

Review of Contextual Mining Techniques with Supervised Learning Algorithms to Enhance the Accuracy and Sophistication of Sentiment Analysis in Textual Data

Abhishek A. Gulhane^{1,a}, Nilesh S. Wadhe^{2,b}, Smeet D. Thakur^{3,c}, Himanshu D. Kale^{4,d}, Shrikant N. Sarda^{5,e}, Pravin R. Nerkar^{6,f}

^{1,2,3,4,5,6} Assistant Professor, PRMIT & R Badnera ,
Maharashtra, India

^{a)} aagabhi@gmail.com

^{b)} nilesh.wadhe05@gmail.com

^{c)} smheet87@gmail.com

^{d)} himkale@gmail.com

^{e)} shrikantnsarda@gmail.com

^{f)} pravinnerkar2007@gmail.com

Abstract. With the rapid growth of web technology and internet usage, an enormous amount of data has become available online. This data encompasses various activities, such as online education and the exchange of ideas and opinions. Social media platforms like Twitter, Instagram, Facebook, and Google+ are widely used for tasks such as sentiment analysis, opinion mining, text summarization, and question-answering. In our study, we explored methods for sentiment analysis using supervised learning and contextual mining. We employed machine learning and deep learning techniques to distinguish between positive and negative tweets. Specifically, classifiers like Naive Bayes (NB), Decision Tree (DT), and Random Forest (RF) were used to classify the sentiment of non-sarcastic statements in the data. Among these methods, Deep Learning Neural Network (DLNN) technology achieved the highest accuracy of 99%. Our study aims to inspire further research and advancements in this promising field. We also examined the reliability of various machine learning methods to ensure the best outcomes, with a focus on supervised learning and contextual mining for sentiment analysis.

INTRODUCTION

Sentiment analysis is a field that combines text classification, natural language processing (NLP), and text analytics [1]. As people increasingly use the internet to express their emotions through text messages, it has become a popular platform for sharing feelings and opinions. Given the sheer volume of these messages, it's difficult for humans to sift through them all to determine which are positive and which are negative. This is where automated sentiment classification techniques become essential [2].

Developing algorithms for sentiment classification is an ongoing effort. While there has been considerable research using machine learning to classify English text sentiment, fewer studies have focused on Chinese text sentiment. Recently, the use of neural networks for classification has grown rapidly across various fields [3]. Deep learning, in particular, has revolutionized machine learning by enabling solutions to previously unsolvable problems. However, deep learning models can sometimes overfit, reducing their ability to generalize [4]. This is especially problematic in critical applications like autonomous driving, medical diagnosis, or financial transactions, where mistakes can have serious consequences [5].

To address these issues, several strategies have been proposed. One promising approach is the Bayesian method, which offers a comprehensive framework for building learning algorithms and assessing neural networks. Unlike the frequentist framework, the Bayesian paradigm provides a more nuanced approach to statistical hypothesis testing [6].

Sentiment analysis is the practice of interpreting and managing subjective content, such as opinions and emotions. By analyzing tweets and other text sources, sentiment analysis can reveal public opinions and evaluations. This method can be used to predict the outcomes of significant events, like movie box office performance and election results. Public reviews on websites like Yelp and Amazon help assess businesses, individuals, places, or products. These

reviews can express positive, negative, or neutral opinions. Sentiment analysis processes user reviews to automatically categorize their opinions based on their expressiveness. As the need to analyze and organize unstructured data from social media grows, sentiment analysis is becoming increasingly popular [8].

General Workflow of Sentiment Analysis: This section outlines a generic machine learning-based process for sentiment analysis, which is typically conducted at three main levels: phrase level, document level, and aspect level.

- **Document-Level Sentiment Classification:** This approach aims to classify an entire text or topic as positive or negative based on an overall sentiment scale.
- **Sentence-Level Sentiment Classification:** This method evaluates the polarity (positive, negative, or neutral) of each individual sentence within a document.
- **Aspect-Level Sentiment Classification:** This technique examines the sentiment of specific aspects or assertions within the text, identifying the polarity of each specific feature mentioned.

Sentiment analysis plays a vital role in opinion mining by providing insights into public sentiment and opinions [9]. Sentiment analysis is often used when consumers need to choose between two product options, taking into account their preferences and the opinions of others. It is a valuable technique for understanding public perceptions of a product. The results of sentiment analysis can help customers make informed purchasing decisions by providing insights into the overall sentiment towards a product or service. A single global rating derived from sentiment analysis can significantly influence one's opinion of a product or service. Additionally, companies can use sentiment analysis to evaluate customer opinions about their products. A typical sentiment analysis procedure is illustrated in Figure 1

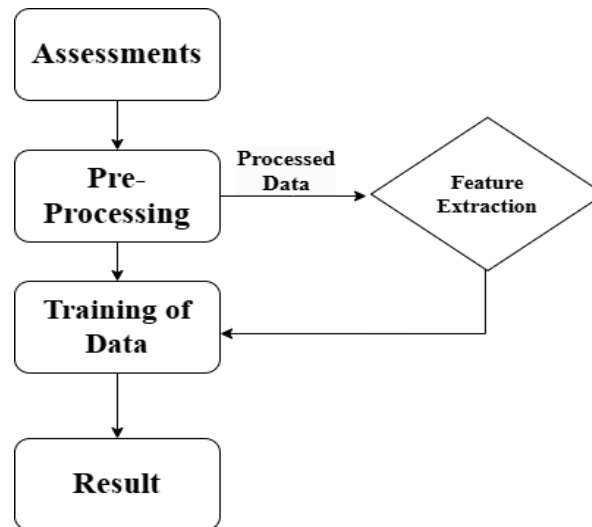


FIG 1: A TYPICAL SENTIMENT ANALYSIS PROCEDURE

Sentiment analysis helps identify which features matter most to customers based on their feedback. Understanding people's emotions opens up numerous possibilities in Human-Computer Interaction. To analyze consumer reviews, the raw data is first pre-processed. After preprocessing, the data is used to train models to achieve the desired results. Sentiment analysis is essential for assessing business activities such as brand management, product development, and other areas [11]. Conceptual, content, and context-based text analysis is gaining popularity in sentiment analysis of natural language text. Context refers to any information that explains the circumstances surrounding a particular item or person. In text, context describes how a phrase is used in relation to the rest of the sentence. The words surrounding a term are essential for understanding the meaning of the statement [12]. Context is crucial in natural language processing (NLP) for tasks like correcting typos and distinguishing between words with multiple meanings. Context-sensitive spelling error tasks address the problem of correcting spelling mistakes within a single sentence. It is not enough to look at a word in isolation to determine its correct spelling; a detailed analysis of the surrounding words is required [13]. Sense disambiguation involves determining the correct meaning of words with multiple definitions. Similarly, the context of a phrase helps clarify the meaning of the terms it contains. For sentiment analysis to accurately determine sentiment polarity, understanding context is essential. Words can have different meanings depending on their context. For example, "high" might be positive when referring to high compensation but negative when referring to high prices [14]. There are various context-related factors to consider, including the degree, nature, and representation of context, as Level, Representation, Type etc.

Supervised Learning for Sentiment Analysis: Supervised machine learning often tackles classification challenges, such as sorting text with numbers or special characters, deciding if a response is positive or negative, and identifying

a person's gender. The core objective of machine learning is to group similar items together. To achieve this, classifiers compute a value from various features and use this value to make categorization decisions [15]. In a 2-class classification problem, a linear classifier functions by creating a hyperplane in a large input space. Points on one side of this hyperplane are labeled as "yes," while points on the other side are labeled as "no." When speed is crucial, especially if the input vector is sparse, the linear classifier can be the fastest option [16].

The process of supervised learning is illustrated in Figure 2.

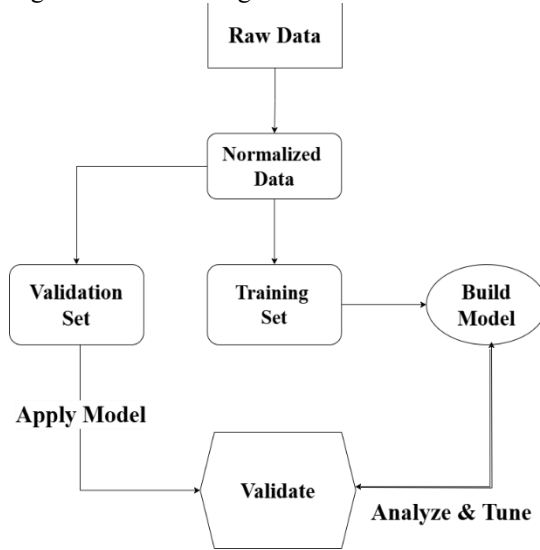


FIG 2: WORKFLOW OF SUPERVISED LEARNING

To evaluate an algorithm's performance, it's important to use a test data set similar to the training data set. Adjusting training control parameters can help improve accuracy depending on the algorithm [17]. When using supervised learning, consider the following:

- Bias in Training Data: Avoid using a training data set that is biased toward a specific output label.
- Overfitting: Be aware of overfitting, where the algorithm becomes too tailored to the training data, which can reduce accuracy on new data.
- Nature of Input Vectors: Consider the type of input vectors being used, whether they are numeric, categorical, or another format.

Some of the most commonly used supervised learning algorithms include Support Vector Machines (SVMs), Neural Networks, Naïve Bayes (NB), Decision Trees (DT), K- Nearest Neighbors, Linear Regression, and Logistic Regression, with SVMs being particularly popular [18]. These methods are commonly used for classifying datasets. Typically, the model is first trained on sample data and then tested using this trained model. The model's performance is assessed based on classification precision. Implementing such solutions often involves extracting phrases and analyzing aspect levels [19].

LITERATURE REVIEW

This section provides a comprehensive review of the existing literature on sentiment analysis.

Taneja et al. (2022) [20] explored the growing popularity of social networking in recent years, focusing on trends in viewing habits for films and television shows. Their research was conducted in two phases. First, they used multilabel categorization to classify movies into different categories. They also performed sentiment analysis on tweets using Text Blob, which proved to be an effective tool. The experimental study utilized a movie dataset and showed that the combined accuracy of the binary relevance and Gaussian Naive Bayes (GNB) classifiers was 85.33%. Additionally, the study found that the comedy genre had the highest percentage of positive tweets, while the horror genre had the highest percentage of negative tweets.

Van et al. (2021) [21] emphasize the importance of sentiment analysis across various studies, from product reviews to social media comments. Their research evaluated machine learning performance using both traditional and deep learning methods on a test set of economic headlines from the Netherlands. The study revealed three key findings:

Crowd Coding Systems (CCS): These systems, involving either crowd-sourced or specially trained individuals, provided the best results.

Dictionary Limitations: The dictionaries used did not meet acceptable standards of validity.

Machine Learning vs. Dictionary-Based Techniques: Machine learning, especially deep learning, significantly outperformed dictionary-based techniques, although overall human performance remained superior.

The Crowd Coders method achieved an impressive coding accuracy of 97%. The study underscored the importance of thoroughly testing automated text analysis algorithms before using them and ensuring the effectiveness and accuracy of automated text analysis projects through a detailed, step-by-step approach.

According to Ruz et al. (2020) [22], sentiment analysis using machine learning and Twitter data has become a prominent topic, especially during times of crisis such as natural disasters or social upheavals. The study applied SVM, Naïve Bayes (NB), and Random Forest (RF) Bayesian network classifiers to assess sentiment in two Spanish datasets: the 2017 Catalan independence vote and the 2010 Chilean disaster. The Bayesian network classifier used the Bayes factor method to automatically control the number of edges in the network based on the training data, resulting in more realistic networks.

The study found that SVM achieved an accuracy of 80% in both datasets, outperforming other machine learning algorithms. The resulting networks also illustrated the relationships between words, providing valuable qualitative data to understand the key aspects of event dynamics from historical and social perspectives.

According to Al-Shabi et al. (2020) [23], the rise of social media has led to a massive amount of user-generated data that can be analyzed to explore opinions and measure emotions. Sentiment analysis aims to determine the polarity of a text based on the author's viewpoint, classifying sentiments as positive, negative, or neutral.

The study identified two main categories of advanced sentiment analysis methods:

Data Mining and Machine Learning: These techniques train a model using various algorithms.

Lexicon-Based Methods: These methods infer sentiments by comparing text phrases to pre-established lexicons.

The research employed a lexicon-based approach, which achieved a classification accuracy of 72%, showing better performance in identifying both positive and negative sentiments.

According to Wazery et al. (2018) [24], many consumers review online opinions before making purchases or travel plans, creating a need for automated text sentiment analysis. Instead of manually reading through reviews, quickly gathering opinions and emotions about a specific topic can be highly beneficial.

The study utilized three main sentiment analysis techniques: Decision Trees (DT), Support Vector Machines (SVM), and Recurrent Neural Networks (RNN). It applied these techniques to three different Twitter datasets related to the Internet Movie Database, Amazon, and Airlines. The results showed that the RNN achieved the highest accuracy among the algorithms tested, with accuracy scores of 88%, 87%, and 93% respectively.

In their literature review on Bayesian networks and their application in sentiment analysis, Gutiérrez et al. (2018) [25] proposed directions for further research. Their study focused on text representation and the use of Bayesian networks for evaluating sentiments in written content. They categorized the publications into two groups: those using Bayesian networks directly for classification and those employing them to enhance categorization by extracting features and identifying correlations between variables. The review ultimately set the stage for future research into developing Bayesian network-based methods for text representation.

Chaturvedi et al. (2018) [26] identified subjectivity detection as a significant challenge in natural language processing, specifically targeting the removal of objective content from online product reviews. For sentiment analysis systems, which differentiate between positive and negative information, effective pre-processing is crucial for enhancing accuracy.

The study introduced a novel framework using an Extreme Learning Machine (ELM) model, integrating subjectivity detection with fuzzy recurrent neural networks (RNNs) and Bayesian network features. Bayesian networks were employed to establish connections between the hidden neurons in the ELM structure, ensuring effective linkage in large volumes of high-dimensional data. The fuzzy RNN was used to capture temporal properties.

The results revealed that the ELM-based framework achieved an accuracy rate of 89%. It effectively addressed challenges related to subjectivity detection and portability across languages in translation tasks.

Mariel et al. (2018) [27] discussed how deep learning represents a new era in artificial intelligence by emulating the structure and function of the human brain. This approach involves creating more complex artificial neural networks (ANNs) with multiple hidden layers. Deep Learning Neural Networks (DLNN) have proven highly effective in recognizing patterns across various types of data, including images, text, and audio.

In their study, the researchers evaluated the performance of DLNN for text classification tasks, specifically sentiment analysis, which involves assessing the sentiment expressed in a text. They compared DLNN's performance with two other well-known algorithms, Naïve Bayes (NB) and Support Vector Machines (SVM). The results demonstrated that DLNN outperformed both NB and SVM in sentiment analysis.

According to Shepelenko et al. (2017) [28], social media platforms provide a space for individuals to express and share their opinions and experiences. While some of this data is subjective and arbitrary, it can be processed to extract

valuable information for analysis and decision-making. The initial step involves understanding and categorizing this information. To achieve this, sentiment analysis was used to extract opinions from tweets.

The study aimed to explore algorithms suitable for estimating public opinion. For this purpose, Naïve Bayes (NB) and Convolutional Neural Networks (CNN) were employed and evaluated on two different datasets. The results showed that CNN achieved an average accuracy of 74%, surpassing the NB method. The study also considered the impact of training data on classifier performance.

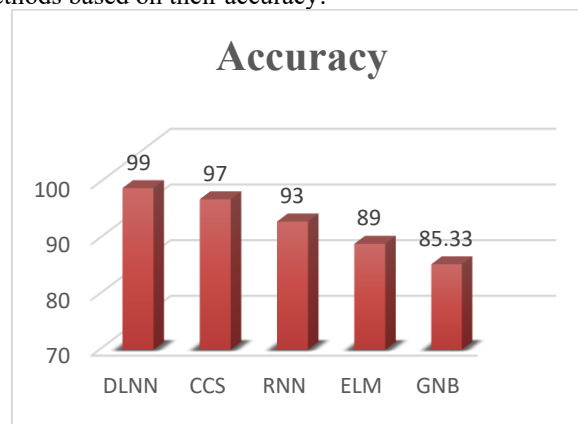
Ain et al. (2017) [29] examined the significant impact of online customer comments on various stakeholders, including readers, product suppliers, and legislators. Social media data, often disorganized, needed thorough analysis and preparation before it could be used effectively. Sentiment analysis played a crucial role in this process by categorizing content into classifications such as negative, positive, favorable, unfavorable, thumbs up, thumbs down, etc.

The study highlighted the challenges of sentiment analysis, particularly due to the scarcity of labelled data in the field of natural language processing (NLP). To address these challenges, the research integrated deep learning models, which excel at self-learning, with sentiment analysis methods. Deep learning models like Convolutional Neural Networks (CNNs) and deep neural networks were employed to overcome many of the difficulties associated with sentiment analysis.

A hybrid learning approach for text classification using natural language processing reframes: Natural languages are inherently complex and disorganized, making it difficult for models to fully grasp nuances such as tone, interpretation, and meaning, which can vary from person to person. As a result, machine learning alone is often inadequate for addressing natural language processing (NLP) challenges. While machine learning models are useful for recognizing entities in text or determining overall sentiment, they struggle with isolating specific themes or linking sentiments to particular concepts or ideas. To address these limitations, a Hybrid NLP approach incorporates rules that reflect language-specific conventions into the machine learning model. By integrating these rules and patterns, the algorithm can better align classification processes with human intuition, leading to more natural and accurate interpretations.

COMPARATIVE ANALYSIS OF EXISTING RESEARCH

This section of the article presents a comparative analysis of various deep learning approaches used for sentiment analysis, highlighting their accuracy rates. Among the popular methods for effective sentiment analysis, several techniques such as CNN, RNN, Deep Learning Neural Networks (DLNN), SVM, and NB are examined. The DLNN achieves the highest accuracy rate at 99%, followed by Crowd Coders with an accuracy of 97%. Graph provides a detailed comparison of these methods based on their accuracy.



CONCLUSION

This review provides a comprehensive examination of various machine learning methods employed for sentiment analysis at both the aspect and sentence levels. It highlights the effectiveness of supervised learning and contextual mining in sentiment analysis. The review systematically evaluates and compares several studies, summarizing their methodologies and findings, and identifies research gaps. Due to inconsistencies in assessment methods, it is challenging to determine which technique yields the best results for each specific subject. The review article presents

different strategies and their accuracy using bar graphs. Techniques reviewed include CNN, RNN, Deep Learning Neural Networks (DLNN), SVM, and NB, with DLNN achieving the highest accuracy at 99%, followed by Crowd Coders with an accuracy rate of 97%. The integration of deep learning methods with a semantic concept-centric approach points toward the development of advanced systems capable of reasoning about language and context. Future research will need to explore optimal methods to enhance accuracy and leverage various machine learning classifiers. This review can serve as a valuable resource for researchers aiming to advance in this field and offers useful guidance for further exploration.

REFERENCES

1. Vinodhini, G., and R. M. Chandrasekaran. "A sampling-based sentiment mining approach for e-commerce applications." *Information Processing & Management* 53, no. 1 (2017): 223-236.
2. Jacob, I. Jeena. "Performance evaluation of caps-net based multitask learning architecture for text classification." *Journal of Artificial Intelligence* 2, no. 01 (2020): 1-10.
3. Wadawadagi, Ramesh, and Veerappa Pagi. "Sentiment analysis with deep neural networks: comparative study and performance assessment." *Artificial Intelligence Review* 53, no. 8 (2020): 6155-6195.
4. Jospin, Laurent Valentin, Wray Buntine, Farid Boussaid, Hamid Laga, and Mohammed Bennamoun. "Hands-on Bayesian neural networks--a tutorial for deep learning users." *arXiv preprint arXiv:2007.06823* (2020).
5. Chang, Daniel T. "Probabilistic Deep Learning with Probabilistic Neural Networks and Deep Probabilistic Models." *arXiv preprint arXiv:2106.00120* (2021).
6. Jones, Evan, Tuan Do, Bernie Boscoe, Yujie Wan, Zooey Nguyen, and Jack Singal. "Photometric Redshifts for Cosmology: Improving Accuracy and Uncertainty Estimates Using Bayesian Neural Networks." *arXiv preprint arXiv:2202.07121* (2022).
7. Poria, E. Cambria, R. Bajpai, A. Hussain, A review of affective computing: from unimodal analysis to multimodal fusion, *Inf. Fusion* 37 (2017) 98–125.
8. Heredia, Brian, Taghi M. Khoshgoftaar, Joseph D. Prusa, and Michael Crawford. "Improving detection of untrustworthy online reviews using ensemble learners combined with feature selection." *Social Network Analysis and Mining* 7, no. 1 (2017): 1-18.
9. Mangat, Venu. "Dictionary based Sentiment Analysis of Hinglish text." *International Journal of Advanced Research in Computer Science* 8, no. 5 (2017).
10. Jain, Kruttika, and Shivani Kaushal. "A comparative study of machine learning and deep learning techniques for sentiment analysis." In *2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, pp. 483-487. IEEE, 2018.
11. Haberzettl, Markus, and Bernd Markscheffel. "A literature analysis for the identification of machine learning and feature extraction methods for sentiment analysis." In *2018 Thirteenth International Conference on Digital Information Management (ICDIM)*, pp. 6-11. IEEE, 2018.
12. Mishra, Abhijit, Diptesh Kanojia, Seema Nagar, Kuntal Dey, and Pushpak Bhattacharyya. "Leveraging cognitive features for sentiment analysis." *arXiv preprint arXiv:1701.05581* (2017).
13. Phan, Minh Hieu, and Philip O. Ogunbona. "Modelling context and syntactical features for aspect-based sentiment analysis." In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 3211-3220. 2020.
14. Shi, Yong, Luyao Zhu, Wei Li, Kun Guo, and Yuanchun Zheng. "Survey on classic and latest textual sentiment analysis articles and techniques." *International Journal of Information Technology & Decision Making* 18, no. 04 (2019): 1243-1287.
15. Ahmad, Munir, Shabib Aftab, Syed Shah Muhammad, and Usman Waheed. "Tools and techniques for lexicon driven sentiment analysis: a review." *Int. J. Multidiscip. Sci. Eng* 8, no. 1 (2017): 17-23.
16. Yusof, Nor Nadiyah, Azlinah Mohamed, and Shuzlina Abdul-Rahman. "A review of contextual information for context-based approach in sentiment analysis." *International Journal of Machine Learning and Computing* 8, no. 4 (2018): 399-403.
17. Dabrowski, Joel Janek, Johan Pieter de Villiers, and Conrad Beyers. "Context-based behaviour modelling and classification of marine vessels in an abalone poaching situation." *Engineering Applications of Artificial Intelligence* 64 (2017): 95-111.
18. Bardool, Kevin, Tinne Tuytelaars, and José Oramas. "A Systematic Analysis of a Context Aware Deep Learning Architecture for Object Detection." In *BNAIC/BENELEARN*. 2019.

19. Alam, Mohammad Arif Ul. "Activity-Aware Deep Cognitive Fatigue Assessment using Wearables." In 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 7433-7436. IEEE, 2021.
20. Santana, Tiago MHC, Keiller Nogueira, Alexei MC Machado, and Jefersson A. dos Santos. "Deep contextual description of superpixels for aerial urban scenes classification." In 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3027-3031. IEEE, 2017.
21. Santana, Tiago MHC, Ricardo da S. Torres, and Jefersson A. dos Santos. "Superpixel Context Description based on Visual Words Co-Occurrence Matrix." In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, pp. 6199-6202. IEEE, 2018.
22. Samal, BiswaRanjan, Anil Kumar Behera, and Mrutyunjaya Panda. "Performance analysis of supervised machine learning techniques for sentiment analysis." In 2017 Third International Conference on Sensing, Signal Processing and Security (ICSSS), pp. 128-133. IEEE, 2017.
23. A A Gulhane "A Study of Deep Learning and NLP", International Journal of Research Publication and Reviews, Vol 3, no 5, pp 2735-2742, May 2022. ISSN 2582-7421.
24. Bilgin, Metin, and İzzet Fatih Şentürk. "Sentiment analysis on Twitter data with semi-supervised Doc2Vec." In 2017 international conference on computer science and engineering (UBMK), pp. 661-666. Ieee, 2017.
25. Ghiassi, Manoochehr, and S. Lee. "A domain transferable lexicon set for Twitter sentiment analysis using a supervised machine learning approach." Expert Systems with Applications 106 (2018): 197-216.
26. Ahmad, Munir, Shabib Aftab, Syed Shah Muhammad, and Sarfraz Ahmad. "Machine learning techniques for sentiment analysis: A review." Int. J. Multidiscip. Sci. Eng 8, no. 3 (2017): 27.
27. Manek, Asha S., P. Deepa Shenoy, and M. Chandra Mohan. "Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier." World wide web 20, no. 2 (2017): 135-154.
28. Taneja, Shweta, Siddharth Bhasin, and Sambhav Kapoor. "Trends and Sentiment Analysis of Movies Dataset Using Supervised Learning." In Proceedings of International Conference on Intelligent Cyber-Physical Systems, pp. 331-342. Springer, Singapore, 2022.
29. Van Atteveldt, Wouter, Mariken ACG van der Velden, and Mark Boukes. "The validity of sentiment analysis: Comparing manual annotation, crowd-coding, dictionary approaches, and machine learning algorithms." Communication Methods and Measures 15, no. 2 (2021): 121-140.
30. Ruz, Gonzalo A., Pablo A. Henríquez, and Aldo Mascareño. "Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers." Future Generation Computer Systems 106 (2020): 92- 104