

# Journal of Hunan University (Natural Sciences)

Vol. 52 No. 5

May 2025

Available online at

<https://jonuns.com>



ELSEVIER  
Scopus



Clarivate  
WEB OF SCIENCE

Open Access Article

 <https://doi.org/10.55463/issn.1674-2974.52.5.18>

## Predict Union Election Results Through Polarity and Subjectivity Detection of Multilingual Opinions

Imitaz Ali Solangi<sup>1\*</sup>, Javed Ahmed Mahar<sup>1</sup>, Ghulam Ali Mallah<sup>1</sup>, Zulfiqar Ali Solangi<sup>2</sup>

<sup>1</sup> Department of Computer Science-Shah Abdul Latif University Khairpur,

<sup>2</sup> Jubail Technical Institute, Education Sector Royal Commission Jubail

\* Corresponding author: [imtiazhphdscholar@gmail.com](mailto:imtiazhphdscholar@gmail.com)

### Article History:

*Received:* April 28, 2025

*Revised:* May 26, 2025

*Accepted:* June 3, 2025

*Published online:* June 30, 2025

**Abstract:** This study examines subjectivity and polarity in multilingual WhatsApp messages exchanged during the employees' union election at the Mehran University of Engineering and Technology (MUET), Khairpur campus. As social-media platforms generate vast volumes of unstructured multilingual text, sentiment analysis an intersection of natural language processing (NLP) and machine learning (ML) offers a rigorous means of extracting collective opinion. We introduce a novel analytical framework for code-switched discourse in Sindhi and Urdu, thereby contributing to both multilingual NLP research and electoral studies. A corpus of WhatsApp messages reflecting anticipated election outcomes was collected, cleaned, and annotated. Five supervised classifiers—feed-forward neural networks, nonlinear support vector machines, random forests, decision trees, and k-nearest neighbours were trained to detect sentiment polarity and subjectivity. Across all models, Candidate 1 received markedly more favourable sentiment, posting an average positive-sentiment score of 61.48 % versus 37.11 % for Candidate 2. Candidate 1 likewise exhibited superior aggregate performance metrics (mean accuracy = 89.92 %, recall = 86.52 %, precision = 85.12 %, F1-score = 85.81 %). These findings demonstrate the feasibility of fine-grained sentiment mining in under-resourced languages and yield actionable insights for campaign strategists. Future research will enhance classification accuracy through advanced deep-learning architectures and will track dynamic sentiment patterns across multiple social-media platforms throughout the electoral cycle.



Copyright: © 2025 by the authors. Licensee JHU

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)

**Keywords:** multilingual sentiment analysis; code-switching; WhatsApp; natural language processing; machine learning; Sindhi; Urdu; electoral behaviour.

## 通过对多语种意见的极性与主观性检测预测工会选举结果

**摘要：** 本研究针对信德语与乌尔都语交替使用的 WhatsApp 多语种讯息，分析其主观性与极性，这些讯息产生于迈赫兰工程与技术大学（MUET）海尔普尔校区职工工会选举期间。当今社交媒体不断产生庞大且非结构化的多语文本，情感分析——结合自然语言处理（NLP）与机器学习（ML）——可用以严谨提取群体观点。本文提出一套全新的分析框架，专门处理信德语与乌尔都语的交替语篇，为多语 NLP 与选举研究提供新见解。研究收集并清洗了反映预期选举结果的 WhatsApp 语料，并进行标注。随后采用五种有监督分类器——前馈神经网络、非线性支持向量机、随机森林、决策树和 k 近邻——来识别情感极性与主观性。所有模型均显示，候选人 1 获得显著更为积极的情感评价，平均正向情感比例为 61.48%，而候选人 2 为 37.11%。候选人 1 在综合性能指标上亦表现更优（平均准确率 89.92%，召回率 86.52%，精确率 85.12%，F1 值 85.81%）。结果证明，在资源匮乏语言中进行细粒度情感挖掘的可行性，并为竞选策略制定者提供了可操作的洞见。未来研究将通过先进深度学习架构提高分类准确性，并在整个选举周期内，跨多个社交平台追踪动态情感模式。

**关键词：** 多语种情感分析；语码转换；WhatsApp；自然语言处理；机器学习；信德语；乌尔都语；选举行为

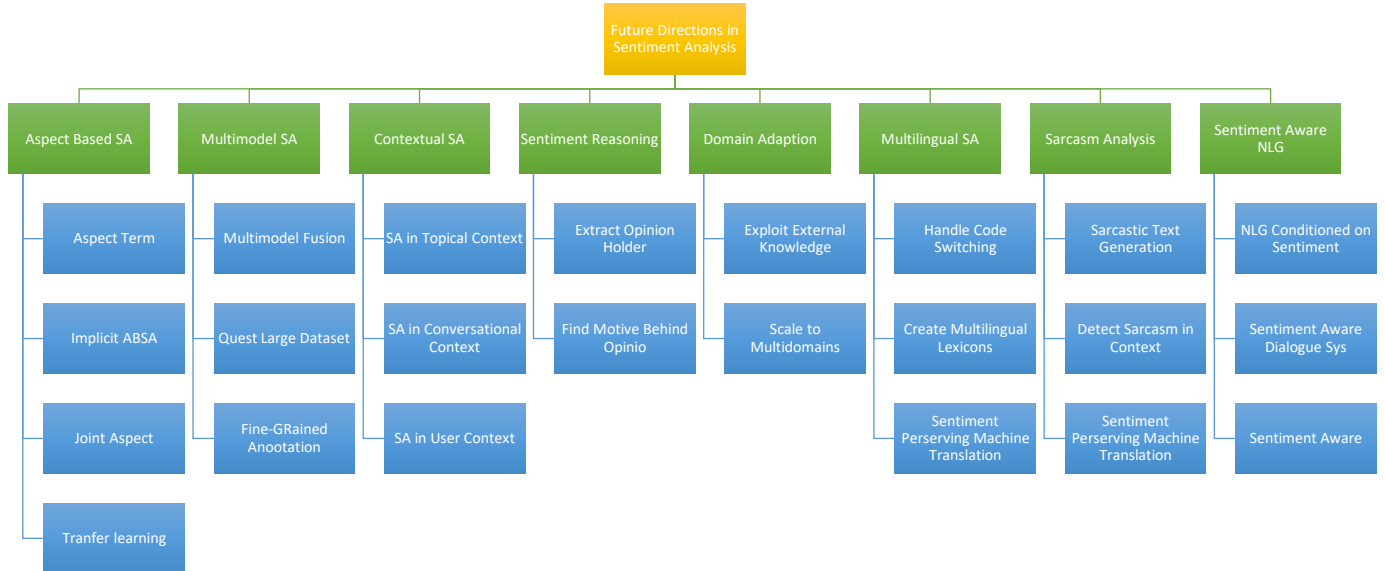
### 1. Introduction

Sentiment Analysis (SA) is a subfield of Natural Language Processing (NLP) and Machine Learning (ML) has been an important tool for the businesses and organizations to interpret and judge the sentiments and emotions of public on social media posts about their products and services. Businesses are utilizing SA to grasp and evaluate dissimilar perspectives of public in improving their services and products scope in the contemporary business race (Naithani & Raiwani, 2023). In current research works, Natural Language Processing (NLP) has become one of the vital research areas related to Artificial Intelligence, Information Technology and Computer Science that has been applied to form human-computer-interface. The major applications may include sentiment analysis, voice recognition, information extraction, Chabot, machine translations etc. (Khurana et al., 2023). NLP system needs to be linked and synchronized with real-time source of information systems and conversational context so that NLP system may easily interpret the true intent behind the words. Therefore, NLP research areas like sentiment analysis, text normalization, speech recognition, and other areas have been subject of thought (Ding et al., 2022). In addition to it, several areas of SA are listed in the Figure 1 beneath. Subjectivity and polarity detection are the initial components of Sentiment Analysis to predict the various emotional states. Subjectivity detection is to identify

biased or unbiased words in natural language processing while the polarity detection is to determine positive or negative opinion categorization in NLP (Xing et al., 2022). Business and organizations are paying a huge amount on data and information interpretation and analysis generally published and posted against their products and services on their blogs, comment boxes, or social media pages. Eventually, the insightful analysis of information and emotions help organizations to develop future business strategies and make acknowledgeable decisions (Vasquez Rojas et al., 2018). There are several studies conducted in literature suggesting techniques for sentiment analysis in categorizing emotions such as lexicon-based methods, machine learning algorithms, and hybrid models (Jassim et al., 2023). One of the important objectives of sentiment analysis is to display emotions, opinions or emotions of individuals. While polarization analysis deals with the process to divide the opinions or sentiments of individuals into positive, negative or neutral. The effectiveness of these techniques can be more efficient in sentiment classification and predictions by sharing the information between polarization and subjectivity components of natural language processing. Most of the people share their thoughts and opinions on diversity of issues and themes using social media sites like Facebook, Twitter, and WhatsApp that have produced a huge unstructured data in latest years. In order to capitalize these shared opinions and comments about different political, social,

commercial and elections events, the organizations require sentiment analysis techniques offered by NLP(Tucker et al., 2018). For example, workers frequently use digital channels like WhatsApp to voice their preferences, annoyances, and expectations during union elections. Examining the tone of these messages might reveal important information about the employees' support for certain candidates, their worries

about the election process, and their general attitudes toward union leadership(Chauhan et al., 2021a). Organizations and unions can improve member engagement, make strategic adjustments, and make well-informed decisions by performing sentiment analysis on such data. Figure 1 displays the list of prospective future progresses in SA.



**Figure 1. Taxonomy of Future Research Directions in Sentiment Analysis**

This paper focuses on performing sentiment analysis on the WhatsApp chat data exchanged during the union elections at Mehran University of Engineering and Technology (MUET), Khairpur. The data consists of multilingual text in Roman Urdu and Sindhi, reflecting the discussions, predictions, and sentiments of employees regarding the upcoming union election results. By analyzing this data using NLP techniques, we aim to uncover the underlying emotional tones expressed in the conversation, providing a deeper understanding of employee opinions. Furthermore, this research highlights the challenges and opportunities of analyzing sentiment in informal, multilingual, and unstructured text data, and demonstrates how these insights can help organizations and unions understand their stakeholders' views and concerns. To achieve this, we conducted a literature review and experiments to determine the most suitable approach. The article is structured as Introduction which contains specifics about sentiment analysis and its significance to recognize the public feelings. The role of NLP in extracting meaningful insights from WhatsApp chat of employees of MUET Khairpur. Background and Literature Review, explores existing research on natural language processing, sentiment analysis, and tools for multilingual sentiment detection using machine learning algorithms. Dataset Analysis, details

the experimental data collection, preprocessing, analysis, and sentiment prediction methods and Discussion and Conclusion, summarizes the findings, discusses implications, and concludes with potential future research directions.

## 2. Background

In the progressive exploration of sentiment analysis (SA) space, NLP academics throughout the world have proposed and implemented broad strategies, tactics, tools, and algorithms. Analyzing the sentiment expressed in text is invaluable for informed decision-making. One approach involves examining viewer responses to social media visual clips to gauge overall attitudes. However, social media comments often employ non-native scripts and deviate from traditional grammatical rules, particularly in low-resource languages. The absence of recognized code-mixed metadata exacerbates this even further. The gold standard dataset of YouTube comments is code-switched between Tamil and English along with annotations for sentiment polarity by (Raja Chakravarthi et al., 2020) in order to address the challenges of non-scripted low-sources languages. The authors reutilized the YouTube dataset and created corpus, inter-annotator settlement and sentiment analysis. Social media users share the identical Emoji and text while writing the comments and how these

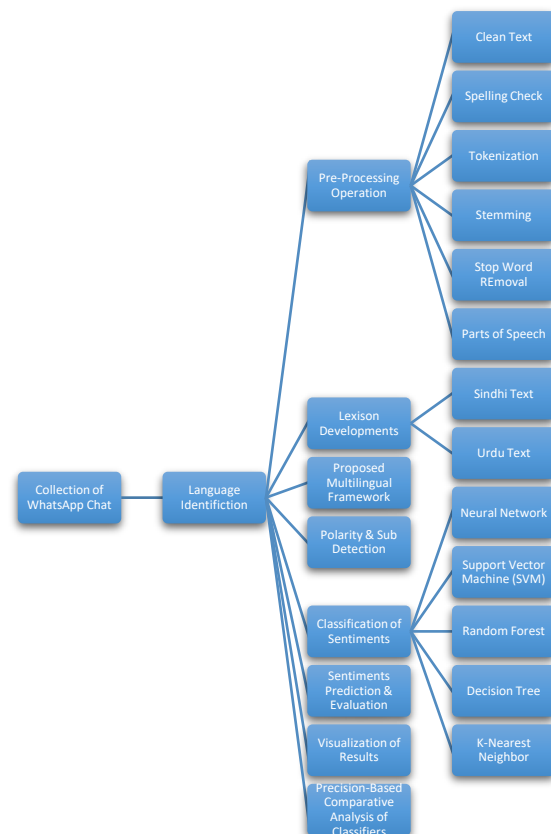
emojis and text patterns differ from man to woman is investigated by Gupta (Gupta et al., 2020). While analyzing the sentiments of people on social media the researchers have observed the both text and emojis represent a vibrant picture how the men and women use it on Twitter(X) media application. Several modern technical ways are suggested by (Devi & Kamalakkannan, 2020) to analyze sentiments on different machine learning classifiers. Kaur and Sharma investigated the feelings of people about COVID-19 virus. Both authors analyzed the positive, negative, and neutral tweets about COVID-19 in their studies using machine learning algorithms (Kaur & Sharma, 2020; Zhang et al., 2020). Prabhakar Kaila endorsed that polarity and subjectivity of these tweets seemed correct for this kind of research (Prabhakar Kaila & Prasad, 2020). Latent Dirichlet Allocation (LDA) was used as a special method for recognizing the document well. They succeeded to identify important sentiments associated with COVID-19 like how medical professionals were dealing the case, general public perception like sometimes frightened and convinced. The social media is constantly being investigated by the researchers in nearly all domains of business, politics, and marketing to return the valuable information from the public sentiments and attitude. Alzubaidi highlighted the difficulties in attaining high sentiment analysis accuracy because of the intricacy of processing text data and emojis simultaneously. Though, the use of alphanumeric characters, emojis, and cursive texts in the public comments and opinions disrupts the accuracy in sentiment analysis yet the author Alzubaidi suggested effective method of analyzing the social media sentiments having both text and emojis with inspiring accuracy rate (Saleh Alzubaidi & Bourennani, 2021).

Emotions and comments shared on social media may include wide annoying sensitive information hence, it requires vigilant use of machine learning branch, natural language processing (NLP) to comprehend the range of feelings voiced about some event, topic, or popular figure. NLP eases researchers to investigate and infer the social media posts comments as positive or negative. Additionally the polarity is classified as positive, negative, or neutral with sentiment analysis classifiers. Due to sensitivity of opinions and comments, constructive condemnation is agreed and processed whereas harsh sentiments are disregarded. Diverse sentiment analysis approaches are utilized in scientific investigations by Kumar (Kumar et al., 2020). The author has presented a detailed review of several studies about sentiment processing techniques applied in the earlier literature of Machine Learning (ML) field. Main contributions are intensively presented how ML techniques are employed to treat specific data types. In addition, performance metrics are suggested to evaluate the research data correctness and reliability. Therefore,

this research has identified the research gap to determine unmapped applications of sentiment analysis that can bring significant future contribution in the machine learning research area.

### 3. Methodology

Lexicon-based Sentiment Analysis on Sindhi-Urdu Texts on Social Media (LSASUTE) framework is processed in series of steps in this section. A multilingual research framework is created to record and examine user opinions from reviews on the social media app WhatsApp in Sindhi and Urdu. This model specifically targets the MUET Employees Union Election results, aiming to predict the outcome of union election and glean insights into employees' perceptions through the systematic analysis of WhatsApp chat messages shared on the group. The LSASUTE model uses a word-based method, relying on a self-made and manually edited word list that contains about 50,000 sentences. This large word list is the base for sentiment analysis (SA), helping to find and sort feelings shown in WhatsApp group chats about Union Election results. The steps of the LSASUTE model are shown in Figure 2, which explains the process from collecting data, multilingual (Sindhi and Urdu) identification, preprocessing steps, lexicon development, polarity detection, and deployment of classifiers, prediction of sentiments about the Union election candidates, visualization of election results and finally comparative analysis of results of all classifiers.



## Figure 2. Workflow of the Proposed Multilingual Sentiment-Analysis for WhatsApp Election Discourse

### 3.1 Data Collection

The election data of employee at MUET Jamshoro and SZABSZAB Khairpur is primarily utilized in this study. The WhatsApp chat group data of employees was shared on request of the authors. The research study evaluated the chat messages of the WhatsApp group shared during the Union election of the employees. This research has tailed the data collection method offered by (Raja Chakravarthi et al., 2020). To ensure there was enough data for research, two employee unions each developed their own datasets. Chat conversations from the most recent and upcoming union elections yielded more than 50,000 lines in Sindhi and Urdu. Language identification software was used to identify and select proper language chat shared in multiple languages, as advised by (Chandio, Asikuzzaman, Pickering, et al., 2020). The chat messages underwent preprocessing, which involved removing reactions and applying a filter to extract the sentences. Sindhi and Urdu language chat is displayed in Figures 3 and 4 respectively. With four to seven words per sentence, the Sindhi and Urdu sentences were deconstructed into individual words. To analyze the polarity and subjectivity of university employee sentiments, over 250,000 words were extracted from the collected phrases.



Figure 1 WhatsApp message written in Urdu language (Source: MEUT WhatsApp group)



Figure 2 WhatsApp message in Sindhi language with symbols, and emoji (Source: MEUT WhatsApp group)

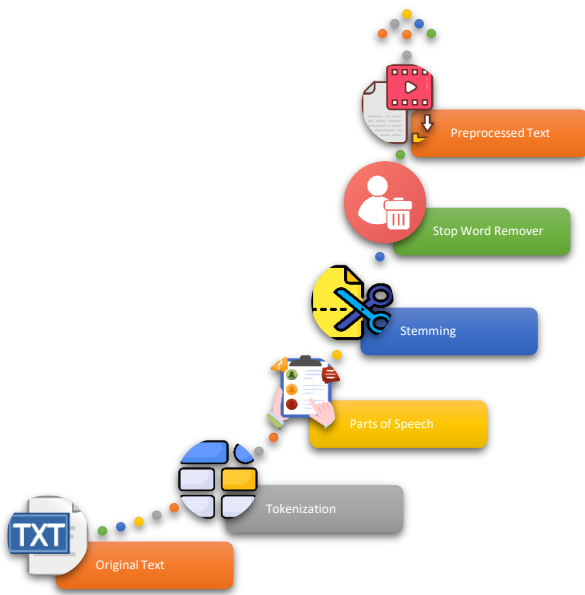
### 3.2 Data Processing

Given that the short messages are written in Sindhi and Urdu, language identification using simulation software is necessary as suggested by (Humayoun et al., 2016). The TextCat software program is utilized to determine the dialect (Chandio, Asikuzzaman, & Pickering, 2020; Nadeem et al., 2022). The chosen and identified words that denote polarity or subjectivity are next subjected to preprocessing processes. Text cleaning is the first preliminary process, which eliminates punctuation, standardizes capitalization, and transforms numerical expressions into regular text. A tool is needed to check the spelling of information authored in Sindhi and Urdu in order to address the prevalent misspelling problems on social media. The word processor recommended by (Dahar et al., 2018) is utilized for this purpose. This study utilizes the tokenization model proposed by (Mahar et al., 2012) to analyze employee attitudes based on their text messages, facilitating sentiment analysis of university employee communications.

Table 1. Statistics of the word instances for each script

Script type	No. of Words
Urdu	75033
Sindhi	17379
Symbols	1055
Others	117

Sindhi and Urdu, renowned for their linguistic richness, heavily utilize prefixes and suffixes. Implementing a robust stemming mechanism is essential reused from (Shah et al., 2016). Following the stemming process, stop word removal is implemented. This step aims to eliminate special characters, generic terms, unnecessary words, and meaningless words.



**Figure 3. Pre-processing tasks (developed by the authors)**

### 3.3. Lexicon & Proposed Multilingual Sentiment Analysis

Following the preprocessing steps, lexicons for Sindhi and Urdu texts are generated independently, as subsequent components rely heavily on these lexicons. The development and recommendation Lexicon-based SA on Sindh-Urdu Texts on Social Media (LSASUTE) is a key component of this study. Lexicon development is a crucial aspect of sentiment analysis, especially when using lexicon-based approaches (Alsemaree et al., 2024). Using a dictionary-based algorithm to find the emotive terms in chat\_text is a common tactic when working with sentiment analysis. (Martins et al., 2020). Each dictionary-based algorithm's primary drawback is that the learned vocabulary is domain-independent, meaning it fails to capture the unique characteristics of each given domain. In order to complete the sentiment analysis of Sindhi lexicon, the author expanded and reutilized the Sindhi Text sentiment analysis dataset from Kaggle and Large Movie Review Dataset (Maas et al., n.d.) is also expanded and reused as Urdu Sentiment Analysis dataset from Kaggle. Sindhi dataset contains 3366 unique data values and the Urdu dataset contains 9997 unique labels. Both datasets were expanded with proper guidance and support of knowledgeable researchers of machine learning. As it is noted earlier that data was collected through the WhatsApp chat groups. Accordingly, the study managed to gather 1000 WhatsApp chat messages detailed in Table 1 with 75033 Urdu, 17379 Sindhi words, 1055 emoji characters with happy and sad related feelings and 117 some other symbols of them labelled, organized the

documented messages of the chat. Some unidentified words were not annotated.

### 3.4 Sentiment Polarity and Subjectivity

The TextBlob Python library was utilized to work with multilingual (Sindhi and Urdu) categorization, classification, and development of lexical resources of Sindhi and Urdu language. Following the assignment of a score to each word, the final sentiment was ascertained using pooling techniques, such as averaging all sentiments (Vel, 2021). By calculating the polarity between  $[-1,1]$ , where -1 indicates a negative attitude and 1 indicates a positive mood, TextBlob can define the subjectivity and polarity of texts written in Sindhi and Urdu. Subjectivity was determined as personal opinions, emotions or factual information contained in the chat text which lied between  $[0,1]$  (Abiola et al., 2023).

## 4. Model Training for Classification of Sentiments

To categorize employee attitudes, five widely recognized machine learning algorithms were employed: Neural Network, Support Vector Machine, Random Forest, Decision Tree, and K-Nearest Neighbor. These classification methods are readily available in most modern high-level programming languages. Five different classifiers were evaluated using recommended following calculations, and the standard metrics were observed using a confusion matrix. Prediction was performed using five representative ML algorithms: Neural Networks, Support Vector Machine (Nonlinear), Random Forest, Decision Tree, and K-Nearest Neighbor (KNN). Neural network models drive numerous complex machine learning applications, with classification and sentiment analysis being prominent examples of their current use. We used Support Vector Machines (SVM), a potent supervised machine learning method, for both regression and binary classification tasks. We used SVM for tasks that involved binary classification and regression. SVM is an effective supervised machine learning method. By aggregating the predictions of multiple decision trees through a majority vote (the most common predicted value), the bagging-based ensemble model Random Forest (RF) achieves its final prediction. This method results in a robust and interpretable model. The ease of comprehension and interpretation of Decision Trees, coupled with the fact that no machine learning experience is needed, makes them a wise choice specifically for situations where the ability to explain the model's predictions is paramount (Bansal et al., 2022). We chose the K-Nearest Neighbor (K-NN) algorithm due to its ease of implementation, making it beginner-friendly. Moreover, its effectiveness and fundamental role in data segregation make it a prime candidate for use, especially when the data exhibits significant

ambiguity.

- Accuracy: The overall perfection of the framework.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- Precision (Positive Predictive Value): Out of all the instances the model predicted as positive, what proportion were actually positive?

$$Precision = \frac{TP}{TP + FP}$$

- Recall (Sensitivity, True Positive Rate): Out of all the actual positive instances, what proportion did the model correctly identify?

$$Recall = \frac{TP}{FN + TP}$$

- Specificity (True Negative Rate): Out of all the actual negative instances, what proportion did the model correctly identify?

$$Specificity = \frac{TN}{TP + FN}$$

- F1-Score: The harmonic mean of precision value and recall value, causing a balanced measure of the two.

$$F1 - Score = 2x = \frac{Precision \times Recall}{Precision + Recall} = \frac{2TP}{2TP + FP + FN}$$

- False Positive Rate (FPR): Out of all the actual negative instances, what proportion did the model incorrectly classify as positive?

$$FPR = \frac{FP}{TN+FP} = 1 - Specificity$$

- False Negative Rate (FNR): Out of all the actual positive instances, what proportion did the model incorrectly classify as negative?

$$FNR = \frac{FN}{TP + FN} = 1 - Recall$$

## 5. Implementation and Results

Based on the polarity score, the sentiment is categorized as 'Negative', or 'Positive', and the results are stored in three new columns: 'TextBlob\_Subjectivity', 'TextBlob\_Polarity', and 'TextBlob\_Analysis'. The sentiment analysis is then applied to the Dataset by calling the sentiment\_analysis function. A Python code is executed to train the proposed model with several effective Python libraries introduced beneath.

```
def sentiment_analysis(W_chat):
def getSubjectivity(text):
    return TextBlob(text).sentiment.subjectivity

#Create a function to get the polarity
def getPolarity(text):
    return TextBlob(text).sentiment.polarity

#Create two new columns 'Subjectivity' & 'Polarity'
W_chat['TextBlob_Subjectivity'] = W_chat['comment'].apply(getSubjectivity)
W_chat ['TextBlob_Polarity'] = W_chat['comment'].apply(getPolarity)
def getAnalysis(score):
    if score < 0:
        return 'Negative'
    elif score == 0:
        return 'Neutral'
    else:
        return 'Positive'
W_chat ['TextBlob_Analysis'] = W_chat ['TextBlob_Polarity'].apply(getAnalysis)
return W_chat
```

Figure 4. Polarity & Subjectivity Python code

### Algorithm 1: Steps to implement Classification Employee Sentiments

1. Importation of Sentiment\_Analyzer from TextBlob Sentiment for text sentiment analysis.
2. Define sentiment\_scores() to input a sentence and calculate sentiment scores using TextBlob
3. Creation of an instance of Sentiment\_Analyzer for sentiment score computation.
4. Apply polarity\_scores() and generating dictionary of sentiments with positive, negative, and neutral.
5. Present the dictionary of sentiment with percentage of negative, neutral, and positive sentiments.
6. Inclusive classification of sentiments based on composite score: positive, negative, or neutral.

## 6. Sentiment Analysis

Employee sentiments regarding the specific election forecast were evaluated based on the five different machine learning model performance of the sentiment analysis. The research experiment utilized both tabular and graphical representations to visualize the anticipated election results. Furthermore, various classification techniques were employed to categorize attitudes, assessing the classifier's effectiveness on Arabic script-based texts like Sindhi and Urdu.

### 6.1 Comparative analysis of classifiers

A comparative accuracy study of the selected five different machine learning models: Research was conducted using neural networks, SVM (Nonlinear), Random Forest, Decision Tree, and KNN approaches to determine the optimal approach for sentiment analysis using Sindhi and Urdu text. Figure 2 explains the comprehensive procedure for identifying subjectivity and polarity, categorizing sentiment, and predicting the elections results.

Table 2. Performance of Supervised Machine Learning Models on Multilingual SA (Source: Authors)

Model	Candidate #1	Candidate #2	Candidate #1	Candidate #2	Candidate #1 Recall	Candidate #2 Recall	Candidate #1	Candidate #2	Candidate #1	Candidate #2
	Accuracy	Accuracy	Precision	Precision			F1 Score	F1 Score	Specificity	Specificity

Neural Networks	90.74%	90.00%	85.58%	85.58%	88.11%	88.11%	86.82%	86.82%	91.67%	91.10%
SVM (Nonlinear)	91.11%	90.37%	86.45%	86.15%	88.56%	87.93%	87.50%	87.03%	92.00%	91.08%
Random Forest	89.63%	88.52%	84.92%	83.91%	85.80%	84.95%	85.36%	84.43%	90.98%	90.00%
Decision Tree	89.99%	88.89%	85.14%	84.16%	86.02%	85.00%	85.58%	84.58%	91.20%	90.61%
KNN	88.15%	87.41%	83.52%	82.83%	84.10%	83.67%	83.81%	83.25%	90.12%	89.55%

Across the evaluated models with training (0.80) and testing (0.20) data, Candidate #1 demonstrates slightly superior and more consistent sentiment classification accuracy compared to Candidate #2 in Table 2. Notably, SVM (Nonlinear) and Neural Networks deliver the most consistently high accuracy for both candidates. In terms of Precision, Candidate #1 consistently outperforms Candidate #2, indicating fewer instances of false positives, while Candidate #2 shows slightly lower Recall, suggesting a higher likelihood of misclassifying actual positive sentiments as negative. This balance between precision and recall is returned in the F1-scores, where Candidate #1 maintains higher values across all models, indicating a stronger overall performance, though Candidate #2's scores are closely matched. Finally, Candidate #1 exhibits higher specificity, demonstrating a better ability to identify negative sentiments, although Candidate #2 also shows good, albeit slightly weaker, performance in this area across the different models. Overall, Candidate #1 performs slightly better, exhibiting higher accuracy, precision, and F1-scores across the learning models, while Candidate #2's sentiment classification could be improved by fine-tuning the models to achieve a better balance between precision and recall.

**Table 3. Comparison Table of Candidate#1 performed with different models (Source: Authors)**

<i>Learning Model</i>	<i>Positive Sentiments</i>	<i>Positive (%)</i>	<i>Sentiment Analyzed</i>
Neural Networks	<b>165</b>	<b>61.11%</b>	<b>270</b>
SVM (Nonlinear)	<b>166</b>	<b>61.48%</b>	<b>270</b>
Random Forest	<b>164</b>	<b>60.74%</b>	<b>270</b>
Decision Tree	<b>164</b>	<b>60.74%</b>	<b>270</b>
KNN	<b>171</b>	<b>63.33%</b>	<b>270</b>

A comparison of Candidate #1's positive and negative sentiment categorization using five distinct learning models—Neural Networks, SVM (Nonlinear), Random Forest, Decision Tree, and KNN—is shown in Table 3. The table clearly counts the number of positive and negative sentiments found for each model, as well as the percentages of each

sentiment out of all the classifications the model made. This makes it possible to compare directly how well each learning model identified the sentiment toward Candidate #1. Positive sentiment percentages are similar for the Neural Networks and SVM (Nonlinear) models, both of which are marginally higher than 61%. Similarly, the Random Forest and Decision Tree models show a close alignment in their sentiment classifications, each reporting 60.74% positive and 39.26% negative sentiments. Notably, the KNN model identifies the highest proportion of positive sentiments at 63.33%, suggesting it classified a greater number of messages as positive in comparison to the other evaluated models for Candidate #1.

**Table 4. Comparison of Candidate #2 performed with different models (Source: Authors)**

<i>Learning Model</i>	<i>Negative Sentiments</i>	<i>Negative (%)</i>	<i>Sentiments Analyzed</i>
Neural Networks	105	<b>38.89%</b>	270
SVM (Nonlinear)	104	<b>38.52%</b>	270
Random Forest	106	<b>39.26%</b>	270
Decision Tree	106	<b>39.26%</b>	270
KNN	99	<b>36.67%</b>	270

Table 4 shows results analyzed across different learning models for Candidate #2, provides a detailed breakdown of the sentiment classification results obtained using five different learning models: Neural Networks, SVM (Nonlinear), Random Forest, Decision Tree, and KNN. For each model, the table clearly presents the total count of positive sentiments identified, the percentage of positive sentiments relative to the total predictions, the total count of negative sentiments identified, and the corresponding percentage of negative sentiments. This structured format enables a direct comparison of how each learning model categorized sentiments expressed towards Candidate #2, highlighting the distribution between positive and negative classifications for each approach. The Random Forest and Decision tree models stand out by identifying the highest percentage of negative sentiment (39.26%) for Candidate #2,

indicating a potentially stronger ability to recognize non-supportive messages. Notably, the SVM (Nonlinear) and Neutral Network models exhibit similar sentiment classification patterns, each reporting 38.52% and 38.89% negative sentiments. Meanwhile, the KNN model reports the lowest negative sentiment recognition rate at 36.67%. Overall, while the models display a general consistency in the distribution of sentiments, and the Random Forest and Decision tree model lean towards identifying sentiment that is more negative for Candidate #2.

## 7. Observation

The aforementioned method was used to analyze the datasets. The outcomes were exciting. Two election candidates in Table 5 were used to introduce the testing. Those are Mr. Candidate#1 and his competitor Candidate#2. Both datasets of candidates were processed with comparable investigative measures. Based on the analysis of sentiment metrics across all models, Candidate#1 emerges as the clear winner due to its consistent performance in positive sentiment recognition, with an average of 61.48%, compared to Candidate#2's 37.11%. Candidate#1 demonstrates stronger key performance metrics, including Recall, Precision, and F1-Score, across Neural Networks, SVM (Nonlinear), Random Forest, Decision Tree, and KNN models, achieving a better balance between minimizing misclassifications and accurately identifying positive sentiments. Additionally, Candidate#1 exhibits slightly higher specificity in negative sentiment classification and overall accuracy, ensuring more reliable sentiment detection and robust classification capabilities.

**Table 5. Comparison between actual MUET Employees final result vs. Model predicted results**

MEUT Employees Election 2023	Candidate#1	Candidate#2
Actual elections final results	63.58%	34.55%
Predicted election results through proposed sentiment analysis	61.48%	37.11%
Accuracy error percentage	3.30%	6.89%

## 8. Discussion

The results for the MUET Employees' election are shown in a table 5. It tells us that candidate #1 got 63.58% of the votes, while candidate #2 got 34.55%. The table also shows that a special model used to guess the election results worked well. It predicted that candidate #1 would have support from 61.48% of the university staff. When the real votes came in, there was a small difference of 3.36% from what the model predicted. Nowadays, social media is really important for figuring out how people feel, especially when

trying to guess how many votes someone might get before the actual voting happens (Chauhan et al., 2021b). This suggested paradigm might also be helpful to politicians, particularly if they need to comprehend their supporters and followers. It will assist in compiling the community's and public's unfavorable views, which will have a big influence. This way, electable candidates can utilize the data as a tool to improve their campaign and increase their vote totals (Khan et al., 2023). However, this paradigm is widely applied in many contexts, including tracking voters' social media activity to learn about their opinions, convictions, and judgment.

## 9. Conclusion & Future Work

This study predicted the results of the MUET Employees' Union election by doing a rigorous emotional analysis of collected WhatsApp conversations. This research study developed and proposed the novel research framework for sentiment analysis of multilingual WhatsApp discussions in Sindhi and Urdu and contributing to both multilingual NLP developments and political analysis. The MUET Employee WhatsApp group supplied the information. Word cloud counts and a sentimental analysis based on polarity were applied to the collected conversation data. Separate files were produced for Sindhi and Urdu texts. The text was categorized using Neural Networks, SVM (Nonlinear), Random Forest, Decision Tree, and KNN based on sentiment analysis, both positive and negative. In order to better comprehend voters' feelings, future research will concentrate on enhancing sentiment analysis's accuracy by using advanced machine learning algorithms and examining dynamic sentiment patterns on different social media platforms throughout election campaigns.

## Declarations

### Author Contributions

Conceptualization, Imtiaz Ali Solangi and Zulfiqar Ali Solangi; methodology, Imtiaz Ali Solangi and Javed Ahmed Mahar; software, Ghulam Ali Mallah; validation, Imtiaz Ali Solangi and Javed Ahmed Mahar; formal analysis, Ghulam Ali Mallah; investigation, Imtiaz Ali Solangi, Javed Ahmed Mahar, and Ghulam Ali Mallah; resources, Imtiaz Ali Solangi; data curation, Imtiaz Ali Solangi; writing—original draft preparation, all authors contributed equally; writing—review and editing, Imtiaz Ali Solangi; visualization, Ghulam Ali Mallah; supervision, Imtiaz Ali Solangi; project administration, Imtiaz Ali Solangi. All authors have

read and agreed to the published version of the manuscript.

### Data Availability Statement

The datasets generated or analysed during the current study are available from the corresponding author on reasonable request.

### Funding

No external funding was received for this research.

### Institutional Review Board Statement

The study was conducted in accordance with the Declaration of Helsinki and approved by the Review Committee of the Research Center, Politeknik Negeri Media Kreatif, Jakarta, Indonesia.

### Conflicts of Interest

The authors declare no conflict of interest regarding the publication of this manuscript.

## References

- [1] Abiola, O.; Abayomi-Alli, A.; Tale, O.A.; Misra, S.; Abayomi-Alli, O. Sentiment Analysis of COVID-19 Tweets from Selected Hashtags in Nigeria Using VADER and Text Blob Analyser. *J. Electr. Syst. Inf. Technol.* 2023, 10, 5. <https://doi.org/10.1186/s43067-023-00070-9>
- [2] Alsemaree, O.; Alam, A.S.; Gill, S.S.; Uhlig, S. Sentiment Analysis of Arabic Social Media Texts: A Machine Learning Approach to Deciphering Customer Perceptions. *Heliyon* 2024, 10, e27863. <https://doi.org/10.1016/j.heliyon.2024.e27863>
- [3] Bansal, M.; Goyal, A.; Choudhary, A. A Comparative Analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short-Term Memory Algorithms in Machine Learning. *Decis. Anal. J.* 2022, 3, 100071. <https://doi.org/10.1016/j.dajour.2022.100071>
- [4] Chandio, A.A.; Asikuzzaman, M.; Pickering, M.R. Cursive Character Recognition in Natural Scene Images Using a Multilevel Convolutional Neural Network Fusion. *IEEE Access* 2020, 8, 109054–109070. <https://doi.org/10.1109/ACCESS.2020.3001605>
- [5] Chandio, A.A.; Asikuzzaman, M.; Pickering, M.R.; Leghari, M. Cursive-Text: A Comprehensive Dataset for End-to-End Urdu Text Recognition in Natural Scene Images. *Data Brief* 2020, 31, 105749. <https://doi.org/10.1016/j.dib.2020.105749>
- [6] Chauhan, P.; Sharma, N.; Sikka, G. The Emergence of Social Media Data and Sentiment Analysis in Election Prediction. *J. Ambient Intell. Humaniz. Comput.* 2021a, 12, 2601–2627. <https://doi.org/10.1007/s12652-020-02423-y>
- [7] Chauhan, P.; Sharma, N.; Sikka, G. The Emergence of Social Media Data and Sentiment Analysis in Election Prediction. *J. Ambient Intell. Humaniz. Comput.* 2021b, 12, 2601–2627. <https://doi.org/10.1007/s12652-020-02423-y>
- [8] Dahar, I.A.; Abbas, F.; Rajput, U.; Hussain, A.; Azhar, F. An Efficient Sindhi Spelling Checker for Microsoft Word. *Int. J. Comput. Sci. Netw. Secur.* 2018, 18, 144–150.
- [9] Devi, G.D.; Kamalakkannan, S. Literature Review on Sentiment Analysis in Social Media: Open Challenges toward Applications. *Test Eng. Manag.* 2020, 83, 2466–2474.
- [10] Ding, Y.; Ma, J.; Luo, X. Applications of Natural Language Processing in Construction. *Autom. Constr.* 2022, 136, 104169. <https://doi.org/10.1016/j.autcon.2022.104169>
- [11] Gupta, S.; Singh, A.; Ranjan, J. Sentiment Analysis: Usage of Text and Emoji for Expressing Sentiments. In *Advances in Data and Information Sciences; Lecture Notes in Networks and Systems*, 94; Springer: Singapore, 2020; pp. 477–486. [https://doi.org/10.1007/978-981-15-0694-9\\_45](https://doi.org/10.1007/978-981-15-0694-9_45)
- [12] Humayoun, M.; Nawab, R.M.A.; Uzair, M.; Aslam, S.; Farzand, O. Urdu Summary Corpus. In *Proceedings of (Conference)*, 2016; pp. 796–800.
- [13] Jassim, M.A.; Abd, D.H.; Omri, M.N. A Survey of Sentiment Analysis from Film Critics Based on Machine Learning, Lexicon and Hybridization. *Neural Comput. Appl.* 2023, 35, 9437–9461. <https://doi.org/10.1007/s00521-023-08359-6>
- [14] Kaur, C.; Sharma, A. Sentiment Analysis of Tweets on Social Issues Using Machine Learning Approach. *Int. J. Adv. Trends Comput. Sci. Eng.* 2020, 9, 6303–6311. <https://doi.org/10.30534/ijatcse/2020/310942020>
- [15] Khan, A.; Boudjellal, N.; Ahmad, A.; Khan, M. From Social Media to Ballot Box: Leveraging Location-Aware Sentiment Analysis for Election Predictions. *Comput. Mater. Contin.* 2023, 77, 3037–3055. <https://doi.org/10.32604/cmc.2023.044403>
- [16] Khurana, D.; Koli, A.; Khatker, K.; Singh, S. Natural Language Processing: State of the Art, Current Trends and Challenges. *Multimed. Tools Appl.* 2023, 82, 3713–3744. <https://doi.org/10.1007/s11042-022-13428-4>
- [17] Kumar, S.; Dhathri Devi, N.; Krishnendhu, T.; Nerthu, K.; Radhakrishnan, S. Review of Sentiment Analysis: A Multilingual Approach. *Int. J. Adv. Res. Comput. Commun. Eng.* 2020, 9, 53–58. <https://doi.org/10.17148/IJARCCCE.2020.9110>
- [18] Maas, A.L.; Daly, R.E.; Pham, P.T.; Huang, D.; Ng, A.Y.; Potts, C. Learning Word Vectors for Sentiment Analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational*

*Linguistics*; ACL: Portland, OR, USA, 2011; pp. 142–150.

[19] Mahar, J.; Shaikh, H.; Memon, G. A Model for Sindhi Text Segmentation into Word Tokens. *Sindh Univ. Res. J. (Sci. Ser.)* 2012, 44, —.

[20] Martins, R.; Almeida, J.; Henriques, P.; Novais, P. Predicting an Election's Outcome Using Sentiment Analysis. In *Trends and Innovations in Information Systems and Technologies*; Rocha, Á. et al., Eds.; Springer: Cham, 2020; Volume 1159, pp. 134–143. [https://doi.org/10.1007/978-3-030-45688-7\\_14](https://doi.org/10.1007/978-3-030-45688-7_14)

[21] Nadeem, M.A.; Bukhari, S.H.H.; Arshad, M.U.; Naeem, S.; Beg, M.O.; Shahzad, W. Language Detection and Localization for Pakistani Languages in Acoustic Channels. In *Proceedings of the 17th International Conference on Emerging Technologies (ICET)*; Islamabad, 2022; pp. 142–147. <https://doi.org/10.1109/ICET56601.2022.10004691>

[22] Naithani, K.; Raiwani, Y.P. Realization of Natural Language Processing and Machine Learning Approaches for Text-Based Sentiment Analysis. *Expert Syst.* 2023, 40, e13114. <https://doi.org/10.1111/exsy.13114>

[23] Prabhakar Kaila, D.R.; Prasad, D.A. Informational Flow on Twitter—Corona Virus Outbreak—Topic Modelling Approach. *Int. J. Adv. Res. Eng. Technol.* 2020, 11, 128–134.

[24] Raja Chakravarthi, B.; Muralidaran, V.; Priyadharshini, R.; McCrae, J.P. Corpus Creation for Sentiment Analysis in Code-Mixed Tamil–English Text. *arXiv* 2020, arXiv:2006.

[25] Saleh Alzubaidi, M.; Bourennani, F. Sentiment Analysis of Tweets Using Emojis and Texts. In *Proceedings of the 4th International Conference on Information Science and Systems (ICISS)*; ACM: New York, NY, USA, 2021; pp. 111–114. <https://doi.org/10.1145/3459955.3460608>

[26] Shah, M.; Shaikh, H.; Mahar, J.; Mahar, S. Sindhi Stemmer for Information Retrieval System Using Rule-Based Stripping Approach. *Sindh Univ. Res. J. (Sci. Ser.)* 2016, 48, —.

[27] Tucker, J.A.; Guess, A.; Barberá, P.; Vaccari, C.; Siegel, A.; Sanovich, S.; Stukal, D.; Nyhan, B. Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature. *SSRN Electron. J.* 2018, 95 pp. <https://doi.org/10.2139/ssrn.3144139>

[28] Vásquez Rojas, C.; Roldán Reyes, E.; Aguirre y Hernández, F.; Cortés Robles, G. Integration of a Text Mining Approach in the Strategic Planning Process of Small and Medium-Sized Enterprises. *Ind. Manag. Data Syst.* 2018, 118, 745–764. <https://doi.org/10.1108/IMDS-01-2017-0029>

[29] Vel, S.S. Pre-Processing Techniques of Text Mining Using Computational Linguistics and Python Libraries. In *Proceedings of the 2021 International Conference on Artificial Intelligence and Smart*

*Systems (ICAIS)*; Coimbatore, India, 2021; pp. 879–884.

<https://doi.org/10.1109/ICAIS50930.2021.9395924>

[30] Xing, Y.; Wang, X.; Qiu, C.; Li, Y.; He, W. Research on Opinion Polarization by Big Data Analytics Capabilities in Online Social Networks. *Technol. Soc.* 2022, 68, 101902. <https://doi.org/10.1016/j.techsoc.2022.101902>

[31] Zhang, J.; Wang, W.; Lu, C.; Wang, J.; Sangaiah, A.K. Lightweight Deep Network for Traffic Sign Classification. *Ann. Telecommun.* 2020, 75, 369–379. <https://doi.org/10.1007/s12243-019-00731-9>

#### 参考文献:

- [1] Abiola O.; Abayomi-Alli A.; Tale O. A.; Misra S.; Abayomi-Alli O. 使用 VADER 与 TextBlob 分析器对尼日利亚选定话题标签的 COVID-19 推文进行情感分析[J]. *Journal of Electrical Systems and Information Technology*, 2023, 10(1): 5. DOI: 10.1186/s43067-023-00070-9
- [2] Alsemaree O.; Alam A. S.; Gill S. S.; Uhlig S. 阿拉伯语社交媒体文本情感分析: 机器学习视角下的客户感知解析[J]. *Heliyon*, 2024, 10(9): e27863. DOI: 10.1016/j.heliyon.2024.e27863
- [3] Bansal M.; Goyal A.; Choudhary A. K. 近邻、遗传算法、支持向量机、决策树与长期记忆算法在机器学习中的比较分析[J]. *Decision Analytics Journal*, 2022, 3: 100071. DOI: 10.1016/j.dajour.2022.100071
- [4] Chandio A. A.; Asikuzzaman M.; Pickering M. R. 基于多级卷积神经网络融合的自然场景图像草写字符识别[J]. *IEEE Access*, 2020, 8: 109054–109070. DOI: 10.1109/ACCESS.2020.3001605
- [5] Chandio A. A.; Asikuzzaman M.; Pickering M. R.; Leghari M. Cursive-Text : 用于自然场景图像端到端乌尔都语文本识别的综合数据集[J]. *Data in Brief*, 2020, 31: 105749. DOI: 10.1016/j.dib.2020.105749
- [6] Chauhan P.; Sharma N.; Sikka G. 社交媒体数据与情感分析在选举预测中的兴起[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2021a, 12: 2601–2627. DOI: 10.1007/s12652-020-02423-y
- [7] Chauhan P.; Sharma N.; Sikka G. 社交媒体数据与情感分析在选举预测中的兴起[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2021b, 12: 2601–2627. DOI: 10.1007/s12652-020-02423-y
- [8] Dahar I. A.; Abbas F.; Rajput U.; Hussain A.; Azhar F. 用于 Microsoft Word 的高效信德语拼写检查器[J]. *International Journal of Computer Science and Network Security*, 2018, 18: 144–150.

- [9] Devi G. D.; Kamalakkannan S. 社交媒体情感分析文献综述：面向应用的开放挑战[J]. *Test Engineering and Management*, 2020, 83(7): 2466–2474.
- [10] Ding Y.; Ma J.; Luo X. 自然语言处理在建筑领域的应用[J]. *Automation in Construction*, 2022, 136: 104169. DOI: 10.1016/j.autcon.2022.104169
- [11] Gupta S.; Singh A.; Ranjan J. 情感分析：文本与表情符号在情感表达中的应用[C]// *Advances in Data and Information Sciences*. Singapore: Springer, 2020: 477–486. DOI: 10.1007/978-981-15-0694-9\_45
- [12] Humayoun M.; Nawab R. M. A.; Uzair M.; Aslam S.; Farzand O. Urdu 摘要语料库[C]// (会议论文集), 2016: 796–800.
- [13] Jassim M. A.; Abd D. H.; Omri M. N. 基于机器学习、词典和混合方法的影评情感分析综述[J]. *Neural Computing and Applications*, 2023, 35: 9437–9461. DOI: 10.1007/s00521-023-08359-6
- [14] Kaur C.; Sharma A. 利用机器学习方法对社会问题相关推文进行情感分析[J]. *International Journal of Advanced Trends in Computer Science and Engineering*, 2020, 9: 6303–6311. DOI: 10.30534/ijatcse/2020/310942020
- [15] Khan A.; Boudjellal N.; Ahmad A.; Khan M. 从社交媒体到投票箱：基于位置感知的情感分析用于选举预测[J]. *Computers, Materials & Continua*, 2023, 77: 3037–3055. DOI: 10.32604/cmc.2023.044403
- [16] Khurana D.; Koli A.; Khatter K.; Singh S. 自然语言处理：现状、趋势与挑战[J]. *Multimedia Tools and Applications*, 2023, 82: 3713–3744. DOI: 10.1007/s11042-022-13428-4
- [17] Kumar S.; Dhathri Devi N.; Krishnendhu T.; Nerthu K.; Radhakrishnan S. 情感分析综述：多语种方法[J]. *International Journal of Advanced Research in Computer and Communication Engineering*, 2020, 9: 53–58. DOI: 10.17148/IJARCCCE.2020.9110
- [18] Maas A. L.; Daly R. E.; Pham P. T.; Huang D.; Ng A. Y.; Potts C. 用于情感分析的词向量学习[C]// 第49届 ACL 年会论文集. Portland (USA): ACL, 2011: 142–150.
- [19] Mahar J.; Shaikh H.; Memon G. 信德语文本分词模型. *Sindh University Research Journal (Science Series)*, 2012, 44(1).
- [20] Martins R.; Almeida J.; Henriques P.; Novais P. 利用情感分析预测选举结果[C]// *Trends and Innovations in Information Systems and Technologies*. Cham: Springer, 2020, 1159: 134–143. DOI: 10.1007/978-3-030-45688-7\_14
- [21] Nadeem M. A.; Bukhari S. H. H.; Arshad M. U.; Naeem S.; Beg M. O.; Shahzad W. 巴基斯坦语言在声道中的语言检测与定位[C]// 第17届新兴技术国际会议 (ICET). Islamabad, 2022: 142–147. DOI: 10.1109/ICET56601.2022.10004691
- [22] Naithani K.; Raiwani Y. P. 基于文本的情感分析：自然语言处理与机器学习方法的实现[J]. *Expert Systems*, 2023, 40(5): e13114. DOI: 10.1111/exsy.13114
- [23] Prabhakar Kaila D. R.; Prasad D. A. Twitter 信息流冠状病毒爆发主题建模方法 *International Journal of Advanced Research in Engineering and Technology*, 2020, 11(3): 128–134.
- [24] Raja Chakravarthi B.; Muralidaran V.; Priyadharshini R.; McCrae J. P. 为泰米尔-英语混合文本情感分析构建语料库 [EB/OL]. arXiv:2006, 2020.
- [25] Saleh Alzubaidi M.; Bourennani F. 结合表情符号与文本的推文情感分析[C]// 第4届信息科学与系统国际会议 (ICISS). New York: ACM, 2021: 111–114. DOI: 10.1145/3459955.3460608
- [26] Shah M.; Shaikh H.; Mahar J.; Mahar S. 基于规则剥离方法的信息检索系统信德语词干提取器. *Sindh University Research Journal (Science Series)*, 2016, 48(4).
- [27] Tucker J. A.; Guess A.; Barberá P.; Vaccari C.; Siegel A.; Sanovich S.; Stukal D.; Nyhan B. 社交媒体、政治极化与政治虚假信息：科学文献综述[R]. SSRN Electronic Journal, 2018: 95 pp. DOI: 10.2139/ssrn.3144139
- [28] Vásquez Rojas C.; Roldán Reyes E.; Aguirre y Hernández F.; Cortés Robles G. 将文本挖掘方法融入中小企业战略规划过程[J]. *Industrial Management & Data Systems*, 2018, 118(4): 745–764. DOI: 10.1108/IMDS-01-2017-0029
- [29] Vel S. S. 利用计算语言学与 Python 库的文本挖掘预处理技术[C]// 2021 年人工智能与智能系统国际会议 (ICAIS). Coimbatore (India), 2021: 879–884. DOI: 10.1109/ICAIS50930.2021.9395924
- [30] Xing Y.; Wang X.; Qiu C.; Li Y.; He W. 在线社交网络中基于大数据分析能力的观点极化研究[J]. *Technology in Society*, 2022, 68: 101902. DOI: 10.1016/j.techsoc.2022.101902
- [31] Zhang J.; Wang W.; Lu C.; Wang J.; Sangaiah A. K. 用于交通标志分类的轻量级深度网络[J]. *Annals of Telecommunications*, 2020, 75: 369–379. DOI: 10.1007/s12243-019-00731-9

**Word Count (excluding references):** 7 575 words

#### Peer-Review Record

- **Fast-track status:** Not fast-tracked
- **First-round reviews received:** 3 reports

- **Revision cycles completed:** 3 rounds
- **Final version submitted:** 3 June 2025

editorial staff accept no responsibility for any injury to persons or damage to property resulting from the ideas, methods, instructions, or products discussed herein.

#### **Publisher's Disclaimer**

The views, opinions, and data expressed in this article are solely those of the authors and do not necessarily reflect those of the *Journal of Hunan University (Natural Sciences)* or its editors. The journal and its