iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XI **Month of publication:** November 2023

DOI: https://doi.org/10.22214/ijraset.2023.56545

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

A Review Paper on Sentiment Analysis using Machine Learning

Shweta Vishwakarma[1], Neeta Nathani[2]
Research Scholar, Professor, Gyan Ganga Institute of Technology and Sciences, Jabalpur

Abstract: ***Opinion Mining (OM) or Sentiment Analysis (SA) can be described as the process of identifying, extracting, and categorizing viewpoints on various subjects. It falls under the domain of natural language processing (NLP) and is commonly employed to gauge public sentiment towards specific laws, policies, marketing campaigns, and more. This involves the development of methodologies to collect and analyze comments and opinions posted on social media platforms concerning legislation, regulations, and other related matters. Information extraction plays a pivotal role in this process, as it is both challenging and invaluable. Therefore, the utilization of automated opinion-mining techniques becomes imperative for extracting sentiments from various online sources. Presently, several methods are employed, including machine learning (both supervised and unsupervised), lexical-based approaches, and sentiment analysis. This study aims to comprehensively examine the methodologies employed in sentiment analysis (SA) and opinion mining (OM), two interrelated yet distinct techniques. Additionally, it delves into the domains of sentiment analysis applications and the associated challenges, building upon previous research efforts in the field. In this review paper different sentiment analysis techniques have been discussed.***
Keywords: IoT, MATLAB, AI

I. INTRODUCTION

Sentiment analysis, a pivotal domain in the realm of natural language processing (NLP), involves the computational analysis of textual data to discern the underlying emotional tone or subjective sentiment expressed within it. With the exponential growth of online content, from social media posts and product reviews to news articles, sentiment analysis has gained paramount importance across various industries. Its applications span from gauging consumer opinions to brand perception, market research, and even understanding public sentiment towards social and political events.

Over time, sentiment analysis methodologies have evolved significantly. Traditional approaches often relied on lexicons and predefined emotion dictionaries to categorize words and phrases as positive, negative, or neutral. However, the complexity of human language and the nuances of context demanded more sophisticated techniques. This prompted the rise of machine learning models, such as Support Vector Machines (SVMs) and Recurrent Neural Networks (RNNs), which could learn and adapt from data to make more accurate sentiment predictions.

In recent years, deep learning techniques like Convolutional Neural Networks (CNNs) have taken sentiment analysis a step further. These models can capture intricate patterns and dependencies within language, enabling them to grasp the subtlest shifts in sentiment and context. This advancement has paved the way for sentiment analysis to move beyond basic polarity detection and delve into more fine-grained sentiment classification and emotion detection.

As sentiment analysis continues to advance, challenges like sarcasm, irony, cultural context, and domain- specific jargon pose ongoing hurdles. Researchers and practitioners strive to develop models that can not only decipher sentiment accurately but also understand the intricacies of human expression. In this era of information overload, sentiment analysis remains a vital tool, bridging the gap between textual data and actionable insights, thereby enhancing decision-making processes across a wide array of sectors.

A. Theoretical Background

Recent years have seen a rise in the importance of sentiment analysis in text mining and natural language processing (NLP) research. Online applications have become more widely available, and there has been an increase in social platforms for exchanging opinions, online review sites, and personal blogs. These developments have attracted the interest of stakeholders like consumers, organisations, and governments to analyse and explore these opinions. As a result, the primary function of sentiment classification is to analyse an online document, such a blog, remark, review, or new item, as a full sentiment and classify it as positive, negative, or neutral [1], [2].

©IJRASET: All Rights are Reserved

Sentimental analysis has recently gained popularity among academic researchers, and several research investigations are being carried out on the topic. It is often referred to as sentiment categorization and opinion mining. The sentimental analysis classifies texts based on their subjectivity and separates feelings for texts that are mostly customer reviews of goods and services.

Positive and negative feelings are the two categories into which sentiments fall. Neutrality is the phrase for situations in which there may be no feelings at all. A complex process, sentiment analysis involves numerous steps, including sentiment analysis (SA), subjectivity analysis, opinion mining (OM), and sentiment orientation.

A complex process, sentiment analysis involves numerous steps, including sentiment analysis (SA), subjectivity analysis, opinion mining (OM), and sentiment orientation. It is regarded as a unique, developing area of research in computational linguistics, natural language processing, and machine learning (ML). The word level, phrase level, and document level are the three main stages of the sentiment analysis. The task needed for the procedure is determined by the level of analysis. Due to the difficulties of doing the analysis, the word level is the most complicated, but the analysis is easier at the sentence and document levels [3]. Machine learning and semantic-based analysis are the two main methods used to review sentiment analysis. Additionally, a strategy is employed to integrate the two approaches. Machine learning techniques have been applied in several research [4] through [10]. One well-known method of sentiment analysis is semantic-based analysis [11], [12].

II. LITERATURE REVIEW

Zulfadzli et al. in [13], had explored the utilization of social media, which contains an extensive volume of unprocessed data contributed by users in the forms of text, videos, images, and audio. By employing sentiment analysis, this data can be transformed into valuable insights. A comprehensive study was conducted, encompassing research published from 2014 to 2019, drawing from reputable databases such as ACM, Emerald Insight, IEEE Xplore, Science Direct, and Scopus. After a meticulous initial screening and in-depth analysis, 24 out of 77 articles were selected in the review process. These chosen articles underwent assessment based on their research objectives. The findings indicate that a majority of the articles employed the opinion- lexicon method to evaluate sentiment in textual content across social media platforms. Particularly, data from microblogging sites, primarily Twitter, was frequently analyzed. The application of sentiment analysis was observed in various domains including global events, healthcare, and more.

Zhou Gui et. al. [14], mentioned that the growing significance of microblogs and online shopping platforms in people's daily lives due to the widespread use of the Internet and rapid advancements in network technology. These platforms have become vital for communication and learning. The content on these platforms is typically concise and lacks standardized grammatical structures, but it does carry the emotional nuances of users. However, conventional machinery-based approaches struggle with extracting meaningful features from this type of content, as they don't effectively capture the semantic and emotional aspects of short texts.

To address these challenges, the paper introduces a novel approach: a bidirectional model that combines long- term and short-term memory networks with emotional multichannel capabilities. This innovative approach merges deep learning techniques like the attention mechanism and convolutional neural networks, along with both shallow and deep learning methods, to comprehensively understand short texts. By doing so, it enhances the comprehension of semantic and emotional elements within the content, thereby improving the identification and classification of emotional features in short texts.

The proposed model's effectiveness is demonstrated through a comparison with existing models on diverse classification datasets like NLPIR and NLPCC2014. The results showcase notable enhancements in accuracy and F1 scores for the proposed model in the realm of short-text sentiment analysis.

Muhammad Irzam et al. in [15], has mentioned that the practice of sentiment analysis in research, which entails processing and assessing sentiments from textual data. While sentiment analysis has been effectively applied to languages like English and French, it remains limited for languages with fewer resources due to a lack of textual data. This systematic review delves into the domain of sentiment analysis for Urdu, a language known for its scarcity of resources. Despite being spoken by approximately 169 million people globally, Urdu faces several challenges in sentiment analysis research. The existing literature highlights deficiencies such as inadequate large datasets, language parsers, and the absence of pre-trained machine learning models, all of which contribute to subpar performance.

The paper thoroughly examines and evaluates various studies that focus on employing machine learning techniques for Urdu sentiment analysis. After conducting a comprehensive search and refining the selection, a total of forty articles were scrutinized. The research introduces specific objectives that subsequently give rise to pertinent research questions. The search process was meticulously organized within digital repositories, encompassing the identification and assessment of pertinent studies. Extracted data was then derived from these selected studies.

The analysis of the existing literature underscores the potential for enhancing sentiment classification performance. This improvement hinges on effectively addressing challenges such as word sense disambiguation and the availability of extensive datasets. Furthermore, the study suggests avenues for refining elements unique to Urdu, including language parsers, emoticons, techniques for context-level sentiment analysis, pre-processing methodologies, and lexical resources.

Yuling Chen et al. [16], mentioned that the contemporary focus on sentiment analysis of online reviews in the realm of Web information mining. Conventional methods of text sentiment analysis primarily rely on emotion dictionaries or machine learning. However, these approaches are limited by their dependence on constructing emotion dictionaries and manually designing and extracting features, leading to constrained generalizability. In contrast, deep learning models possess greater expressive prowess, enabling them to adeptly comprehend intricate mappings from data to emotional semantics. To this end, the paper introduces an innovative approach that amalgamates Convolutional Neural Networks (CNNs) with Support Vector Machine (SVM) for text sentiment analysis. Experimental results underscore that this proposed methodology significantly enhances sentiment classification accuracy when compared to conventional CNNs. The study validates the potency of sentiment analysis founded on CNNs and SVM, underscoring their potential in advancing sentiment classification accuracy. Comparative study is shown in the table below:

Table 1. Comparison of Literature survey

S. No.	Author Names	Year	Description	Outcome	Strength	Weakness
1	Zulfadzli	2019	This research paper focused finding of the articles from different domain.	The findings indicate that a majority of the articles employed the opinion- lexicon method to evaluate sentiment in textual content across social media platforms	Comprehensive review	Focussed only on a particular domain.
2	Zhou Gui	2022	This innovative approach merges deep learning techniques like the attention mechanism and convolutional neural networks, along with both shallow and deep learning methods, to comprehensively understand short texts.	Improving the identification and classification of emotional features in short texts	Machine learning based	Sophisticated to the user
3	Liaqat MI, Awais Hassan M, Shoaib M, Khurshid SK, Shamseldin MA.	2022	systematic review delves into the domain of sentiment analysis for Urdu, a language known for its scarcity of resources	for enhancing sentiment classification performance	Machine learning based	This improvement hinges on effectively addressing challenges such as word sense disambiguation and the availability of extensive datasets
4	Y. Chen and Z. Zhang	2018	The paper introduces an innovative approach that amalgamates Convolutional Neural Networks (CNNs) with Support Vector Machine (SVM) for text sentiment analysis.	Enhances sentiment classification accuracy when compared to conventional CNNs	CNN and SVM based approach	Requires much processing power

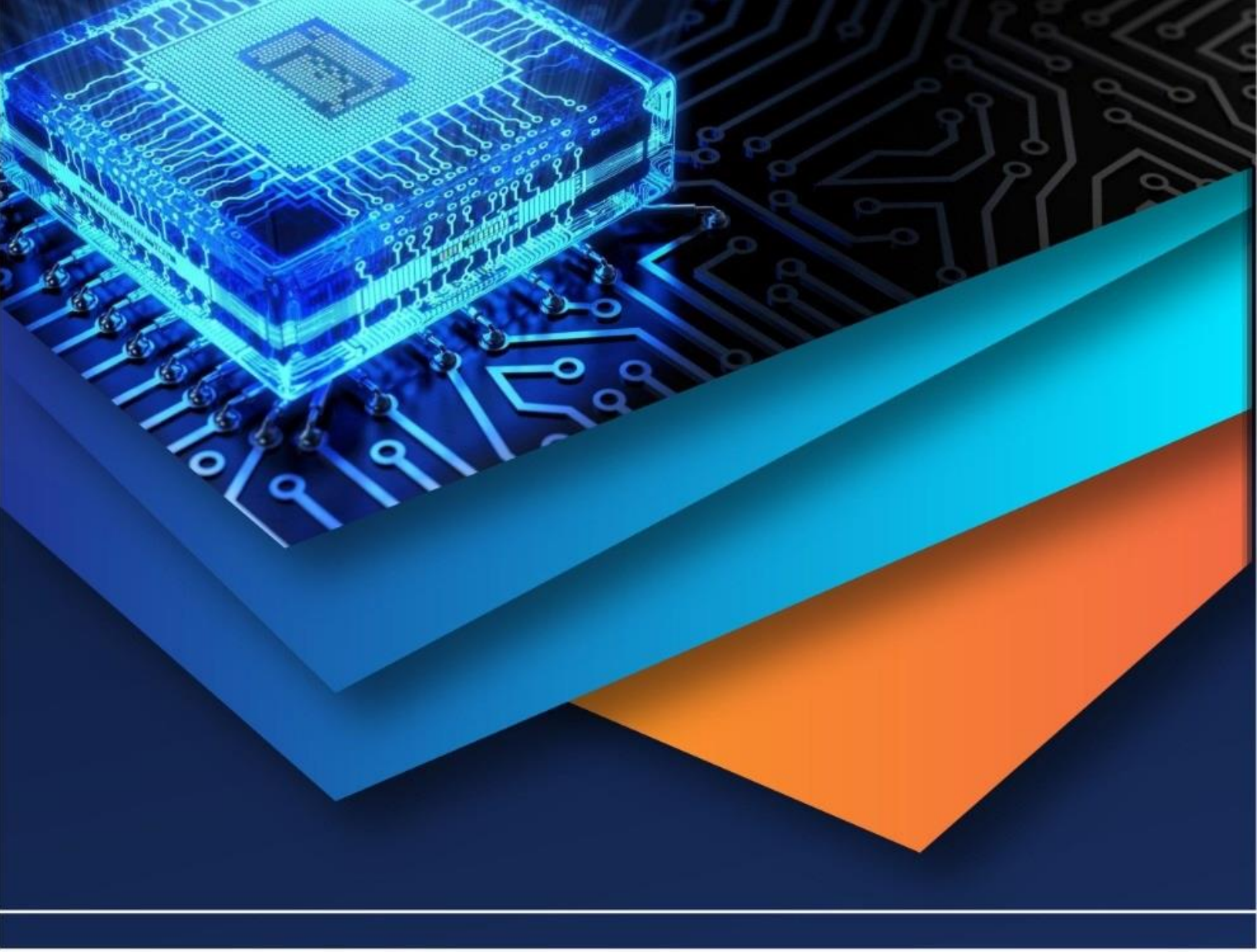
III. CONCLUSION

Sentiment analysis stands as a pivotal field within the realm of natural language processing, holding immense potential for diverse applications. The evolution from traditional methods, reliant on emotion dictionaries and manual feature extraction, to the dynamic landscape of deep learning models, showcases the continuous drive for improved accuracy and generalizability. The fusion of Convolutional Neural Networks (CNNs) with Support Vector Machine (SVM) techniques, as demonstrated in the discussed article, exemplifies the innovative spirit driving sentiment analysis advancements. The pursuit of more effective and nuanced sentiment classification remains a driving force, especially in an era where the digital realm inundates us with a deluge of textual information. As sentiment analysis techniques continue to evolve, they not only empower businesses to gauge consumer perceptions but also hold the potential to decipher complex human emotions, thereby bridging the gap between technology and human understanding.

REFERENCES

[1] M. M. Altawaier & S. Tiun (2016)“Comparison of Machine Learning Approaches on Arabic Twitter Sentiment Analysis”. International Journal on Advanced Science, Engineering and Information Technology, Vol. 6 (2016) No. 6, pages: 1067-1073.
[2] K. Dave, S. Lawrence and D. M. Pennock. (2014), "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews". In Proceedings of the 12th international conference on World Wide Web, pp. 519-528.
[3] E. Boiy and M. F. Moens (2009), "A machine learning approach to sentiment analysis in multilingual Web texts". Information retrieval, 12(5): 526- 558.
[4] M. Ghiassi, J. Skinner and D. Zimbra.(2013) "Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network". Expert Systems with applications, 40(16): 6266-6282.
[5] B. Liu, M. Hu and J. Cheng. (2005),"Opinion observer: analyzing and comparing opinions on the web". In Proceedings of the 14th international conference on World Wide Web, pp. 342-351.
[6] J. Brooke, M. Tofiloski and M. Taboada,(2009) "Cross-Linguistic Sentiment Analysis: From English to Spanish". In RANLP, pp. 50-54.
[7] P. Arora, A. Bakliwal and V. Varma .(2012), "Hindi subjective lexicon generation using WordNet graph traversal". International Journal of Computational Linguistics and Applications, 3(1): 25-39, .
[8] R. Feldman, (2013)."Techniques and applications for sentiment analysis". Communications of the ACM, 56(4): 82-89, .
[9] M. Taboada, J. Brooke, M. Tofiloski, K. Voll and M. Stede (2011) Lexicon based methods for sentiment analysis. Computational linguistics, 37(2): 267-307. .
[10] R. Xia, C. Zong and S. Li.(2011) "Ensemble of feature sets and classification algorithms for sentiment classification". Information Sciences, 181(6): 1138-1152, .
[11] J. Kamps, M. Marx, R. J. Mokken and M. d. Rijke.(2004) "Using WordNet to measure semantic orientations of adjectives"..
[12] Zulfadzli Drus, Haliyana Khalid, 2019, The Fifth Information Systems International Conference Sentiment Analysis in Social Media and Its Application: Systematic Literature Review . Elseveir.
[13] Zhou Gui, 2022 “Research on Sentiment Analysis Model of Short Text Based on Deep Learning”, Scientific Programming Hindawi.
[14] Liaqat MI, Awais Hassan M, Shoaib M, Khurshid SK, Shamseldin MA, 2022. “Sentiment analysis techniques, challenges, and opportunities: Urdu language-based analytical study”. PeerJ Comput Sci. Aug 31;8:e1032. doi: 10.7717/peerj-cs.1032. PMID: 36091980; PMCID: PMC9454799.
[15] Y. Chen and Z. Zhang, (2018) "Research on text sentiment analysis based on CNNs and SVM," 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, China, , pp. 2731-2734, doi: 10.1109/ICIEA.2018.8398173.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)