent convolutional neural network model', *Information Processing & Management*, vol. 57, no. 4, p. 102233, 2020.
- [45] N. Alyas, M. H. Malik, and H. Ghous, 'Sentiment analysis Using Machine Learning and Deep Learning: A Survey', 2021.
- [46] H. Rizwan, M. H. Shakeel, and A. Karim, 'Hate-speech and offensive language detection in roman Urdu', in *Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP)*, 2020, pp. 2512–2522.
- [47] L. Nemes and A. Kiss, 'Social media sentiment analysis based on COVID-19', *Journal of Information and Telecommunication*, vol. 5, no. 1, pp. 1–15, 2021.
- [48] M. A. Manzoor, S. Mamoon, S. K. Tao, Z. Ali, M. Adil, and J. Lu, 'Lexical Variation and Sentiment Analysis of Roman Urdu Sentences with Deep Neural Networks', *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 2, 2020.
- [49] H. Ghulam, F. Zeng, W. Li, and Y. Xiao, 'Deep learning-based sentiment analysis for roman urdu text', *Procedia computer science*, vol. 147, pp. 131–135, 2019.
- [50] L. Yang, Y. Li, J. Wang, and R. S. Sherratt, 'Sentiment analysis for E-commerce product reviews in Chinese based on sentiment lexicon and deep learning', *IEEE access*, vol. 8, pp. 23522–23530, 2020.
- [51] W. Xue and T. Li, 'Aspect based sentiment analysis with gated convolutional networks', *arXiv preprint arXiv:1805.07043*, 2018.
- [52] F. Mehmood, M. U. Ghani, M. A. Ibrahim, R. Shahzadi, W. Mahmood, and M. N. Asim, 'A precisely xtreme-multi channel hybrid approach for roman urdu sentiment analysis', *IEEE Access*, vol. 8, pp. 192740–192759, 2020.
- [53] R. K. Behera, M. Jena, S. K. Rath, and S. Misra, 'Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data', *Information Processing & Management*, vol. 58, no. 1, p. 102435, 2021.
- [54] Z. Nasim and S. Ghani, 'Sentiment analysis on Urdu tweets using Markov chains', *SN Computer Science*, vol. 1, no. 5, p. 269, 2020.
- [55] M. Sohail, A. Imran, H. U. Rehman, and M. Salman, 'Anti-social behavior detection in Urdu language posts of social media', in *2020 3rd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, 2020, pp. 1–7.
- [56] M. Hassan and M. Shoaib, 'Opinion within opinion: segmentation approach for urdu sentiment analysis', *Int. Arab J. Inf. Technol.*, vol. 15, no. 1, pp. 21–28, 2018.
- [57] M. U. Arshad, M. F. Bashir, A. Majeed, W. Shahzad, and M. O. Beg, 'Corpus for emotion detection on roman urdu', in *2019 22nd International Multitopic Conference (INMIC)*, 2019, pp. 1–6.
- [58] A. Alrumaih, A. Al-Sabbagh, R. Alsabah, H. Kharrufa, and J. Baldwin, 'Sentiment analysis of comments in social media', *International Journal of Electrical & Computer Engineering (2088-8708)*, vol. 10, no. 6, 2020.
- [59] M. H. Malik, H. Ghous, and S. Perveen, 'A Hierarchical Part of Speech Tag set for Saraiki Language', *Southern Journal of Arts & Humanities*, vol. 1, no. 02, pp. 1–11, 2023.
- [60] M. A. Sarwar, M. M. Hussain, S. Qadri, and Others, 'Descriptive Analysis of Human Emotions Based on Eye pupils', *Lahore Garrison University Research Journal of Computer Science and Information Technology*, vol. 6, no. 2, pp. 1–13, 2022.