Open Access
Full Length Article

ORIGINAL ARTICLE

Analyzing Temporal Variations and Public Sentiment on X: A Clustering Analysis of the 2024 Pakistan Elections

Veena Kumaria[1], Areej Fatemah Meghjib[2], Sania Bhattic[3], Dua Aghad[4]

[1,2,3,4] Department of Software Engineering, Mehran University of Engineering and Technology, Pakistan

ABSTRACT

X serve as vital arena for public discourse on current affairs, politics, and geopolitical events, as it encapsulate individual's experiences and preferences, analysing public opinion within these digital spaces provides critical insights for shaping informed decision-making. This research examines public sentiments surrounding the 2024 Pakistan General Elections through the analysis of tweets collected within the due period. We employ a multi-faceted approach, integrating temporal analysis to pinpoint peak engagement periods, K-means clustering to identify distinct topical clusters, and VADER sentiment analysis to assess the prevailing emotions within each thematic group. The research explores the utility of common election-related hashtags to ensure the retrieval of an equally representative and non-influence opinion. Our findings reveal four prominent topics of discussion revolving around two main political parties, electoral processes, and public enthusiasm. Sentiment analysis indicated predominantly mixed sentiments towards the political parties. The integration of temporal patterns, topical clustering, and sentiment analysis exhibits practical and hands-on recommendations, offering political stakeholders and analysts a nuanced understanding of public sentiment dynamics, which can shape electoral strategies and decision-making processes.

Keywords: *Clustering, Machine learning, Sentiment analysis, Social media, User engagement*

Author`s Contribution
[1,2,3,4] Data analysis, interpretation, and manuscript writing, Active participation in data collection
[1,2,3,4] Conception, synthesis, planning of research. Interpretation and discussion

Address of Correspondence
Veena Kumari, Email: vina.mehrani@gmail.com

Article info.
Received: 15 April 2024
Accepted: July, 2024
Published: 30 Dec 2024

Cite this article: *kumaria V, Fatemah M A, Bhattic S, Aghad D. A Novel Morphological Rule-based Approach for Urdu Text Sentiment Analysis. J. inf. commun. technol. robot. appl.2024; 1-16,https://doi.org/10.51239/jictra.v15i1.342*

Funding Source: *Nil*
Conflict of Interest: *Nil*

INTRODUCTION

Social media platforms like X, formerly known as Twitter, have become an important part of modern life in the digital era, significantly influencing communication, consumption of information, and view of the world [1]. Due to the two-way communication between verified organizations, individuals (including political figures), and the public, X is considered a more formal and reliable platform than Facebook to discuss current affairs, news, politics, and geopolitical events, as well as share opinions on such matters [2]. Recently, engagement on X of people from Pakistan has gradually increased; numbers published in X's advertising resources indicate that it had 4.50 million users in Pakistan in early 2024 [3]. Focusing on elections, the General Pakistan Election of 2018 saw the extensive usage of social media like Facebook and X as a platform for political campaigning and to inspire and sway voter sentiment [4]. For the Elections of 2024, it was anticipated and later deemed accurate that X would serve as a key platform for propagating and disseminating party

slogans, election-related hashtags, and important policy directives.

Though the engagement on X is enormous, only 21% of the total amount of data generated using this platform is organized or structured, with the remaining data being severely unstructured [5]. Unstructured data lacks a predefined format due to the variability and randomness found in raw data and thus requires initial processing before meaningful information and insights can be extracted. Although text, images, sounds, and video formats together make up the unstructured content over X, textual content is found in a larger ratio. If correctly handled, the hidden information within this content has the potential to be a valuable source of insights [6].

The rapidly growing political tweets provide real-time public sentiment and reaction to political policies. The analysis of these tweets can give political parties unprecedented access to understand the concerns of the masses and their opinions regarding policies and issues while making campaign decisions. Analysing voter groups, their opinions, and general preferences regarding certain changes may help make or break an election. Prior research has limited its exploration to sentiment analysis, whereas there is enough data to also analyze the dynamics of public opinion, systematically categorize and analyze the wide assortment of subject matter to uncover underlying patterns and connections within electoral discourse [7-9].

This research proposes filling these gaps by utilizing data clustering, sentiment analysis, and natural language processing techniques to look into data collected through X related to the Pakistan General Elections 2024. The aim is to study the temporal patterns of user engagement and their topical clusters, as well as the group of users present during the election period. In this research, sentiment analysis is integrated with clustering in an attempt to understand the diverse range of opinions out there on the X platform.

We also explore the feasibility of using common election-related hashtags for retrieving unbiased tweets for this experiment. The examination of these hashtags can help serve as possible clues of public sentiment and be leveraged to improve sentiment accuracy; this would also enable a better understanding of the complex nature of user engagement on X in the elections for enhanced political decision-making processes.

Focusing on user-generated content during the electoral period, a series of research questions have been formulated to define the scope and direction of the current study. The research questions proposed for this study are:

RQ1. What are the temporal patterns of user engagement on X during the election period, and how do these patterns vary across different timeframes during the event?

RQ2. What distinct clusters are identified using clustering techniques?

RQ3. What are the sentiments revealed using sentiment analysis surrounding major election topics, and how do these insights help understand voter behavior or trends?

RQ4. How can sentiment and topical cluster analysis on X during the election provide actionable insights for political communication?

In this work, we present a comprehensive analysis of social media discourse surrounding the 2024 Pakistan Election to explore the temporal dynamics, community structures, patterns, and perspectives evident in tweets captured from X. The key contributions of this work are as follows:

- Creation of a dataset comprising public tweets on the 2024 Pakistan Election.
- Use of temporal analysis techniques to determine peak user engagement periods, enhancing our understanding of voter interest trends.
- Employing clustering to identify distinct topics of discussion, bringing clarity to prevalent election-related concerns.
- Sentiment analysis on the clustered tweets to provide a nuanced view of public opinion across different issues.
- Analysis to gain actionable insights that can contribute to political scenarios aligned with public interests.

This research paper is structured as follows: Section 2 discusses existing work in this field through a

brief Literature Review. Section 3 delves into the Research Methodology, focusing on the data collection feature extraction, cluster analysis, and sentiment analysis approaches used in the research. Section 4 presents the Results of the discovered temporal patterns of user engagement and the clustering outcomes with sentiment insights. Section 5 details a Discussion of the process, pre-processing techniques, temporal analysis approach,
results, followed by Section 6 presenting the Conclusion and Future Directions

LITERATURE REVIEW

Ref#	Research Title & Published year	Dataset	Technique	Findings
[10]	Sentiment Analysis of Tweets using Unsupervised Learning Techniques and the K-Means Algorithm (2022)	18k tweets were collected from the coordinates corresponding to the city of lima, Peru.	employed the K-Means algorithm to investigate the sentiment distribution of X users towards managing pension funds.	resulted in the distribution of attitudes, showing 22% of the tweets being good, 4% being neutral, and 74% being negative.
[11]	Studying Public Perception about Vaccination: A Sentiment Analysis of Tweets (2020)	Uses the sample of 9581 vaccine-related tweets	(TF-IDF) and then NLP techniques were applied, followed by machine learning algorithms for sentiment classification.	The number of tweets with negative sentiments was only slightly higher than those with positive or neutral sentiments.
[12]	Using K-Means Clustering Method with Doc2Vec to Understand the Twitter Users' Opinions on COVID-19 Vaccination (2021)	31k English tweets were collected containing the discussion over COVID19 vaccine from Australian X users.	Used k-means clustering approach to observe the public perception on COVID19 vaccine and LDA topic model to identify commonly discussed topics	Four potential topics of discussion were identified.
[13]	The emergence of social media data and sentiment analysis in election prediction (2020)	Dataset is collected from a variety of sources using data scraping methods and REST API as well.	Explored the use of sentiment analysis and social media as a basis to forecast election results, examining several approaches, challenges, and possible future directions.	Highlights the challenges related with predicting election results and open issues associated to sentiment analysis.

[14]	An unsupervised approach for Twitter Sentiment Analysis of USA 2020 Presidential Election (2022)	Used Twitter API to collect tweets focusing on tweets about the candidates.	Clustering (Agglomerative), TF-IDF, hash index, and Vader lexicon for sentiment labeling.	Hash index representation showed best clustering accuracy. Vader lexicon demonstrated high accuracy in sentiment classification indicating positive sentiments towards Joe Biden.
[15]	Sentiment Analysis of Tweets to Gain Insights into the 2016 US Election (2021)	Twitter4J library was used to fetch tweets. Keywords were used to search relevant tweets.	Sentiment analysis algorithm used to analyze tweets and determine the sentiment (positive, negative, or neutral) towards each candidate.	The paper found 4,044,162 tweets mentioning one candidate and 2,810,051 tweets mentioning other candidate, indicating more activity around candidate one.
[16]	Prediction and analysis of Pakistan election 2013 based on sentiment analysis (2014)	Using relevant keywords 61k tweets were collected using X API.	Used various classification algorithms to classify tweets based on their sentiments.	70% classification accuracy in the prediction of positive and negative sentiments, as well as 50% classification accuracy in the prediction of neutral sentiments.
[17]	Sentiment Analysis of Twitter Discourse on the 2023 Nigerian General Elections (2024)	English tweets were extracted with the help of the Tweepy library.	TextBlob for sentiment classification, ML models (RF, SVM, XGBoost)	43% tweets were neutral, 33% positive, 24% negative. XGBoost and RF achieved 93% accuracy, outperforming SVM's 89%. Positive sentiments often showed optimism, while negative ones expressed dissatisfaction.
[18]	Sentiment Analysis of Twitter Data: Understanding Public Opinion before the 15th General Elections in Malaysia (2024)	Collected data included keywords, frequently used words, hashtags.	Sentiment analysis, word cloud analysis, descriptive statistics.	72% of tweets indicated positive sentiment, 24% neutral, and 3% slightly negative. A word cloud showed key topics related to the election.
[19]	Social Media, Sentiments and Political Discourse? An Exploratory Study of the 2021 Canadian Federal Election (2024)	Communalytic Pro Research tool is used to collect English tweets, containing popular hashtags.	First, sentiment analysis using Google Natural Language API (Application Programming Interface) then text clustering.	Negative content was more viral; neutral sentiments correlated better with election outcomes than positive or negative ones; provincial leaders and hot-button issues like public health drove strong sentiments.
[01]	Sentiment Analysis before and after Elections using Twitter Data of U.S.	Tweets were collected, employing streaming	TF-IDF (Term Frequency-Inverse Document Frequency) for feature extraction	Before and after the election, the winning candidate experienced a rise in positive sentiment,

	Election 2020 (2021)	Tweepy API across the United States.	from tweets. Naive Bayes classifier for sentiment classification.	while the losing candidate saw a decline in positive sentiment. The paper achieve an accuracy of 94.58% with a precision of 93.19% and an F1 score of 94.81%.
[05]	Sentimental Analysis of Twitter Data with respect to General Elections in India (2020)	X's API is being used to collect the tweets of the two respective Candidates.	Lexicon-based approach for sentiment analysis and NRC (National Research Council) Dictionary-Based Approach, using the Tidytext package for sentiment analysis based on eight emotions and positive/negative sentiment.	The results and the analysis concluded that Candidate-1 is more liked and is popular among the public, as compared to Candidate-2.
[20]	Analysis Content Type and Emotion of the Presidential Election Users Tweets using Agglomerative Hierarchical Clustering (2023)	Data collected using Python module snscrape using hashtags related to the presidential candidates of Indonesia.	Bottom-Up approach of Agglomerative Hierarchical Clustering algorithm is used & Word Frequency Analysis to interpret the patterns of Content.	The paper concluded on 10 identified clusters, 4 containing opinion content and 6 containing informational content.

The diverse methodologies and findings presented here demonstrate the potential of sentiment analysis as a tool for assessing public opinion in various contexts, whereas there is sufficient room to examine the dynamics of public opinion within the proposed subject matter, systematically grouping and analyzing the wide assortment of electoral tweets to uncover underlying subject matter and connections within electoral discussions.

RESEARCH METHODOLOGY

The current study began its process by retrieving the tweets relating to the Pakistan General Elections 2024 with the use of specific hashtags. Since the corpus is linguistic, it required the performance of pre-processing to remove outliers and irrelevant materials. The level of activity was then analyzed over time so as to identify times when a high level of activity was experienced in tweets. This was followed by a structured feature extraction process where interesting parameters were determined, and K-means and hierarchical clustering were applied to cluster the discrete subjects on similar subtopics. The rule-based and aspect-based sentiment analysis were applied in those clusters. These processes were then used to come up with sentiment scores, which were then compared to set metric milestones to be able to assess the effectiveness of the analysis. The above methodological sequence is shown in Figure 1.

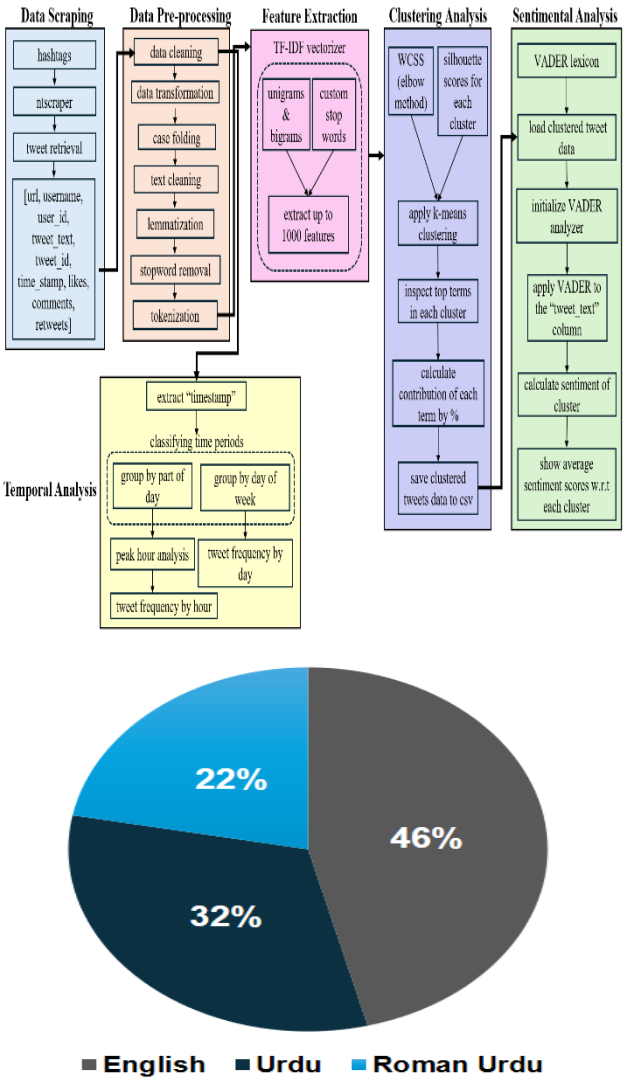
Figure 1

Data Scraping

As explored within the X-related political discourse in the context of Pakistan, the transformation of using hashtags becomes prominent [21]. In the current research, tweets, which mentioned the six most significant hashtags used in the 2024 Pakistan general election, namely, #PakistanElections2024, #Pakistan Election, #Election 2024 pakistan, #general election2024, #Pak elections, and #Elections Results, #Pakistan Elections Results, have been retrieved, filtered, and analysed. In order to gather the data, the open-source Python library

ntscraper was used, which makes it possible to submit the Hypertext Transfer Protocol (HTTP) requests to the public web interface of X and gather information about the tweets encoded in the complex JavaScript Object Notation (JSON), using a custom parser. A total of 10,000 tweets, which included any of the target hashtags, were retrieved between 5th February 2024 and 11th February 2024. The extracted tweets were in the English language, where 46 percent were written in English and the rest in either Urdu or a mix of Urdu and Roman Urdu. The major language of the Pakistani social media discourse is Roman Urdu [22]. The distribution of the collected tweets is shown in Figure 2.

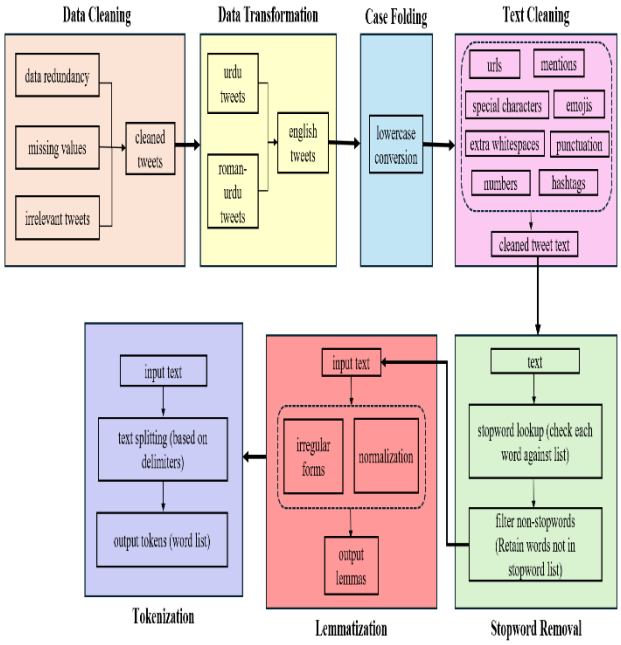
Figure 2.Tweet Distribution in Each Language



Pre-Processing:

Data pre-processing aims to clean and structure raw data before further analysis. This step ensures that the data is suitable and optimized for analysis. Data pre-processing was conducted in seven steps, which comprised data cleaning, data transforming, case folding, text cleaning, stop word removal, lemmatization, and tokenization as illustrated in Figure 3.

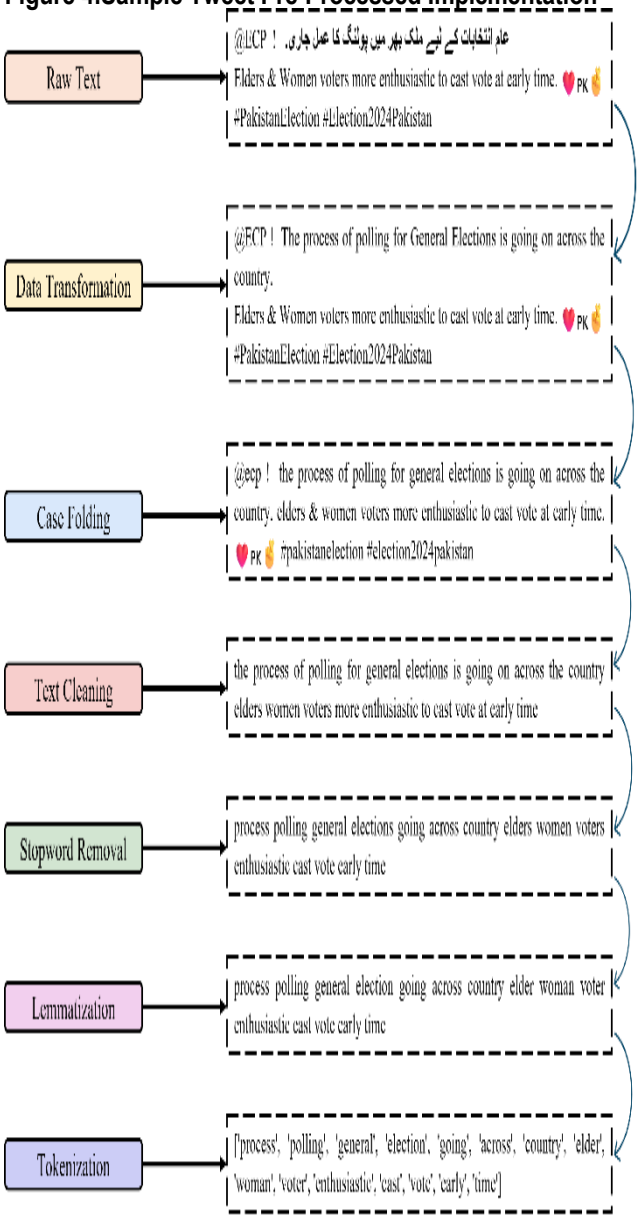
Figure 3.re-processing Steps:



In data cleaning, missing values, irrelevant tweets, and redundant tweets were removed, resulting in roughly 7200 tweets. In data transformation, tweets originally written in Urdu and Roman Urdu were processed through automated language translation algorithms to standardize the language and convert all text into English. Case folding was applied to standardize text data by converting all characters to lowercase to reduce the complexity of textual data [20]. In text cleaning, URLs, mentions, special characters, emojis, extra whitespaces, punctuation, and numbers were systematically removed to reduce noise and enhance text uniformity. We also removed the hashtags used to retrieve the data because their frequency in the dataset could affect the performance of the clustering algorithm. Stopword removal aimed to remove common words that

add no significant information value in the text, followed by lemmatization, which was used to reduce words to their base forms and thus normalize variations. Lastly, the Natural Language Toolkit (NLTK) package was utilized to tokenize content data to simplify text processing by separating words into separate entities known as tokens, to understand and manipulate text in a more structured and effective manner [1, 5]. To better understand the pre-processing steps, Fig. 4 depicts the step-by-step implementation and resultant outcome of each preprocessing step as applied to a sample raw tweet taken from the dataset.

Figure 4.Sample Tweet Pre-Processed Implementation



Temporal Analysis

Temporal analysis is a method of analyzing the time, frequency and the direction of the activities of tweeting during a specific period [23]. In our analysis of the temporal immigration of the population throughout the General Paksitan Election 2024, we applied a time-series analysis method, which determines the activity window the user favors [24]. After the preprocessing phase of data cleaning, the data was loaded and the timestamp column formatted using the date and time of tweets to enable easy manipulation of the dataset. To see the different facets of the temporal activity such as the tweet activity by each day of the event week, we did this by extracting the day of the week of each tweet through the expression data['timestamp].dt.day_name(). The arrangement of days of the week was predefined and it started with Monday and ended on Sunday.

We also pictorized the frequency of the tweets per hour of the day as we retrieved the hour of the day in the form of the expression data['timestamp].dt.hour. The distribution of the number of tweets in an hour was properly analyzed and the number of tweets in an hour calculated and plotted, through the method of peak hour analysis, in order to determine the uppermost hour which showed the highest number of tweets during the campaign.

Feature Extraction

To employ machine learning methods, we had to first transform the textual data into numerical features for cluster analysis. This approach was utilized to lessen the weight of frequently used terms that might not have much value while highlighting necessary keywords that are particularly significant to specific tweets but less common overall [1]. To achieve this, a TF-IDF vectorizer was employed to convert the cleaned tweets into a numerical format suitable for analysis [25]. TF-IDF quantifies the relevance of a particular term within a document. Equation (1) shows how the Term Frequency (TF) calculates the frequency of a word w:

$$tf(w) = \frac{Tweet\ count(w)}{Total\ words\ in\ all\ tweets} \quad (1)$$

Inverse Document Frequency (IDF) down weighs words that are common across many documents. To reduce their effect, such words must be removed. IDF is used to

remove the most frequent words, as mathematically shown in Equation (2).

$$idf(w) = log \frac{Total\ Number\ of\ tweets}{Number\ of\ tweets\ Containing\ word\ w} (2)$$

The most important word in the document is the one with its highest TF-IDF score. Finally, the tweets' words will be shown in vector format, with each word in a tweet having its own TF-IDF value. Equation (3) shows the calculation of TF-IDF.

$$tf\text{-}i(w) = idf(w) * tf(w) \quad (3)$$

The vectorizer was configured to extract up to 1000 features for this experiment, focusing on unigrams and bigrams (single words and two-word combinations) from the pre-processed data [25, 26]. Custom stop words such as "election", "vote", "Pakistan", and "result" were excluded to avoid overlapping of clusters and to ensure prominent separation in clustering analysis. The resulting feature matrix now represents the tweets in a high-dimensional context, which would enable clustering and sentiment analysis.

Clustering Analysis

Clustering is a widely embraced technique used in machine learning and data analytics to gather similar objects or data points [27]. Clustering is a typical example of unsupervised learning that uncovers visual classifications that test the hypothesis [28]. K-means is a widely used centroid-based clustering algorithm that separates data into K distinct clusters based on feature correspondence [29]. Following feature extraction with the TF-IDF vectorizer, the Elbow method was used to identify the optimal number of clusters for clustering analysis by iterating counts from 1 to 10. The algorithm was applied for each cluster count, and the within-cluster sum of squares (WCSS) was calculated. For each cluster count K, the algorithm minimizes the WCSS.

$$WCSS = \sum_{i=1}^{k} \sum_{x \in Ci} || x - \mu_i ||^2 (4)$$

Finally, to identify the optimal cluster number, a graph was plotted using the WCSS value of each cluster, and based on the sharp drop-off curve in the elbow, the preferred number of clusters was decided for analysis. With the optimal number of clusters acknowledged, we proceed to perform K-means clustering on the reduced feature set with the preferred cluster count of 4. The algorithm works by minimizing the following objective function.

$$argmin_C \sum_{i=1}^{k} \sum_{x \in Ci} || x - \mu_i ||^2 \quad (5)$$

Together with the WCSS value, the Silhouette score was also used to evaluate the quality of each cluster. Silhouette score is a measure of cluster quality that numerically measures the inclusiveness of each cluster and the extent to which defining one cluster does not overlap with most other clusters; this calculates the mean of distances between each data point and all other cluster members and compares the result to the mean of distances each point has to all the members of every other cluster [30]. A Silhouette score near to +1 indicates that a particular data point probably belongs within the cluster in which it is found; a score close to zero indicates a possible misclassification of the data point, and a score close to -1 indicates that a particular data point does not belong within its cluster [31, 32]. Equation (6) was used for the silhouette score calculation:

$$Sillhouette\ Score = \frac{1}{n} \sum_{i=1}^{n} \frac{b_i - a_i}{max\ a_i b_i} \quad (6)$$

Clustering the tweet dataset by iteratively improving cluster centroids to minimize variance within each cluster. For visualization purposes, Principal Component Analysis (PCA) was used to transform the high-dimensional TF-IDF data (X) into two principal components (X_pca), allowing for a two-dimensional plot of the tweet clusters to project the cluster's centroid. Mathematically, PCA is defined through Eq. (7).

$$Z=XW \quad (7)$$

After the algorithm iteratively refined the cluster centroids to reduce the variance within each cluster, each tweet was eventually allocated to one of the four clusters. As a result, tweets with similar characteristics were grouped into a cluster, allowing us to discover distinctive topics and patterns within the dataset.

Sentiment Analysis

Sentiment analysis was carried out to categorize text by its sentiment, that is, by the feeling, emotion, or opinion

as conveyed in the piece of writing [33]. This entailed interpretation of the nature of the feeling of the provided text and was either positive, negative or neutral. As this work involves an unlabeled corpus of texts, a lexicon-based sentiment analysis algorithm was employed. It is a useful method of unlabeled analysis since it involves ready-made dictionaries of words that are used to define the sentiment based on how much certain emotions were conveyed in the piece. The method enabled the quantitative analysis of emotions in individual tweets since the frequency of the words of a sentiment lexicon was counted [34].

To conduct this analysis, the Valence Aware Dictionary and sEntiment Reasoner (VADER) sentiment lexicon were used to discern the sentiment of the tweets. VADER suits social media text very well [17]. VADER is based on a dictionary which will match lexical properties with emotional strength to assess the polarity of the textual input, referred to sentiments scores [35]. When the sentiment polarity is less than zero (<0), the text is classified as being a negative one. In the case where the polarity is zero (0), there is a display of a neutral sentiment. In case the sentiment polarity is above zero (>0), then it will be labeled as "Positive". An overall text has got a sentiment score which is a summation of the strength of the sentiment of individual words in the text. The score is obtained by summing the positive, neutral, and negative scores and applying a compound sentiment score calculated with the aid of Equation (8).

$$S_{compound} = \frac{sum\ of\ sentiment\ scores}{\sqrt{sum\ of\ squared\ scores + \alpha}} \quad (8)$$

Where α is a constant that stabilizes the value.

The VADER sentiment analyzer tokenizes the input text into individual components. Each token is then corresponded against a pre-defined lexicon that designates sentiment scores. The sentiment scores from individual tokens are accumulated to calculate the overall sentiment score of the text, calculated using Equation (9).

$S_{compound,i}$= vader_ analyzer. Polarity_ scores (Ti)[′compound′] (9)

The analysis aimed to uncover the average VADER sentiment score for each cluster to glean insights into the general sentiment trends within each group of tweets. By aggregating the sentiment scores per cluster, the mean sentiment score for each cluster was calculated using Equation (10).

$$\bar{S}_{cluster} = \frac{1}{n_c}\sum_{i=1}^{n_c} S_{compound,i} \quad (10)$$

RESULTS AND DISCUSSION

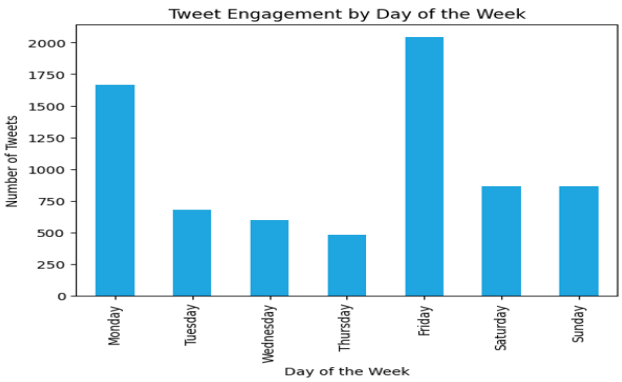
This type of analytical approach can interpret whether certain clusters are predominantly positive, neutral, or negative, and is useful in identifying public opinion trends.

Using the systematic grouping and analysis of a wide assortment of election-related tweets, temporal and clustering techniques allowed the analysis of underlying subtopics in electoral discussions.

Temporal Patterns of User Engagement

The analysis of temporal patterns of user engagement revealed notable trends during the election period. To begin with, engagement by day of the week shown in Figure 5, it was observed that users' activity peaked on Friday, February 9th, 2024, which was the day after the election. This surge followed the lowest engagement levels observed on Election Day - February 8th, 2024, likely caused by partial internet shutdowns across Pakistan.

Figure 5.Tweet engagement by day of the week.



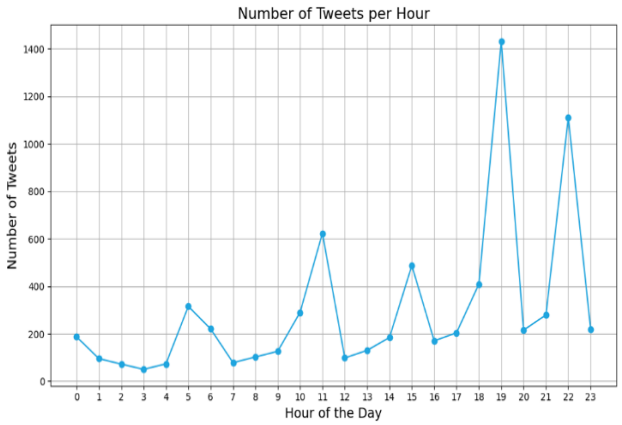
By delving a step deeper into the time series analysis, it became evident that the level of user activity on X was consistently high during the evening hours, while

moderate engagement was observed in the morning, afternoon, and night-time hours, as shown in Figure 6.

Figure 6,Tweet engagement by part of the day.

Observing Figure 7, we can see that tweet activity after a certain decline in the afternoon was consistently high around 7 pm, reflecting a strong engagement window during the evening. There were 1400 tweets posted around the time of 7 pm.

Figure 7.Number of Tweets based on Time of the Day.



Clustering Outcomes with Sentiment Insights

The current study attempts to cluster a collection of data sets into identifiable groups, thus improving the understanding and systematic rationalization of the large-scale text-related corpora. One of the main complications lies in determining the most appropriate number of partitions. To answer this question, two commonly used measures are often used to determine the number of clusters, including the Silhouette Score and the Elbow Method [36]. From the Elbow Method, the WCSS decreases as the number of clusters increases, as shown in Figure 8, indicating that adding more clusters significantly reduces the WCSS after a point. This suggests that the optimal number of clusters may lie up to 4.

Figure 8.Elbow Point

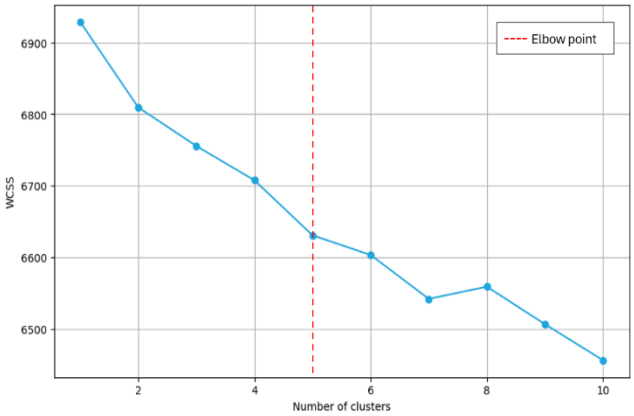


Table 1 shows the silhouette scores for different cluster counts. The silhouette score declined at k=5 with a value of 0.028490, which aligns with the Elbow Method's indication. Considering both methods, we have discussed the clusters up to k=4 in the results.

Table 1.Silhouette Score for Clusters.

Clusters	Silhouette Score
2	0.016780
3	0.017838
4	0.021390
5	0.028580

When k=2, two distinct clusters were identified. These two clusters were separated as visualized in Figure 9, with clear thematic distinctions between them.

Figure 9.Cluster visualization when k=2.

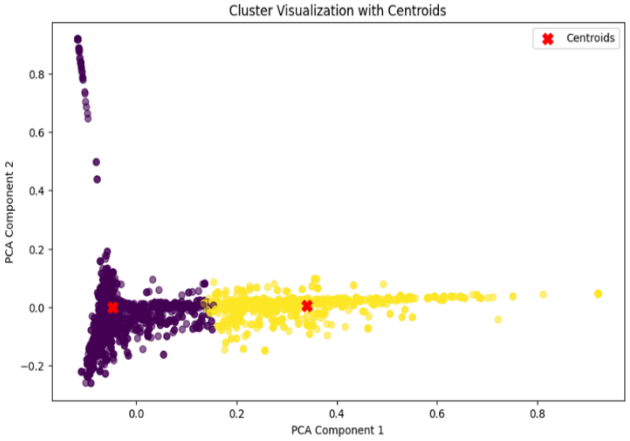


Table 2 depicts the percent of term relevance within each cluster from the least to the most relevant.

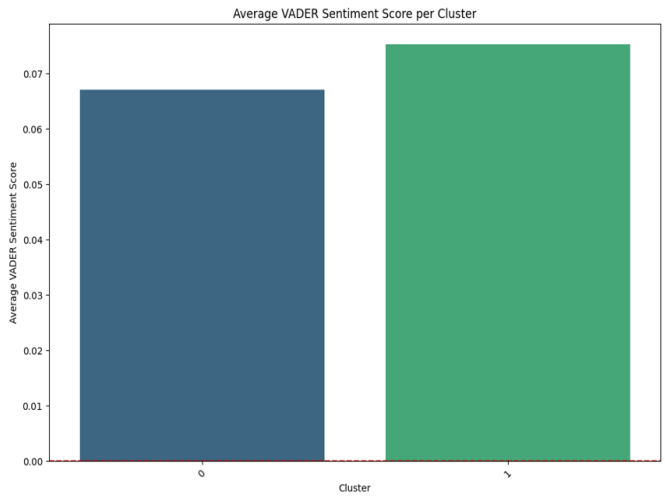
Table 2.Top Terms and their Percentages (k=2).

Cluster	Term	Percentage

0	candidate	13.37%
	go	11.74%
	pmln	16.28%
	people	17.72%
	pti	19.62%
	na	21.27%
1	prime minister	3.63%
	minister	3.72%
	pti	3.60%
	khan	29.12%
	imran khan	29.01%
	imran	30.92%

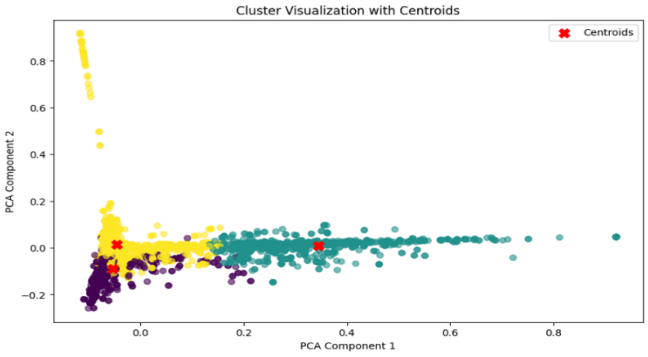
With k=2, the sentiment analysis findings combined with clustering revealed insights into voter sentiment as illustrated in Figure 10. Cluster 0 with an average sentiment score of 0.067011, with neutral to mildly positive sentiment, and Cluster 1 with an average sentiment score of 0.075299, with slightly more positive sentiment. At this point, the clusters had a very distinct partition; Cluster 0 contained terms related to the first political party, and Cluster 1 comprised terms referencing the second political party.

Figure 10.Average VADER Sentiment Score per Cluster when k=2.



Going a step deeper to better analyze the public sentiments, explore the political landscape concerning the user's perspective, and explore different dimensions, k was set to 3. Figure 11 visually illustrates three distinct clusters identified.

Figure 11.Cluster visualization when k=3.



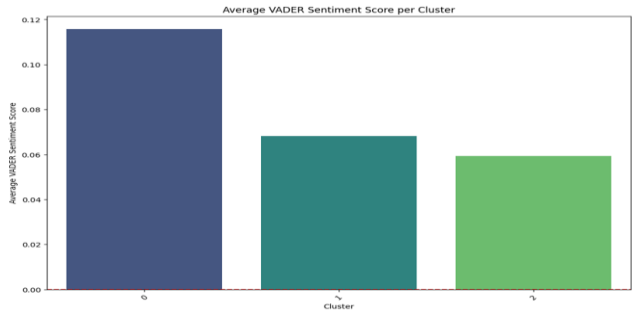
Each term in the cluster highlights distinct aspects of the political discourse within the cluster. Table 3 shows the percentage of term relevance within each cluster from least to the most relevant.

Table 3.Top Terms and their Percentages (k=3).

Cluster	Terms	Percentage
0	independent	9.80%
	candidate	11.45%
	go pmln	13.72%
	go	12.22%
	pmln	21.87%
	na	30.94%
1	prime minister	4.12%
	minister	3.48%
	pti	4.55%
	khan	27.71%
	imran khan	29.65%
	imran	30.48%
2	army	10.13%
	time	12.82%
	party	12.58%
	right	14.22%
	pti	23.86%
	people	26.39%

When the sentiment analysis is conducted without specifying a particular number of clusters, it provides more in-depth information on the election-related discussion by highlighting the specific voter sentiments in each cluster, as the correspondence can be seen in Figure 12. Cluster 0, has a mean sentiment score of 0.115773 portraying the most positive sentiment within the three clusters. Cluster 1, having an average score of 0.068255, exhibits a moderately positive orientation. Overall, the sentiment is positive; however, it is not as enthusiastic as Cluster 0. Lastly, with a score of 0.059275, Cluster 2 comes out as the cluster with the least intense positive sentiment of the three.

Figure 12.Average VADER Sentiment Score per Cluster when k=3.



With k=4, the clustering analysis reveals four distinct clusters, each characterized by its own set of dominant terms. Figure 13 visually illustrates the distinct clusters identified. It can be observed that the cluster boundaries have started to overlap. k=4 demonstrates an increased complexity in the clusters, allowing for a more granular exploration. Each term in the cluster captures distinct aspects of public sentiments with a more detailed examination of the multifaceted political discourse within the data, as shown in Table 4.

Figure 13.Cluster visualization when k=4.

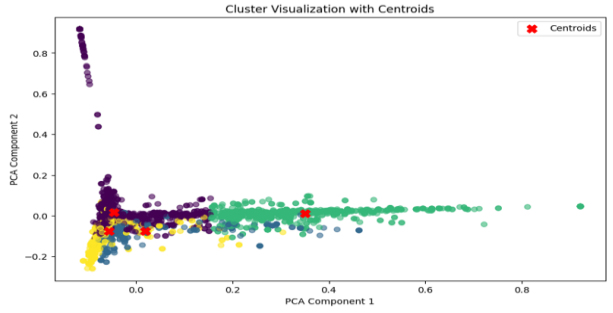


Table 4.Top Terms and their Percentages (k=4).

Cluster	Terms	Percentage
0	time	11.72%
	right	11.96%
	candidate	12.38%
	party	12.80%
	people	25.93%
	na	25.21%
1	khan	4.59%
	prime	5.89%
	pmln	5.37%
	nawaz sharif	25.73%
	sharif	29.01%
	nawaz	29.41%
2	boycott	1.95%
	minister	4.78%
	pti	4.69%
	khan	28.18%
	imran khan	28.09%
	imran	32.31%
3	candidate	8.10%
	seat	11.36%
	go pmln	11.62%
	go	14.02%
	pmln	22.31%
	pti	32.59%

The analysis for k=4 has variations revealing the discussions around major election topics. These clusters highlight different focal points within the political discourse, as well as the associated voter sentiments as shown in Figure 14.

Cluster 0 shows most notably overwhelmingly positive sentiment with an average score of 0.076729; cluster 1 shows the most notable negative sentiment with an average score of -0.071654; cluster 2, with an average score of 0.071303, defines mild positive sentiments, and cluster 3, with an average score of 0.068185, defines a moderately positive sentiment. The analysis of tweet sentiments reveals that out of a total of 7,200 tweets,

27.98% were negative, 33.31% were neutral, and 38.71% were positive, as indicated in Figure 15. These insights illuminate the complex landscape of voter sentiment and behaviour. Overall, these insights illuminate the complex landscape of voter sentiment and behaviour.

Figure 14.Average VADER Sentiment Score per Cluster k=4.

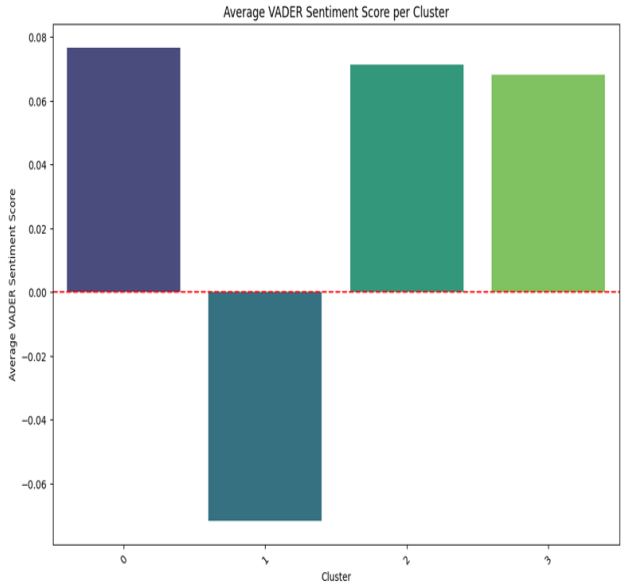
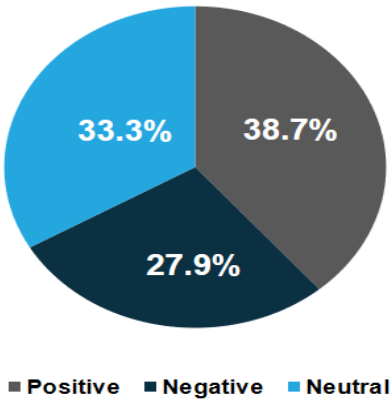


Figure 15.Distribution of Sentiment Categories.



Overall, the findings of the research indicate that:

- The temporal analysis shows that the day after Election Day had the highest engagement of users, moreover in peak evening hours.
- The k-means clustering revealed the three are a few most used topic of discussion among the public indicating that the people of Pakistan are involved in the electoral event.
- The sentimental analysis of each cluster shows the variety of emotions among the people of Pakistan.
- Political landscape is met with notable scepticism, indicating widespread public discontent, and concerns over longstanding influences within political structures.

Overall, these insights illuminate the complex landscape of voter sentiment and behaviour.

By combining clustering techniques with sentiment analysis, this study offers insights into how temporal engagement, topical clusters, and voter sentiment shape electoral discourse. These insights are crucial for understanding voter behavior and guiding strategies for political communication and outreach.

Users' activity peaked on Friday, February 9, 2024, the day after the election, suggesting that users were keen on reviewing the overall election results, sharing their thoughts, and discussing their experiences. It is important to note that the internet remained partially shut across Pakistan on election day itself, which led to the lowest engagement levels on the main day [38]. Consequently, the day following the shutdown saw a surge in engagement on X as people rushed to X to discuss everything they had missed. This post-election peak was marked by users sharing their experiences at polling stations, voicing opinions on the election results, and discussing any irregularities they witnessed. The pent-up eagerness to talk about election day resulted in an intense spike in activity on Friday. On average, a moderate engagement observed in the morning and at noon indicates that users engage sporadically during breaks or transitions between tasks. In contrast, engagement during nighttime hours is notably lower, which aligns with typical user behavior of reduced activity during periods of leisure or sleep. The evening peak not only highlights the time when users are most available but also represents a crucial window where discussions are most dynamic, as evidenced by the elevated levels of participation on the topic during this period. The evening peak, particularly, suggests that targeting users during these hours would maximize visibility and interaction, making it an important consideration for policymakers and political candidates looking to effectively reach the audience during high-engagement periods.

Analyzing sentiment within each cluster, we can uncover insights into how different segments of the population feel about political parties, candidates, and election dynamics. This sentiment analysis using VADER provides a clearer picture of voter sentiment, whether positive, negative, or neutral, across clusters for different values of k. Our findings uncover four prominent topics dominating public discourse: Imran Khan, electoral processes, Nawaz Sharif, and public enthusiasm. These themes shed light on the dynamics of political sentiment in the current landscape. There were positive as well as negative sentiments focused on both the political parties, and public sentiment seemed divided. There was also doubt among the masses and great widespread discontent. Some of the disappointment is related to demands on political arrangements which are considered to be rooted and self gratifying. The disposition of such feelings is mostly one of distrust of the old breed of political leaders and a hope that there is a superior mode of ruling. One of the major themes, the electoral process is an obsession to have a clear and accountable rule of participation in power. It talks about fear of potential manipulation and realizes necessity to change democratic conduct. Lastly, there is evidence of the high level of political interest and involvement of the citizens through public enthusiasm. the anxiety to witness the working of these dynamics, and the influence they may have on the expectations of the people that such a working of them will result in any real change.

The observations expose the role of divergent emotions on voters, as well as the interest in reforms and righteousness. Our data is not limited to user expressions, and 2,788 had a sentiment value of 0.524, which corresponds to general optimism, which we could categorize the tweets under the positive category. Conversely, the 2,014 negative tweets had a sentiment score of -0.482, which is relatively low. In the meantime, the number of neutral tweets with a score between 0.000 and 0.5 amounts to 2398, suggesting that users are very objective or indifferent.

It is evident that as the number of clusters increased, the complexity of the clusters increased as well. However, this led to more overlapping between clusters, as evidenced by the PCA scatter plot visualization. The centroids, which represent the center of each cluster, shifted with each increase in k, resulting in less distinct boundaries between clusters. This movement of centroids occurred because the k-means algorithm adjusts them iteratively to minimize the sum of squared distances between data points and their respective centroids. In higher-dimensional or more complex data, centroids may converge closer to each other, leading to clusters that are less clearly separated.

Analyzing sentiment and topical clusters on X during the election period can yield several actionable insights that are crucial for understanding voter behavior and informing campaign strategies. Some key points on how this analysis can provide value include:

Identifying Key Topics of Interest

By clustering discussions around specific topics, political parties and candidates can identify which issues resonate most with voters. For instance, if a cluster indicates a high volume of discussion around electoral integrity or economic concerns, parties can prioritize these topics in their messaging and policy proposals.

Understanding Voter Sentiment

The sentiment analysis reveals the voters' curiosity in the particular topic of discussion either it's a leader or goverence strategies. The campaigns can used to adopt their communication strategies based on the liking of voter. If the sentiment is negative i.e., for somebody, it may lead to a rethinking of messaging and even an altered course of a campaign in an attempt to rebuild trust.

Segmenting Voter Groups

Different voter groups represent different voter demographics or interest groups, which leads to different topic of discussion. By understanding the sentiment regarding each topic of discussion, campaigns can be designed in such a way that can reach out to that particular audience and also customize their outreach efforts to make them more effective at turning voters.

Predicting Electoral Outcomes

Both sentiment scores and topic of discussion can be associated along with traditional voting behaviour to

accurately predict how various types of people will vote. This can indicate the resources better, reaching out to areas with a positive sentiment, and checking issues in the areas with a negative sentiment.

Optimizing Communication Channels

Political conversations are taking place on X, through the online forums, surveys, tags, and posts can help the stakeholders to identify which platform remains the most influential in the political discourse. Mapping the campaign with the concern medium of communication enables the broadcast of political messages to their most receptive audiences, maximizing the effect of its messages.

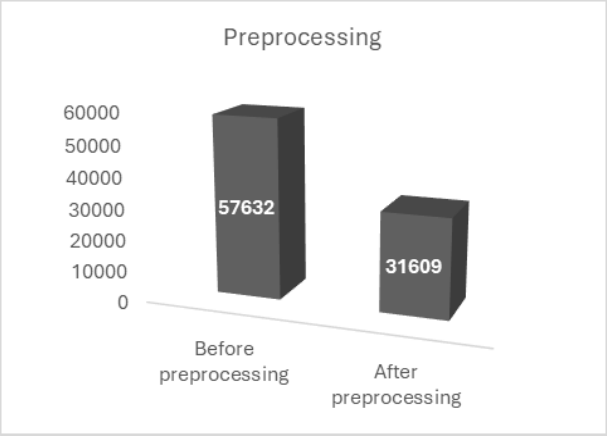
The analysis revealed that the k-means clustering and sentiment analysis provides a significant source of strategic insights which can used by political parties to navigate through the complex electoral dynamics. Moreover, such comprehensive study and its results can be meaningfully applied on a national level as well as regional level to highlight the public interests and goverence priorities that can be implemented while decision-making.

CONCLUSION

In this study, we employed temporal analysis, K-means clustering, and VADER sentiment analysis to explore public sentiment on the 2024 Pakistan General Elections. Temporal analysis was used to determine peak user engagement periods, while K-means clustering identified distinct topics of discussion. Sentiment analysis provided deeper insights into the emotions expressed within each topic. The findings of our analysis indicate that user engagement was notably higher the day after the election, especially during the evening hours. The heightened engagement observed immediately after the event suggests a highly engaged electorate eager to participate in discussions around their future. Clustering results revealed four major topics, and sentiment analysis showed overall positive or mixed sentiment towards the political parties and mixed emotions regarding the electoral process. The insights from clustering showed that increasing the number of clusters allowed a better understanding of topics being discussed by the public; however, increasing the clusters beyond four led to diminishing clarity in topic definition. The critical or mixed views regarding the political parties indicate an evolving political landscape with significant voter preferences. These insights provide opportunities for political stakeholders to effectively target communications, leverage positive sentiment, and enhance voter engagement.

Future research could extend this analysis by incorporating data over an extended period, expanding the data complexities like integrating geographic data to enable spatial analysis of voter sentiment, and providing valuable insights into regional political dynamics. Additionally, incorporating advanced NLP techniques like BERT or transformer-based models could improve the accuracy of sentiment classification. Since the tweets also includes visual content such as images, videos, memes, gifs, etc in future the research may include the incorporation of multimodal analysis to explore and analyze public opinions via other means of communication.

Declaration of Competing Interest



The authors declare that they have no potential competing interests. There are no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Conflict of Interest

The author(s) declare no potential conflicts of interest concerning the authorship, and/or publication of this article. All authors contributed and approved the manuscript.

Author's Contribution.

Veena Kumari: Writing – original draft, methodology, experimentation, visualization, literature review, result reporting. Areej Fatemah Meghji: methodology, writing – editing and review, validation, result reporting. Sania Bhatti: methodology, writing – editing and review, validation, result reporting. Dua Agha: methodology, writing – original draft, validation, result reporting.

REFERENCE

[1] U. Naqvi, A. Majid, and S. Ali Abbas, "UTSA: Urdu Text Sentiment Analysis Using Deep Learning Methods," IEEE Access, 2021, doi: 10.1109/ACCESS.2021.3104308.

[2] T. Nasukawa and J. Yi, "Sentiment analysis: Capturing favorability using natural language processing," in Proceedings of the 2nd International Conference on Knowledge Capture, 2003, pp. 70–77.

[3] N. Mukhtar and M. A. Khan, "Effective lexicon-based approach for Urdu sentiment analysis," Artif Intell Rev, vol. 53, no. 4, pp. 2521–2548, 2020.

[4] N. Mukhtar and M. A. Khan, "Effective lexicon-based approach for Urdu sentiment analysis," Artif Intell Rev, vol. 53, no. 4, pp. 2521–2548, 2020.

[5] A. Z. Syed, M. Aslam, and A. M. Martinez-Enriquez, "Sentiment analysis of Urdu language: handling phrase-level negation," in Mexican International Conference on Artificial Intelligence, 2011, pp. 382–393.

[6] Y. Dubinsky, T. Catarci, S. R. Humayoun, and S. Kimani, "Integrating user evaluation into software development environments," in 2nd DELOS conference on digital libraries, Pisa, Italy, 2007.

[7] Vasudha Dalmia, Hindu Pasts. Sunny Publisher, 2017.

[8] F. Katamba and J. Stonham, Morphology: Palgrave Modern Linguistics, 2nd ed. Macmillan Education UK, 2006.

[9] S. Mukund and R. K. Srihari, "Analyzing Urdu social media for sentiments using transfer learning with controlled translations," in Proceedings of the Second Workshop on Language in Social Media, 2012, pp. 1–8.

[10] M. Humayoun, R. M. A. Nawab, M. Uzair, S. Aslam, and O. Farzand, "Urdu summary corpus," in Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), 2016, pp. 796–800.

[11] A. Z. Syed, M. Aslam, and A. M. Martinez-Enriquez, "Lexicon-based sentiment analysis of Urdu text using SentiUnits," in Mexican International Conference on Artificial Intelligence, 2010, pp. 32–43.

[12] A. Muaz, A. Ali, and S. Hussain, "Analysis and development of Urdu POS tagged corpus," in Proceedings of the 7th Workshop on Asian Language Resources (ALR7), 2009, pp. 24–31.

[13] A. Z. Syed, M. Aslam, and A. M. Martinez-Enriquez, "Associating targets with SentiUnits: a step forward in sentiment analysis of Urdu text," Artif Intell Rev, vol. 41, no. 4, pp. 535–561, 2014.

[14] A. Z. Syed, M. Aslam, and A. M. Martinez-Enriquez, "Associating targets with SentiUnits: A step forward in sentiment analysis of Urdu text," Artif Intell Rev, vol. 41, no. 4, pp. 535–561, 2014, doi: 10.1007/s10462-012-9322-6.

[15] M. Daud, R. Khan, A. Daud, and others, "Roman Urdu opinion mining system (RUOMiS)," arXiv preprint arXiv:1501.01386, 2015.

[16] Z. U. Rehman and I. S. Bajwa, "Lexicon-based sentiment analysis for Urdu language," in 2016 Sixth International Conference on innovative computing technology (INTECH), 2016, pp. 497–501.

[17] N. Mukhtar and M. A. Khan, "Effective lexicon-based approach for Urdu sentiment analysis," Artif Intell Rev, no. 0123456789, 2019, doi: 10.1007/s10462-019-09740-5.

[18] N. Mukhtar and M. A. Khan, "Urdu Sentiment Analysis Using Supervised Machine Learning Approach," Intern J Pattern Recognit Artif Intell, vol. 32, no. 02, p. 1851001, 2018, doi: 10.1142/S0218001418510011.

[19] M. Z. Asghar, A. Sattar, A. Khan, A. Ali, F. Masud Kundi, and S. Ahmad, "Creating sentiment lexicon for sentiment analysis in Urdu: The case of a resource-poor language," Expert Syst, vol. 36, no. 3, pp. 1–19, 2019, doi: 10.1111/exsy.12397.

[20] U. Sehar et al., "A hybrid dependency-based approach for Urdu sentiment analysis," Sci Rep, vol. 13, no. 1, p. 22075, Dec. 2023, doi: 10.1038/s41598-023-48817-8.

[21] N. Mukhtar, M. A. Khan, and N. Chiragh, "Lexicon-based approach outperforms Supervised Machine Learning approach for Urdu Sentiment Analysis in multiple domains," Telematics and Informatics, vol. 35, no. 8, pp. 2173–2183, 2018, doi: 10.1016/j.tele.2018.08.003.

[22] P. Qi, Y. Zhang, Y. Zhang, J. Bolton, and C. D. Manning, "Stanza: A {Python} Natural Language Processing Toolkit for Many Human Languages," in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, 2020. [Online]. Available: https://nlp.stanford.edu/pubs/qi2020stanza.pdf