International Multidisciplinary Research Journal Reviews (IMRJR)

International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 2, Issue 8, August 2025
DOI 10.17148/IMRJR.2025.020804

EMERGING APPROACHES IN SENTIMENT ANALYSIS AND OPINION MINING: A CRITICAL REVIEW

M.Lakshmi Priya[1], Dr.E.Helen Parimala[2], M.Janaki[3], likhitha.G[4]

Assistant Professor, Department of Computer Science, Bangalore City College, Bangalore, India[1]

Assistant Professor, Department of Computer Science, Gitam Deemed to be University, Bangalore, India[2]

Senior Assistant Professor, Department of Computer Application, New Horizon College, Marathalli Bangalore, India[3]

Assistant Professor, Department of Computer Application, Cambridge College Bangalore, India[4]

Abstract: In the current digital era,mining the user opinions has gained substantial attention due to the growing volume of social media contentand the availability of massive data sources. The rise of big data technologies has enabled more advanced sentiment analysis by combining multiple analytical approaches. Thisstudyprovides a comprehensivereview of recent developmentsin sentiment analysis. This paper explores various sentiment analysis techniques,including lexicon-driven models, machine learning algorithms, and rule-oriented method along with methods for managing negations and intensifiers commonly known as modifiers.This study also highlights the main challenges and outlinesfuture research directions for opinion mining. The aim is to offer a comprehensive understanding of recent trends and developments in opinion mining serving as a valuable reference for researchers and practitioners.

Keywords: opinion mining, social media content, lexicon-based, big data.

I. INTRODUCTION

In today's digital era the rapid growth of social media and online product review has resulted in an explosion of user generated content[6,7]. These user generated reviewsarevaluable for businessesand individuals who want to understand public opinions on various topics[6].Sentiment analysisalso known as opinion mining, is a branch of natural language processing that extractssubjective information from text[6,17]. As the volume of text data continues to increase, traditional sentiment analysis techniques face challenges in processing it efficiently[17]. Emerging Big Data tools and techniques provide new opportunities for large-scale sentiment analysis, but they also introduce additional complexities[1]. This paper focuses on summarizing opinion mining methods and the key challenges they present. Several approaches for sentiment analysis are discussed, including machine learning models, lexicon-based methods, and rule-based techniques. Methods for handling negations and intensifiers are also examined. This study aims to offer a comprehensive review of current opinion mining techniques and highlight key challenges and future research directions.

II. PREVIOUS STUDIES

Ali et al.,[24]explored various sentiment analysis tools that integrate computationallearning algorithms. They usedpolitical opinion datasets to evaluate which methods yield higher accuracy. The authors compared the opinionlexicons such as WordNet, SentiWordNet, and Text Blob. The experimental results indicated that Text Blob achieved higher accuracy when combined with machine learning models. However, their study focused solely on textual data excluding emoticons [3,18].

A Multimodal opinion mining system have been presented by Soleymani et al., [2]. Thisstudyfurther investigated the recentadvances in opinion extraction and provided a swift of itsimplementations in affective computing, visual computing, and multimodal communication. The multimodal opinion mining in different domains, such as images, videos, human-human interaction, machine-machine interaction, and spoken reviews were also evaluated. Additionally, the study highlighted the limitations and challenges associated with multimodal opinion mining. However further improvements are needed to achieve more efficacy on large datasets.

Ray et al., [19] proposed an opinion mining framework. The framework uses the Twitter API to collect tweets. Tweets are extracted using the specific keywords within defined time frames.Since the raw data contains noise they applied pre-processing techniques for cleaning. To analysed user sentiment using lexicon based methods.They analysed user

International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 2, Issue 8, August 2025
DOI 10.17148/IMRJR.2025.020804

sentiment using lexicon-based methods. Experimetns and accuracy tests were conducted in R. The framework classifies the tweets as positive, negative or neutral using lexicon-based dictionaries. However, their framework was tested in a business context only; Applying it to multiple domains could improve its accuracy[5]. To rise the accuracy of framework, this study could also be implemented in multiple domains.

Imran Khan et al., [4] have developed an opinion mining framework aimed at analysing the major Big Data sources. Tweets are extracted using Twitter API. The authors have focused the political tweets based on the place and categorised them as political tweet or not. In the proposed framework, there were four main phases namely tweets collection, data processing, classification and storage, and opinion mining. The framework classifies the opinions from the tweets to recognize the sentiments on several political issues. Though, the major limitation of this framework is, it considers only on political tweets.However, the major limitation of this framework is that it only considers political tweets; incorporating domain-specific weighting approaches could extend its applicability to other domains [5].

In [20]Suresh et al.,presented a frameworkfor categorizing user opinions on competing products. Their technique uses a lexicon-based opinion mining method to identify neutral reviews and ensure they are not misclassified as positive or negative reviews. The proposed framework mainly comprises three components. Such as opinion identification, opinion classification and data visualizing. The proposed algorithm concurrently identifies and classifies the opinion values. Unstructured and huge volume of data are handled by using Hadoop tool, which resulted better output. However, the level of accuracy is not up to the level compared to existing works.However, the level of accuracy does not match that of more advanced domain-specific models, which use customized weighting techniques [5].

III. OPINION MINING

Opinion Mining also referred to as Sentiment Analysis[6,17]. As a subfield of NLP, opinion mining is dedicated to the detection and extraction of subjective information from written text[6,17]. The rise of social media and various online platforms has led to an explosion of user-generated content such as reviews, comments, and posts[6,7]. Through the automatic classification of sentiment—whether positive, negative, or neutral—opinion mining supports businesses and organizations in deriving insights from large volumes of user-generated content. The rise of big data technologies has recently boosted this progress, supporting real-time analysis of massive textual datasets[1]. This paper provides a comprehensive study of the state-of-the-art techniques and challenges of opinion mining using big data.

A. *Levels of Opinion Mining*

Opinion mining, also known as sentiment analysis, can be performed at various levels, including document-level, sentence-level, aspect-level and entity levelas depicted in Figure 1[6,7].

- At the document-level, the sentiment of a whole document or text is determined. This can be useful for tasks such as identifying the overall sentiment of a movie review or customer feedback.
- At the sentence-level, individual sentences areanalyzed offering a more granular view of the expressed sentiment. For accurate results, it is important to handle negations and intensifiers effectively [14].
- Aspect-level opinion mining focuses ondetermining the sentiment for specific aspects or features of a product, service, or topic. For example, food quality, service, or ambience in a restaurant review.
- Entity-level opinion mining is similar to aspect-level analysis, but targets the sentiment of specific entities or objects mentioned in the text. For instance, a hotel reviewmight evaluate aspects like staff, room quality, cleanliness and amenities.

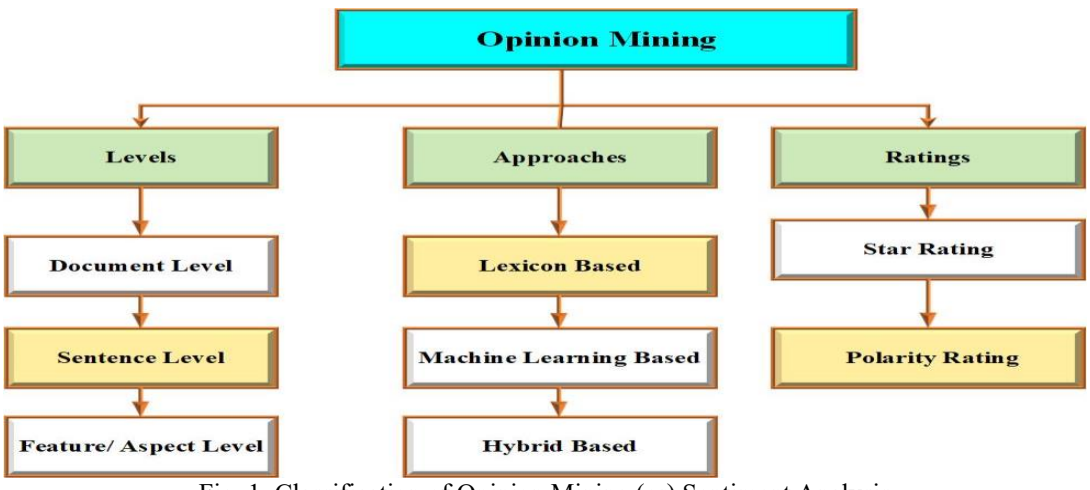

Fig. 1. Classification of Opinion Mining (or) Sentiment Analysis

International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 2, Issue 8, August 2025
DOI 10.17148/IMRJR.2025.020804

B.*Opinion Mining Approaches*

There are two main approaches to sentiment analysis: lexicon based and Machine Learning based methods[6,7]. Some studies have combined these two methods which is known the Hybrid Approach[6,7,11]. Table 1 and Figure 2 depicts the various techniques of Opinion Mining.

1) Lexicon Based:

This method also known as the dictionary-based approach relies on a pre-defined lexicon or dictionary containing phrases with pre-assigned polarity scores [6,7,3,18]. Such lexicons can be generated manually or derived from existing corpora. In an unsupervised technique, classification compares text features with pre defined sentiment scores in the opinion lexicons which can be refined for specific domains using customized weighting methods [3,6,7].

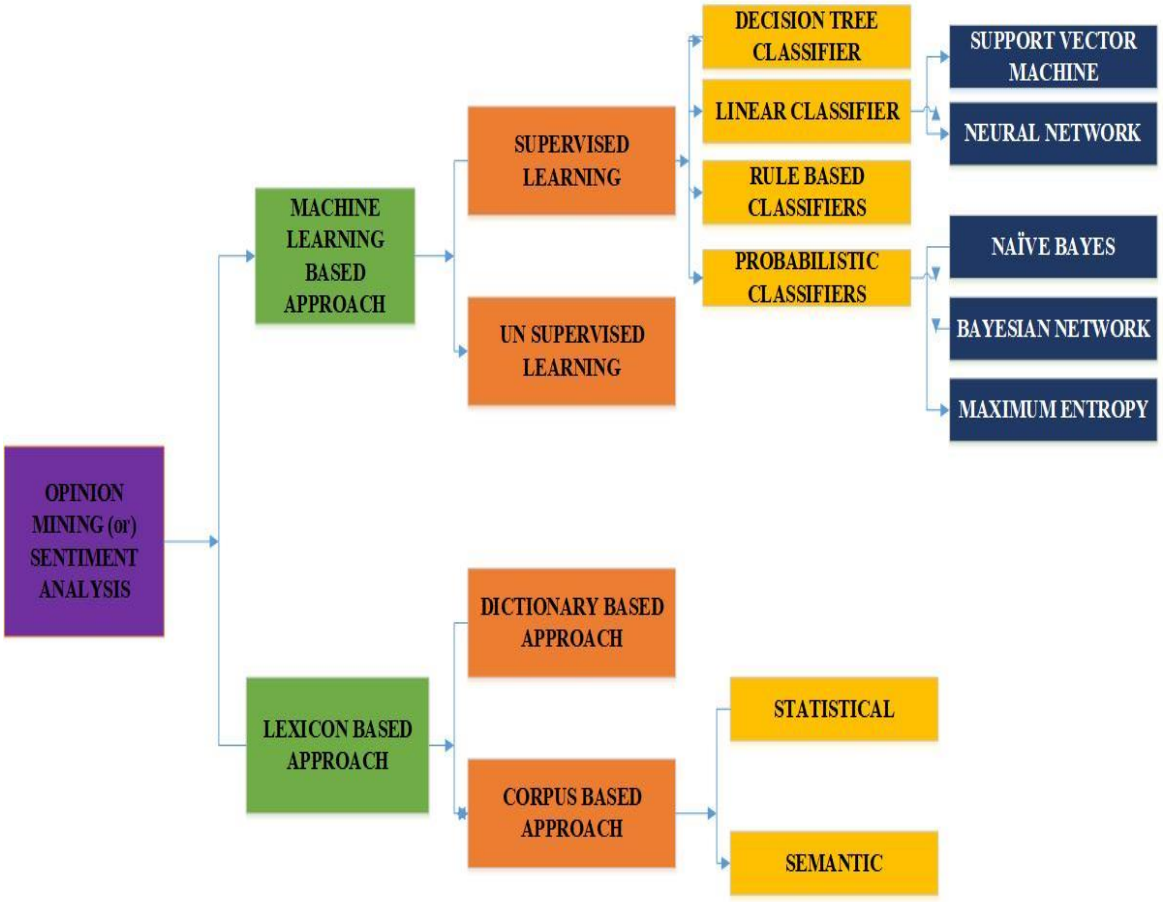

Fig.2. Opinion Mining Approaches

2) Machine Learning (ML):

Machine Learning(ML) tasks often involve statistical and probabilistic methods[6,7]. ML techniques are widely used in sentiment analysis [6,7]. An ML-based method typically uses two datasets: a training set and a test set[27,28].A classifier learns the distinguishing features of documents during training.Several ML techniques have been adopted to classify the reviews. Techniques such as Naive Bayes, Maximum Entropy, and Support Vector Machines (SVM) have proven effective for sentiment analysis [6,7].

3) Hybrid Approach:

Combining lexicon-based and machine learning-based methods results in a hybrid approach [6,7]. This approach has been shown to improve accuracy when classifying sentiment terms in review texts [11]. The advantage of this method is that it combines supervised and unsupervised algorithms, which can enhance classification performance [11].

International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 2, Issue 8, August 2025
DOI 10.17148/IMRJR.2025.020804

Table I.Comparison of Sentiment Classification Approaches: Features, Advantages, and Limitations.

Sentiment Classification		**Features/ Techniques**	**Advantages and Limitations**
Lexicon based	Dictionary-based approach, Corpus-based method, Sentiment lexicons, Ensemble lexicon methods[3,5,6,7]	Manual creation, Corpus based approach, Dictionary based approach[3,5,6,7].	**Advantages** Effective for domain-specific applications [5]. Provides high accuracy within known domains. Covers a wide range of sentiment terms [6], [7] **Limitations** Limited by the fixed size of the lexicon [6]
Machine Learning based	Bayesian networks, Naïve Bayes classification, Maximum Entropy, Neural Networks, Support Vector Machine, Fuzzy classifiers [6,7,11]	Term Presence and Frequency, Part of speech information, Negations, Opinion words and phrases.	**Advantages** Ability to adapt and create trained models for specific purposes and contexts[6,7] **Limitations** Low applicability for new data, since it is necessary of labeled data[11]
Hybrid	Machine Learning approach, Lexicon based approach[6,7,11]	Sentiment lexicon constructed using public resources for initial sentiment detection. Sentiment words are features in machine learning methods.	**Advantages** High accuracy for new data[11] Utilizes both supervised and unsupervised learning for improved results [11] **Limitations** inconsistent Data[6]

C. Opinion Mining Ratings

Opinion mining ratings can be categorized[6,7] as follows:

1)Star Rating:

Stars are assigned according to the sentiment values of the opinionated text. The ratings are assessed using an n-star scale, where *n* can range from 1 to 5. A rating of 5 denotes strong positive sentiment, 2.5 indicates neutrality, and 0 indicates strong negative sentiment[6].

2)*Polarity Based Rating:*

This is the widely used rating method in opinion mining. The polarity is calculated based on the value of emotion words. Afterassigning the polarity value, sentiment class will be classified depending on the polarity value[7].

3)Applications of Opinion Mining:

The growth of the internet has encouraged users to contribute opinions and recommendations about products, services, and events [6,7]. This shift has changed consumer culture, making individuals more open about their views. As a result, content providers now benefit from a wealth of user-generated content, which enables several important applications in opinion mining (OM) [6]. When purchasing a product, consumers can now make informed choices more easily. Opinion mining techniques help individuals evaluate the views and experiences of others and compare competing products [6].Some key OM applications are listed below.

1. Analyzing consumer trends for specific products or services using sentiment analysis [6].
2. Classifying online content into spam and non-spam categories [6,11].
3. Detecting aggressive language, abusive phrases, or hate speech in emails, forums, and tweets [6,11].
4. Helping manufacturers understand critics' reviews and positive customer feedback [6].
5. Assisting policymakers in understanding public perspectives on policies and using this data to create new policies [6].
6. Providing individual opinions that can guide decision-making for similar future issues [6].

IV. CHALLENGES IN OPINION MINING

Much of the user feedback available online tends to be noisy [6]. Opinions are often expressed using informal or non-standard grammar. Opinion sources vary widely in structure and are often presented informally. For example, Twitter

International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 2, Issue 8, August 2025
DOI 10.17148/IMRJR.2025.020804

users frequently rely on abbreviations, misspellings, slang, poor punctuation, and emoticons. These include slang terms, emoticons, and special characters, which add complexity to text extraction. Noise elimination is a critical step in the opinion mining process [6]. Key challenges include the following: Word sense disambiguation is aroused during extraction of exact meanings from words used in a particularcontext.

1.Word sense disambiguation is required to extract precise meanings of words used in different contexts.
2. Identifying spam content, duplicates, outliers, and assessing reviewer credibility.
3. Limitations in classification filters should be addressed to improve sentiment detection accuracy.
4. Limited accessibility and inconsistency of opinion mining tools.
5. Managing the inherent subjectivity and variability of opinion data.
6. The domain-dependent nature of sentiment terms [5].
7. Handling natural language challenges such as ambiguity, co-reference, implicit meaning, and inference.

Since each approach and mining level has its own limitations and challenges, there is a need to develop effective techniques for extracting user sentiments, especially using lexicon-based methods for sentence-level sentiment analysis [14].

V. STATE-OF-THE-ART IN OPINION MINING

With the advancement of sentiment analysis, many techniques are now used to classify and extract user emotions from reviews across massive data sources [6,7]. These data sources include Twitter, news articles, product reviews, and other social networking sites [6]. Machine learning-based classification techniques are among the most widely used approaches; they involve training models to identify and classify opinions using labeled data [6,7]. Deep learning methods, including neural networks, have also been applied to opinion mining with promising results [17].

Other advanced techniques include sentiment analysis for categorizing opinions as positive, negative, or neutral, and aspect-based sentiment analysis, which extends this by identifying specific aspects or features being discussed [6,7].Moreover, domain-dependent approaches have been introduced to enhance the effectiveness of opinion mining models for specific domains or sectors [5].

Overall, the state-of-the-art in opinion mining continues to evolve, with new techniques and models being developed to improve the accuracy and efficiency of extracting opinions from large-scale data sources [6, 17].

VI. FUTURE DIRECTIONS OF OPINION MINING

There are several promising future directions in opinion mining that researchers continue to explore [6].

- Increasing attention is being given to developing advanced algorithms that can detect sarcasm, irony, and figurative language, which remain challenging for current sentiment analysis techniques [6,7].
- There is an ongoing need for better techniques to handle domain-specific language, as terminology can vary greatly across industries and contexts [5].
- There is growing interest in methods for identifying and analyzing opinions of specific subgroups, such as age groups, genders, or cultural communities [6].
- There is a need for more effective techniques to detect sentiment in multimodal content, including images and videos [2].

Finally, improved methods are needed to integrate sentiment analysis with other analytical techniques, such as network and time series analysis, to gain deeper insights into how opinions and sentiments evolve over time [6, 17].

VII. CONCLUSION

This paper presents a comprehensive review of opinion mining in the context of big data, covering the key techniques and challenges involved. It begins with an overview of opinion mining and its growing importance in today's data-driven world, then discusses its levels of analysis, including document-level, sentence-level, aspect-level, and entity-level approaches. It also reviews state-of-the-art techniques such as machine learning-based methods, lexicon-based methods, and hybrid approaches, while highlighting the main challenges of working with noisy, domain-dependent, and large-scale data. Finally, this study outlines promising future directions for advancing opinion mining research. Overall, it provides a thorough understanding of the current landscape and serves as a useful foundation for further research and development in this field.

International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 2, Issue 8, August 2025
DOI 10.17148/IMRJR.2025.020804

REFERENCES

[1] A.AngelPreethi, S.B.R. Kumar, "Visualizing Big Data Mining: Issues, Challenges and Opportunities" International Journal of Control Theory and Applications, Volume 9, Issue 27, Pages: 455-460, 2016.
[2] Mohammad Soleymani, David Garcia, Brendan Jou, Björn Schuller, "A survey of multimodal sentiment analysis", Elsevier, 2017, pp. 1-12.
[3] A.AngelPreethi, and Dr.S.Britto Ramesh Kumar, "A Dictionary based Approach for improving the accuracy of opinion mining on big data", International Journal of Research and Analytical Reviews (IJRAR), Vol. 5, Issue 4, Oct-Dec 2018, pp- i836-844.
[4] Imran Khan, S. K. Naqvi, MansafAlam, and S. N. A. Rizvi, "An efficient framework for real-time tweet classification", springer, March 2017, pp. 1-7.
[5] A.AngelPreethi, and Dr.S.Britto Ramesh Kumar, " Dom_Classi: An Enhanced Weighting Mechanism for Domain Specific Words using Frequency based Probability", International Journal of Applied Engineering Research, Vol.14, Issue 1, 2019, pp 140-148
[6] Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5(1), 1-167.
[7] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval, 2(1-2), 1-135.
[8] Cambria, E., & Hussain, A. (2012). Sentic computing: A common-sense-based framework for concept-level sentiment analysis. Springer.
[9] Dave, K., Lawrence, S., &Pennock, D. M. (2003, August). 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).
[10] Wang, H., Lu, Y., &Zhai, C. (2010, July). Latent aspect rating analysis without aspect keyword supervision. In Proceedings of the 23rd International Conference on Computational Linguistics (pp. 988-996).
[11] Angelpreethi, " Fuzzy Based Sentiment Classification Using Fuzzy Linguistic Hedges for Decision Making", Mapana Journal of Sciences,Vol. 22, Special Issue 2, pp 63-79, 2023 DoI:|https://doi.org/10.12723/mjs.sp2.4
[12] Jindal, N., & Liu, B. (2006). Opinion spam and analysis. In Proceedings of the 2008 International Conference on Web Search and Data Mining (pp. 219-230).
[13] Ding X., Liu B., and Yu P. S, "A Holistic Lexicon-Based Approach to Opinion Mining", In Proceedings of the International Conference on Web Search and Data Mining, ACM., 2008, pp. 231-240.
[14] A.AngelPreethi, and Dr.S.Britto Ramesh Kumar, "NIC_LBA: Negations and Intensifier Classification of microblog data using Lexicon Based Approach "Journal of Emerging Technologies and Innovative Research (JETIR), Volume 6, Issue 6 , PP 753-759, June 2019.
[15] P. S, K. R., F. Iram, and S. R. M V, N. R, "Improved Moth Flame Optimization with Deep Convolutional Neural Network for Colorectal Cancer Classification using Biomedical Images," 2024 Second International Conference on Advances in Information Technology (ICAIT), Chikkamagaluru, Karnataka, India, 2024, pp. 1-6, doi: 10.1109/ICAIT61638.2024.10690631.
[16] Esuli, Andrea and FabrizioSebastiani, "SentiWordNet: A PubliclyAvailable Lexical Resource for Opinion Mining", In Proceedings of 5thInternational Conference on Language Resources and Evaluation,2006, pp. 417-422.
[17] R. Feldma, "Techniques and Applications for Sentiment Analysis, ReviewArticles", Communications of the ACM, Vol. 56, Issue 4, 2013, pp. 82-89.
[18] A.AngelPreethi, and Dr.S.Britto Ramesh Kumar, "A methodological framework for opinion mining", International Journal of Computer sciences and Engineering, Vol. 6, special Issue 2, 2018, pp- 6-9.
[19] A. Angelpreethi and S. B. R. Kumar, "An Enhanced Architecture for Feature Based Opinion Mining from Product Reviews," 2017 World Congress on Computing and Communication Technologies (WCCCT), Tiruchirappalli, 2017, pp. 89-92.
[20] Dibakar Ray, "Lexicon based Sentiment Analysis of Twitter Data", International Journal for Research in Applied Science and Engineering Technology, Volume 5 Issue 10, October 2017, pp.910-915.
[21] Nivetha, Suresh Ramanujam, Nancymala, and kokila, " sentiment analysis using big data",International conference on computation of power, energy, information, and communication, 2015, pp. 480-484
[22]Zhang, D., & Lee, W. S. (2019). Opinion mining and sentiment analysis. In Data Analytics for Intelligent Transportation Systems (pp. 129-144). Springer, Cham.
[23]Duwairi, R.M. and Alshboul, M.A, "Negation-aware framework forsentiment analysis in Arabic reviews", 3rd International Conference onFuture Internet of Things and Cloud, 2015, pp. 731-735.
[24]Ali Hasan, Sana Moin, Ahmad Karim, and ShahaboddinShamshirband, " Machine Learning-Based Sentimental Analysis for Twitter Accounts", Mathematical and computational Applications, Volume 23, Issue 11, 2018, pp. 1-15.
[25]Maheswaran K, Aloysius A, "An Interface based Cognitive Weighted Class Complexity Measure for Object Oriented Design", International Journal of Pure and Applied Mathematics, Vol. 118, No. 18, pp. 2771-2778, 2018.

International Multidisciplinary Research Journal Reviews (IMRJR)
A Peer-reviewed journal
Volume 2, Issue 8, August 2025
DOI 10.17148/IMRJR.2025.020804

[26] Syed TanzeelRabani, Maheswaran K, "Software Cognitive Complexity Metrics for OO Design: A Survey", International Journal of Scientific Research in Science, Engineering and Technology, Vol. 3 , No. 3, pp. 691-698, June 2017.

[27] M. Janaki, Sarojini Balakrishnan, S.N. Geethalakshmi, "Human Activity Recognition Using Chaotic Logistic Map Guided Grey Wolf Optimization with Decision Tree," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 12, no. 1, pp. 142-150, 2025. *Crossref,* https://doi.org/10.14445/23488379/IJEEE-V12I1P113

[28] M. Janaki, Sarojini Balakrishnan , "Leveraging Ensemble Learning Models for Human Activity Recognition", Indonesian Journal of Electrical Engineering and Informatics (IJEEI) Vol. 13, No. 1, March 2025, pp. 249~260 ISSN: 2089-3272, DOI: 10.52549/ijeei.v13i1.6140