



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 12, December 2025

Adaptive Machine Learning Models for Real Time Twitter Sentiment Monitoring: A Review

Samadhan Ramesh Yadav¹, Prof. Mr. A.M Dyade²

PG Student, SVERI's College of Engineering, Pandharpur, Maharashtra, India¹

Faculty, SVERI's College of Engineering, Pandharpur, Maharashtra, India²

ABSTRACT: Every day, billions of people share their opinions on social media sites like Facebook and Instagram. Posts are a simple and essential way for people to communicate their emotions. Thus, in this study, we focused on sentiment analysis of social media data. Most of the sentiment analysis techniques available today only consider the textual content of messages and work best with brief and ambiguous exchanges. The polarity of social media interactions and the methods in which emotions are spread on the site are strongly correlated, according to recent research. In order to enhance sentiment analysis performance in social media data, this study focuses on combining sentiment dissemination models with textual data from social media conversations. Therefore, in order to accomplish this, the proposed technique first looks at the inversion of feelings as a phenomenon to evaluate the dispersion of feelings and finds several interesting features of the reversal of feelings. We provide random forest machine learning that considers the linkages between the textual content of social media messages and the patterns of dispersion of feelings in order to predict the polarities of the feelings expressed in social media communications. To the best of our knowledge, this is the first attempt to use sentiment dissemination models to improve sentiment analysis on social media.

KEYWORDS: Text Mining, Machine learning, Sentiment analysis, sentiment diffusion, Social Media

I. INTRODUCTION

Facebook, a globally recognized microblogging site, has impacted and transformed how people get information from people or organizations that interest them. Facebook users can share their opinions, actions, and current happenings with their followers through postings, or status update messages. By republishing or responding to each other's posts, users can also interact with each other. Since its establishment in 2008, Facebook has expanded to become one of the largest online social media networks globally. Due to its numerous applications and the increasing amount of data that Facebook provides, the polarity of the emotions of Facebook message mining users has gained popularity as a research topic. For example, a number of technologies have been developed to assess Facebook users' opinions about political parties and candidates in order to suggest strategies for political elections. Another rapid and effective method used by commercial companies to monitor consumer opinions of their products and brands is Facebook sentiment analysis.

This analysis is done by looking for ideas or viewpoints in several sentences or posts. Facebook's text data stack is consequently quite valuable since it contains significant information. To uncover this information, data mining must be done using particular techniques. This data can be mined using text mining techniques in conjunction with the Natural Language Pre-processing methodology. The significant data that has been mined must also be identified by the sort of emotion. To do this, analytical emotions are employed.

Facebook is one well-known social networking site. Users utilize Facebook to interact with the general public. There are already 330 million Facebook users worldwide, and 18,000 new data points are created every second. The chirp can contain news, viewpoints, debates, and a wide range of other phrase structures. As a result, Facebook becomes rich in text with particular information. In general, people want other people's perspectives to aid in their decision-making. Direct inquiry is one method to obtain this opinion. Asking someone directly takes time and effort. Opinions can also be found on Facebook. However, this perspective must be distinguished in terms of neutral, negative, and positive attitudes. Additionally, these entries have not been categorized using the criteria you are searching for. It is therefore still necessary and widespread.

II. RELATED WORK

This study examines people's perspectives on climate change and the environment over a ten-year period. We use the Pointwise Mutual Information (PMI) technique to evaluate sentiment in environmental tweets on Twitter, Reddit, and YouTube. The accuracy of a human-annotated dataset was 0.65, which was lower than an expert rater's score (0.90) but higher than Vader's. Our findings show that negative environmental tweets are far more common than neutral or positive ones. Recycling, plastic, emissions, air quality, and climate change are the most discussed issues on social media, indicating a global issue. Environmental-related tweets frequently express a range of public emotions, such as joy, fear, and trust [1].

By examining social media interactions and identifying the most prevalent emotions, opinions, and misleading information that are disseminated on multiple platforms, this study investigates people's feelings over the recent monkeypoxoutbreak. By providing insight into public perceptions of monkeypox and recognizing the potential for sentiment-driven panic and stigmatization, this study aims to inform decision-makers and health authorities. The analysis process consisted of two main steps: first, we collected and annotated over 500,000 multilingual tweets using VADER and TextBlob, categorizing attitudes as neutral, negative, or positive. In the second stage, we developed and evaluated 56 classification models using a range of machine learning techniques, including Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbor (KNN). Lemmatization, stemming, TF-IDF, and CountVectorizer were employed for data preparation and vectorization [2].

The rapid advancement of mobile devices and wireless communication technology has led to a surge in the quantity of text data posted on social media, which captures a range of emotions in unstructured formats. Due to the quantity of such sentiment data, accurate sentiment classification has become crucial for applications in both personal and professional contexts. This work examines how profane data affects the accuracy of sentiment classification using deep learning. In order to determine if profanity influences categorization accuracy by functioning as noise, we compared model performance with and without profanity data using online movie review data. The results showed that removing profanity data reduced classification accuracy by roughly 2%, suggesting that cursing may have contextual value in sentiment analysis rather than just being noise [3].

The emergence of Internet 2.0 has brought about a new era of open sharing and communication through social media and web technologies, increasing the connectivity of individuals, corporations, and governments. This study investigates the vast potential of "Big Social Data" for data-driven decision-making through an examination of current advancements in social media, data science, and machine learning. We introduce the "Sunflower Model of Big Data," which integrates the five V's of big data with the ten essential components of analytics. We also provide a taxonomy of social media analytics to assist scholars. Additionally, we outline the basic tools, tactics, and machine learning techniques that are best suited for social data analysis, enhancing researchers' ability to select effective analytics for their specific needs [4].

The exponential proliferation of Arabic social media and e-commerce content has made aspect-based sentiment analysis (ABSA) in Arabic essential for granular sentiment identification. However, Arabic natural language processing has unique challenges due to the absence of annotated datasets. This systematic review, which examines 21 studies published between 2015 and 2021, focuses on the methods, approaches, and datasets used in Arabic ABSA. The findings demonstrate that, despite improvements, there is still a dearth of annotated Arabic datasets and that the ones that are available only cover a limited number of specialized fields. This assessment enables future research to produce new models and resources for Arabic ABSA [5].

Social media's growth has raised the need for sentiment analysis in informal Arabic language, which poses unique challenges due to dialectal differences and morphological complexity. This paper offers the Emoji Sentiment Lexicon (Emo-SL) for Arabic-language tweets, which blends machine learning (ML) with emoji-based characteristics to improve sentiment categorization. A dataset of 58,000 Arabic tweets with emojis was used to calculate sentiment scores for 222 prominent emojis in order to construct Emo-SL. When text-based lexicons were merged with emoji features, ML classifiers performed more accurately than text-only features. Our results demonstrate the potential of Emo-SL, which integrates emoji semantics with traditional text features, to enhance sentiment analysis in Arabic [6].

The COVID-19 pandemic has caused significant societal and economic problems, which have led to an increase in research efforts to understand its impacts through in-depth data analysis. This work provides an automated Latent Dirichlet Allocation (LDA) algorithm to assess and visualize sentiment and topic trends in a large COVID-19 data corpus. This approach identifies significant public concerns using a 10-topic LDA model; Topic 8, which is focused on prevention and transmission, is the most popular. Among the methods that improve insights are WordCloud, Word2Vec, and t-SNE, which reveal semantic clusters surrounding key terms like "Origin," "Symptom," and "Transmission." This paradigm aims to help policymakers and healthcare professionals better manage pandemic-related concerns by providing timely, data-driven insights [7].

Sentiment analysis of text messages is now essential for gaining business insights, comprehending public opinion, and developing personalized apps due to social media's explosive growth. This work focuses on sentiment analysis through Emotional Recognition (ER) by creating a novel lexicon-based system that considers longer words (like "happyyyyyyy") as indicators of elevated emotions. By calculating sentiment ratings for these longer words instead of normalizing them, the method provides a more complex evaluation of user sentiment. When evaluated on casual chat data from Facebook, Twitter, and private discussions, the proposed technique achieves an F-measure rate of 81% to 96% across datasets, outperforming traditional models [8].

Twitter and other microblogging platforms provide valuable insights into public opinion on a variety of global issues, including the state of affairs in conflict zones. This study employs geotagged tweets to conduct sentiment and volume analysis across two datasets from different time periods using Latent Dirichlet Allocation (LDA) for topic modeling and a hybrid feature engineering strategy to enhance sentiment classification performance. According to the results, the nations with the most tweets are Afghanistan, Pakistan, India, the United States, and the United Kingdom. While tweets from Pakistan and Afghanistan are more likely to be positive, the majority of tweets from the US and India are negative. Topic modeling shows that majority Afghan Twitter users are content with the Taliban's rule, whereas negative conversations center on US concerns [9].

Social media platforms, particularly Twitter (now X), have a significant impact on public opinion, particularly during divisive events like elections. In order to examine Twitter/X's impact on the 2021 Peruvian presidential elections, this study examines public sentiment toward candidates and election-related topics amid political instability and the COVID-19 pandemic. By classifying tweets into favorable, negative, or neutral attitudes using advanced sentiment analysis algorithms, the study finds significant trends and issues influencing public opinion. By comparing sentiment dynamics with official voting results, the study assesses how social media sentiment might predict election outcomes, providing political analysts and policymakers with useful information [10].

The overt and covert forms of racism that have grown on social media platforms include memes, fraudulent identities, and racist remarks that endanger social stability. This study attempts to detect racist content in tweets using sentiment analysis and a stacked ensemble deep learning model known as Gated Convolutional Recurrent Neural Networks (GCR-NN). The model uses Gated Recurrent Units (GRU), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) to effectively extract and classify racist utterances. Experiments show that GCR-NN outperforms other machine learning models by detecting 97% of tweets with racist content with an accuracy of 0.98 [11].

Sentiment analysis (SA), an essential usage of natural language processing, has grown in popularity in e-commerce because it can disclose the emotions that underlie customer reviews and comments. This report provides an in-depth overview of the techniques currently used in sentiment analysis across e-commerce platforms and outlines future research opportunities in this field. Out of 271 acknowledged papers, 54 experimental investigations were selected. It was discovered that 48% of the studies employed machine learning methods, 44% employed deep learning strategies, and 7% employed a mixed strategy. Amazon and Twitter were the most widely used data sources. Intriguing directions for further study are also identified in the article, including the development of universal language models, sarcasm detection, aspect-based sentiment analysis, and fine-grained sentiment analysis [12].

Low-resource sentiment analysis has advanced significantly in recent years, particularly with regard to deep learning models. This study evaluation evaluates several approaches and data sources utilized in low-resource sentiment analysis between 2018 and 2023, with a focus on multilingual scenarios. The review states that word embedding learning,

machine translation, and transfer learning are the most often used methods. Social media was the primary data source, followed by product, movie, and hotel reviews. The growing popularity of pre-trained transformers indicates their potential for low-resource languages. Despite these developments, a major challenge remains the absence of annotated datasets. The paper introduces a conceptual framework for sentiment analysis in low-resource languages and provides significant insights for future research in the area [13].

Due to developments in natural language processing (NLP) and the increasing quantity of linguistic data available on social media platforms, sentiment analysis (SA) has emerged as a crucial technique for contemporary applications. This study looks at the main concepts, techniques, and steps needed to implement SA, with a focus on its application in social media data analysis. The efficiency of SA approaches in detecting and classifying sentiment in online content is demonstrated by the study, which looks at how these techniques were utilized to diagnose mental health decline during the COVID-19 epidemic. It concludes with a review of the benefits and drawbacks of the methods examined and emphasizes potential future uses of SA in a range of industries [14].

Sentiment categorization of social media posts is quite challenging, especially in scriptio continua languages like Thai, where words are not separated by spaces and punctuation is not used to indicate the end of sentences. This work introduces SETAR, a unique stacking ensemble learning method for sentiment categorization for the Thai language. By merging features from Word2Vec, TF-IDF, and bag-of-words with a pre-trained Thai language model, WangChanBERTa, the technique generates a hybrid sentence representation. This feature vector is coupled with seven complex machine learning algorithms to generate meta-learners. Extensive benchmarking on four datasets, including a newly developed sentiment corpus annotated by subject-matter experts, shows that the SETAR model significantly outperforms baseline techniques in all classification measures [15].

The global HIV/AIDS epidemic puts millions of people living with HIV/AIDS (PLWHA) at higher risk of victimization and stigmatization. The COVID-19 epidemic and its strict social controls have exacerbated these issues. By analyzing the attitude of tweets and other social media posts from March 2020 to April 2022, this study aims to determine how the pandemic affected PLWHA. Following the use of machine learning techniques, such as textual mining and thematic analysis, to analyze 2,839,091 tweets, 25 subjects were identified and narrowed to 14 key themes. Sentiment study using the VADER Sentiment study Library revealed worries about the challenges PLWHA experience, including high medical costs, stigmatization, limited treatment access, and delayed HIV diagnosis [16].

III. PROPOSED APPROACHES

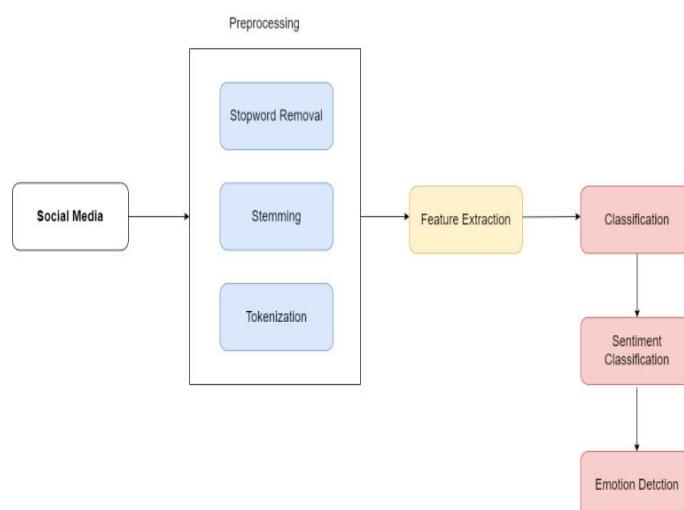


Figure1. System Architecture

To forecast the sentiment polarity of each message, we propose a machine learning method that takes into account the interrelationships between sentiment diffusion patterns and the textual content of Facebook messages. If the sentiment

polarity of a post and its reposts match the sentiment reversal prediction result, the probability that a textual information-based sentiment classifier will accurately classify messages will increase. If not, there will be less chance. This enables emotion reversals to be combined with textual data from Facebook messages.

IV. CONCLUSION

Determining the polarity of opinions expressed in Facebook messages is a significant and challenging problem. The majority of sentiment analysis algorithms now in use only consider the textual content of messages and are unable to produce satisfactory findings because of the unique characteristics of Facebook communications. Despite recent research demonstrating a strong correlation between patterns of feeling diffusion and the polarities of Facebook messages, current approaches mostly rely on textual data from Facebook messages, neglecting the dissemination of information about sentiments. Inspired by fresh research on the synthesis of knowledge from multiple domains, take the first step toward combining textual data and spreading emotions to enhance Facebook's sentiment analysis capabilities.

REFERENCES

- [1] DANIYAR AMANGELDI 1 , AIDA USMANOVA2 , AND PAKIZAR SHAMOI" Understanding Environmental Posts: Sentiment and Emotion Analysis of Social Media Data" IEEE ACCESS 2024.
- [2] STAPHORD BENGESI 1 , TIMOTHY OLADUNNI2 , RUTH OLUSEGUN1 , AND HALIMA AUDU1, "A Machine Learning-Sentiment Analysis on Monkeypox Outbreak: An Extensive Dataset to Show the Polarity of Public Opinion From Twitter Tweets," February 2023.
- [3] Cheong-Ghil Kim1 , Young-Jun Hwang1 and Chayapol Kamyod2,* , "A Study of Profanity Effect in Sentiment Analysis on Natural Language Processing Using ANN," March 2022
- [4] Md. Saifur Rahman and Hassan Reza, "A Systematic Review Towards Big Data Analytics in Social Media," , September 2022
- [5] RUBA OBIEDAT 1 , DUHA AL-DARRAS2 , ESRA ALZAGHOU1 , AND OSAMA HARFOUSHI 1, "Arabic Aspect-Based Sentiment Analysis: A Systematic Literature Review," November 2021.
- [6] MANAR ALFREIHAT1 , OMAR SAAD ALMOUSA1 , YAHYA TASHTOUSH 1 , ANAS ALSOBEH 2,3, (Member, IEEE), KHALID MANSOUR 4 , AND HAZEM MIGDADY 5, "Emo-SL Framework: Emoji Sentiment Lexicon Using Text-Based Features and Machine Learning for Sentiment Analysis," June 2024
- [7] AMREEN BATool 1 AND YUNG-CHEOL BYUN 2 , "Enhanced Sentiment Analysis and Topic Modeling During the Pandemic Using Automated Latent Dirichlet Allocation", June 2024.
- [8] ASHIMA KUKKAR 1 , RAJNI MOHANA2 , AMAN SHARMA2 , ANAND NAYYAR 3 , AND MOHD. ASIF SHAH 4,5, "Improving Sentiment Analysis in Social Media by Handling Lengthened Words", February 2023.
- [9] ERNESTO LEE 1 , FURQAN RUSTAM 2 , IMRAN ASHRAF 3 , PATRICK BERNARD WASHINGTON 4 , MANIDEEP NARRA 5 , AND RAHMAN SHAFIQUE 2 "Inquest of Current Situation in Afghanistan Under Taliban Rule Using Sentiment Analysis and Volume Analysis" January 2022.
- [10] VÍCTOR ANDRÉS AYMA QUIRITA , (Senior Member, IEEE), JUAN DAVID CÁRDENAS , WALTER ALIAGA , (Member, IEEE), VÍCTOR HUGO AYMA QUIRITA , (Senior Member, IEEE), ARAMIS PALACIOS , AND RAFAELA B. SIERRA, "Peruvian Presidential Debates in the Elections of 2021 in Twitter/X: A Sentiment Analysis Approach," October 2024.
- [11] ERNESTO LEE 1 , FURQAN RUSTAM 2 , PATRICK BERNARD WASHINGTON 3 , FATIMA EL BARAKAZ4 , WAJDI ALJEDAANI5 , AND IMRAN ASHRAF 6, "Racism Detection by Analyzing Differential Opinions Through Sentiment Analysis of Tweets Using Stacked Ensemble GCR-NN Model," January 2022.
- [12] HUANG HUANG , ADELEH ASEMI ZAVAREH , (Member, IEEE), AND MUMTAZ BEGUM MUSTAFA , (Member, IEEE), "Sentiment Analysis in E-Commerce Platforms: A Review of Current Techniques and Future Directions," August 2023.
- [13] YUSUF ALIYU 1 , ALIZA SARLAN 2 , (Member, IEEE), KAMALUDDIN USMAN DANYARO 1,3, (Member, IEEE), ABDULLAHI SANI B. A. RAHMAN 1 , AND MUJAHEED ABDULLAHI 1, "Sentiment Analysis in Low-Resource Settings: A Comprehensive Review of Approaches, Languages, and Data Sources," May 2024
- [14] FERNANDO ARIAS 1,2,3, (Member, IEEE), MAYTEÉ ZAMBRANO NÚÑEZ 1,2, (Senior Member, IEEE), ARIEL GUERRA-ADAMES 1 , (Member, IEEE), NATHALIA TEJEDOR-FLORES 2,4, (Member, IEEE), AND MIGUEL VARGAS-LOMBARDO 2,5, (Senior Member, IEEE) "Sentiment Analysis of Public Social Media as a Tool for Health-Related Topics" July 2022.



[15] PREE THIENGBURANATHUM 1,3 AND PHASIT CHAROENKWAN2,3 “SETAR: Stacking Ensemble Learning for Thai Sentiment Analysis Using RoBERTa and Hybrid Feature Representation “ September 2023.

[16] RICHARD K. LOMOTEY 1 , (Member, IEEE), SANDRA KUMI 2 , (Graduate Student Member, IEEE), MAXWELL HILTON1 , RITA ORJI 3 , (Member, IEEE), AND RALPH DETERS 2 , (Member, IEEE), “Using Machine Learning to Establish the Concerns of Persons With HIV/AIDS During the COVID-19 Pandemic From Their Tweets” April 2023.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com



Scan to save the contact details