

Hadoop-Based Big Data Sentiment Analysis Using Machine Learning

Hamsitha Challagundla^a, Athulya Biju^{b,*}, Aaditri Mittal^c, Philip Eugene Abraham^d

^aSoftware Engineer 2, Sabre India, India.

^bSoftware Engineer-1, Honeywell, India.

^cSoftware Engineer 1 B, Bank of America, India.

^dSWE, Wipro, India

Abstract

Sentiment Analysis (SA) and Opinion Mining have emerged as critical research areas in the exponential increase in sentiment-rich social media information on the web. SA has changed into a challenging task with the emergence of Big Data. This study presented an effective method to perform SA on large-scale datasets of tweets employing Machine Learning algorithms in the Hadoop ecosystem. In particular, we developed and executed Naive Bayes(NB), Recurrent Neural Networks (RNN), and Support Vector Machines (SVM) classification algorithms for evaluating datasets, and the effectiveness of the strategy was measured in retrieval matrices—Precision, Recall, F-score, and Accuracy. The experimental outcomes show that our method handles vast sentiment datasets with exceptional efficiency, and notably, SVM outperformed other classifiers by achieving an outstanding accuracy of 95% and a ROC of 93%. According to the results of our analysis, SVM stands out as a reliable choice in important sectors where SA plays a crucial role. This work substantially contributes to the area of SA in the Big Data realm by providing a flexible, quick, and scalable method for massively evaluating sentiment-rich social media content. Additionally, it will be more helpful for increasing significant value to businesses, the government, and individuals.

Keywords: Machine Learning, Big Data, Sentiment Analysis, Twitter, Lexicon, Hadoop distributed file system.

1 Introduction

The 21st century has brought forth an era of unheard-of data growth when information has exploded across numerous fields and given rise to Big Data. Big Data refers to enormous and varied amounts of information that can be analyzed to produce valuable knowledge and data [1]. SA relies heavily on this popular field of computer science study because it incorporates the vast volumes of organized, semi-structured, and unstructured data accessible through the internet, social media, remote sensing, and medical records [2]. Rapid technological advancements in the digital sphere have made society a tech-savvy populace that depends on digital sensors, communication tools, social media applications, and data processors [3, 4, 5]. The Internet of Things (IoT) results from this digital revolution, where interconnected gadgets produce massive volumes of real-time data that keep growing as we interact with web-connected digital devices. Utilizing conventional database management software and hardware to manage and analyze unstructured data is difficult due to the amount of data [6, 7].

Large datasets have been generated by widely used social media platforms like Twitter, Facebook, YouTube and e-commerce sites like Flipkart and Amazon [8] as a result of rising digitalization [9]. Consequently, businesses have turned to Big Data for insights, revolutionizing fields like marketing, fraud detection [10], healthcare, and e-reputation management [11]. People can generate enormous amounts of big data through online communities, smartphones, laptops, and PCs [12]. In particular, with the vast amounts of user-generated content, Twitter alone sends 500 million tweets daily, of which 40 million are shared. At the same time, Facebook sees 4.3 billion messages received daily and 5.75 billion likes [13]. The proliferation of digital technologies is the primary cause of this ongoing growth, which guarantees that the amount of data will keep growing, reshaping the information landscape and highlighting the significance of cutting-edge data analysis techniques like SA to generate insightful conclusions [14] [15].

SA [16] seeks to ascertain a speaker's or writer's perspective on a particular subject or the general contextual polarity of a publication. The attitude could represent the person's judgment, evaluation, affect, or purposeful emotional expression [17]. SA can be used to extract useful information in this large ocean of big social data. It is a new area of research that involves forecasting attitudes toward goods or social entities using feelings, and it is essential for modeling business strategies to meet business objectives [18]. People can better grasp the findings of sentiment categorization with the aid of sentiment visualization [19].

Big Data encompasses not just the sheer amount of data but also its velocity and variety, which can be structured or unstructured and arrive in various formats, including real-time streaming data [20], [21]. Industries depend on robust solutions like Apache Hadoop and Apache Spark [22] to handle this data flood. Data is stored using Hadoop's fault-tolerant, scalable, and modular architecture, and processing speed is increased using Spark's lightning-fast cluster computing engine with in-memory cluster computation [23]. Hadoop is a Java-based Apache open-source platform that enables distributed processing of huge datasets across computer clusters using a straightforward programming model [24].

2 Related Work

Ramesh et al.[3] concentrated on SA and opinion mining about social media data, notably on platforms like Twitter, and discovered effective methods for doing sentiment analysis on Big Data sets utilizing tools like Hadoop [25]. The explosive expansion of sentiment-rich social media information offers invaluable insights for businesses to understand how consumers view their brands. It highlights the growing interest among researchers in this field. According to the experimental findings, the technique used in the study is highly effective at processing big sentiment data sets.

The exponential growth of text data from various sources, including social networks, e-commerce sites, and scientific investigations, has become useful for companies looking for consumer insights in the big data era. BARZANJI et al.[26] proposed a system for big data SA that handled massive sentiment datasets with great accuracy and provided substantial value to businesses, governments, and people. Open-source big data technologies and machine learning methods, such as NB, SVM, and Apache Spark, are crucial for processing this enormous volume of text data in real-time. Another study [27] employed multiple SA approaches (Linear SVC, Logistic Regression, and NB) that were investigated on sizable Amazon Fine Food review datasets using the data processing capabilities of Apache Spark and the MLlib package. The results showed Linear SVC's outstanding performance, delivering over 80% accuracy and offering priceless insights for decision-making.

According to AWAJAN et al. [28], the massive volume of online consumer reviews (OCRs) produced by social media, web content, and microblogs have increased the significance of SA or opinion mining. An innovative method for evaluating products based on online reviews combines neutrosophic set theory with sentiment analysis and multi-attribute decision-making. The SNNWA operator and the cosine similarity measure rank the alternatives, whereas Neutro-VADER efficiently handles neutral data. For real-time, massively scalable data analytics in cybersecurity, authors[29] offered MOLESTRA, a Multi-Task Learning model utilizing the "Kappa" architecture. It successfully differentiates short-term and long-term memory for better performance using k-NN Classifier with Self-Adjusting Memory (k-NN SAM). The difficulty of real-time stream data processing and the requirement for online analytical algorithms to handle this sort of data stress the significance of preprocessing data to keep just summaries in main memory while ensuring that high-speed data arrival does not lead to data loss. Authors [30] introduced Sentinel, a Java-based distributed system. Sentinel used the parallel decision-tree-learning Vertical Hoeffding Tree learner method for distributed processing and learning on top of Apache Storm as its fundamental distributed computing platform [31, 32, 33]. A synopsis data structure is maintained using SpaceSaving to summarize the data stream [34, 35].

3 Proposed Method

The research activity described in this study is divided into five stages: data preprocessing, feature selection from the dataset, uploading the preprocessed data to Hadoop distributed file system (HDFS), sentiment analysis of the data, categorization of their polarity, and evaluation of results using classifiers. Figure 1 depicts the proposed method's design. Figure 2 depicts the fundamental roles of preprocessing approaches.

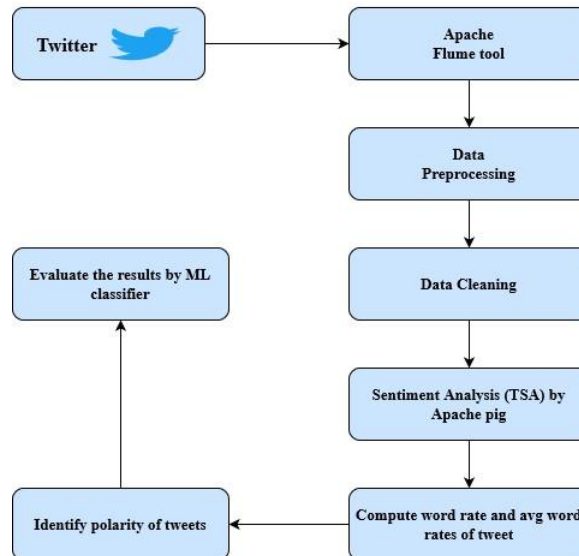


Fig 1 Framework of the proposed method

3.1 Dataset Description & Preprocessing

114940 data instances and 15 tweet-related attributes are included in the demonization dataset. Using a dataset, tweets are divided into three categories: positive, negative, and neutral. The data from Twitter is unsuitable for analysis. As a result, Table 1 displays the preprocessed data that Algorithm 1 applied. The technique of reducing a stream of sentences to a group of words is known as tokenization. Stemming, Second refers to the process of reducing words to their stems. For grammatical purposes, the same term can be used in various synonyms, such as Search and Searches Searching. Case The text is converted into either lowercase or uppercase at this phase. To remove the punctuation marks from the term, punctuation must be removed from a text because it lacks information. Stop words like prepositions, help verbs, articles, etc., are removed using this technique. The data was preprocessed, and the final dataset was uploaded to HDFS. Later, the dataset is analyzed by Apache Pig, shown in Algorithm 2.

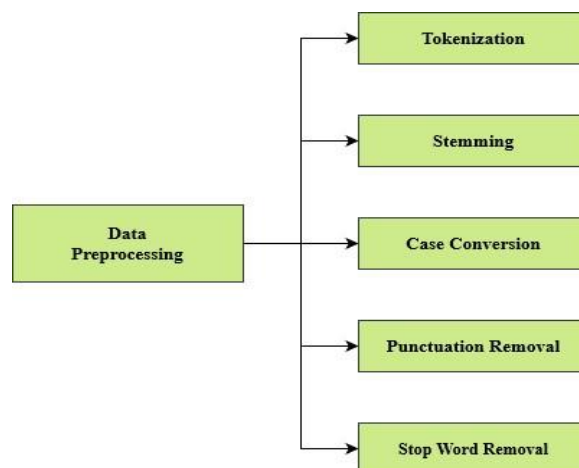


Fig 2 Basic preprocessing

Algorithm 1 Twitter Dataset Preprocessing

Require: D : Twitter dataset
Ensure: RD : Preprocessed dataset

```
1: function PREPROCESSTWITTERDATASET( $D$ )
2:    $RD \leftarrow \{\}$ 
3:   for  $tweet$  in  $D$ , do
4:      $cleaned\ tweet \leftarrow CLEANUPTWEET(tweet)$ 
5:     if HASEMOTICONTAGS( $cleaned\ tweet$ ) then
6:        $RD \leftarrow RD \cup \{cleaned\ tweet\}$  ▷ Add the cleaned tweet to the pre-processed dataset
7:   return  $RD$ 
8: function CLEANUPTWEET( $tweet$ )
9:   /* Add tweet cleaning process here */
10: function HASEMOTICONTAGS( $tweet$ )
11:   /* Add emoticon tag detection process here */
```

Algorithm 2 Twitter Sentiment Analysis (TSA)

Require: dataset: number of tweets, dictionary words, and their rates
Ensure: categorization: positive, negative, or neutral sentiment

```
1: function TWITTERSENTIMENTANALYSIS( $dataset, dwr, Twt, Cs$ ) 2: Load the dataset into Pig-x
local mode using pig storage.
3: Extract id and  $Twt$  from the  $dataset$ .
4: for each  $Twt$  in  $Cs$  do ▷ Loop through each attribute  $Twt$  in the corpus  $Cs$  5: Divide  $Twt$ 
text into words.
6: Load Dictionary and join token and dictionary.
7: for each word in words do
8:   if word == any word in  $dwr$  then
9:      $wr \leftarrow dwr$ 
11:  $Avg\ R \leftarrow Avg(wr)$ 
12: if  $Avg\ R \geq 0.0$ , then
13: Categorization: Tweet is positive 14: else if  $Avg\ R < 0.0$ , then
15:   Categorization: Tweet is negative
16:   else
17:     Categorization: Tweet is neutral
18: Return  $Avg\ R$ 
```

3.2 Apache Hadoop Ecosystem

- HDFS HDFS is a Hadoop file system that implements a distributed file system design. It can contain a lot of data and provide multiple clients an easy way to access it over the network. It is extremely fault-tolerant and designed to run on commodity hardware, which is inexpensive hardware. A file system called HDFS uses a block structure and loads files onto Hadoop clusters in fixed-size blocks. A Master-slave architecture is implemented by the HDFS using Name and Data nodes. While the Name node operated as the master, several Data nodes functioned as slaves.
- Map Reduce (Distributed Data Processing) The MapReduce system model is used to process massive amounts of data. The Map/Reduce algorithm uses the divide-and-conquer strategy, in which a large, complex problem is divided into smaller ones that are solved concurrently and independently by distributed clusters. Finally, the problem's complete answer was created by combining all the interim solutions.

- Apache Spark, The distributed computing engine Apache Spark, may be used with various workloads and platforms. Spark leverages a variety of paradigms, such as Spark Streaming, Spark ML, Spark SQL, and Spark Graph x, to bind to different networks and process varied data workloads. Apache Spark is a speedy in-memory data processing engine that enables data workers to effectively complete streaming machine learning or SQL tasks that call for immediate interactive access to data sets. Apache Spark also has beautiful and descriptive development APIs. Spark offers real-time streaming, querying, machine learning, and graph processing. Before Apache Spark was developed, several methods were employed for various workloads. Each method for batch analytics, interactive queries, real-time streaming processing, and machine learning has its database. Apache Spark can complete all of these tasks rather than relying on several technologies that aren't always compatible.

3.3 Sentiment Analysis

This research aimed to develop a method for performing SA more quickly because a large volume of data must be processed. Additionally, care had to be taken to avoid seriously sacrificing precision in favor of speed. Combining Big Data and Hadoop makes SA on Big Data possible. An author's expression or opinion on any subject or component is their sentiment. The main goals of SA are finding the opinion word in the text and parsing it. Opinion words are identified, and then sentiment scores are given to these terms. Lastly, one must establish the polarity of the sentence. Polarity can be positive, negative, or neutral. Use the Lexicon approach to classify sentiment. The parsing procedure involves breaking the sentence up into words. This action is also known as tokenization. These tokenized words are used to categorize input words that express opinions. Using distributed processing by Apache Pig, NLTK was used to classify all comments into good, negative, and neutral.

Algorithm 3 Sentiment Analysis

```
1: procedure SENTIMENTANALYSIS(Tweets, SentiWord Dictionary)
2:   for tweet  $T_i$  in Tweets do
3:      $Sentiscore[T_i] \leftarrow 0$ 
4:     for the word  $W_j$  in  $T_i$  if  $W_j$  is listed in SentiWord Dictionary do
5:        $SentiScore[W_j] \leftarrow 0$ 
6:       if  $polarity[W_j] = \text{blind negation}$  then
7:         return "negative"
8:       else
9:         if  $polarity[W_j] = \text{positive}$  and  $strength[W_j] = \text{Strongsubj}$  then
10:           $Sentiscore[T_i] \leftarrow Sentiscore[T_i] + 1$ 
11:        else if  $polarity[W_j] = \text{positive}$  and  $strength[W_j] = \text{weak}$  then
12:           $Sentiscore[T_i] \leftarrow Sentiscore[T_i] + 0.5$ 
13:        else if  $polarity[W_j] = \text{negative}$  and  $strength[W_j] = \text{Strongsubj}$  then
14:           $Sentiscore[T_i] \leftarrow Sentiscore[T_i] - 1$ 
15:        else if  $polarity[W_j] = \text{negative}$  and  $strength[W_j] = \text{weaksubj}$  then
16:           $Sentiscore[T_i] \leftarrow Sentiscore[T_i] - 0.5$ 
17:        else if  $polarity[W_j] = \text{negation}$  then
18:           $Sentiscore[T_i] \leftarrow Sentiscore[T_i] - 1$ 
19:        if  $Sentiscore[T_i] > 0$  then
20:          return "positive"
21:        else if  $Sentiscore[T_i] < 0$  then
```

22: return" negative"
 23: else
 24: return" unaffected"

4. Result

The information was cleansed, preprocessed, and categorized into three groups: neutral, negative, and positive. Apache Pig and Apache Flume are utilized in the Hadoop ecosystem. After the dataset had been categorized, three algorithms, including SVM, NB, and RNN, were selected to carry out the classification task. Every tuple in the dataset is mapped during classification to create the final variables. It is claimed that the output variables are class labels. Every tuple from the supplied dataset has its class label predicted by the mapping method. The accuracy of the class variable prediction for the test data is used to evaluate the classifiers' performance analysis. As was previously stated, this research aimed to develop a technique that could swiftly compute the feelings of big data sets without significantly sacrificing accuracy. In terms of speed, the suggested method has performed exceptionally well. Using the following information retrieval matrices—Precision, Recall, F-score, and Accuracy—we will assess the outcomes of our experiment. Time to choose an algorithm, split our data into training and testing sets (80% and 20%, respectively), and get started! The first algorithm to be used is the NB classifier. This text categorization algorithm is highly popular. The training package will comprise the first 80% of randomly selected reviews, including good and negative feedback. The model's accuracy was then determined by comparing the outcomes to the final 20%. Simply executing the NB was the first step, followed by training everything in a single line using Train (). After being educated, it is simple enough. After that, you may try it: We can still" test" the data because we know the right responses. We can still" test" the data because we know the right responses. As a result, display the data to the machine while checking without giving the right answer. The machine is correct if it properly guesses what it knows the answer to be. Given the shifting done, various degrees of accuracy can be reached. The results are displayed in Table 1. The findings are shown in

Table 1 Performance Analysis of Classification

Metrics	SVM	NB	RNN
Precision	90%	75%	87%
Recall	92%	77%	89%
F-score	91%	76%	88%

Table 1 of our study clearly shows each method's advantages and disadvantages when it comes to SA on large datasets. SVM attained a remarkable precision of 90%, demonstrating its superior ability to distinguish between positive and negative emotions. The SVM can successfully capture a significant amount of both positive and negative attitudes in the dataset, as evidenced by the high recall value of 92%. The corresponding F-score of 91% further supports its overall competence in sentiment analysis tasks. On the other hand, NB demonstrated lower precision and recall.

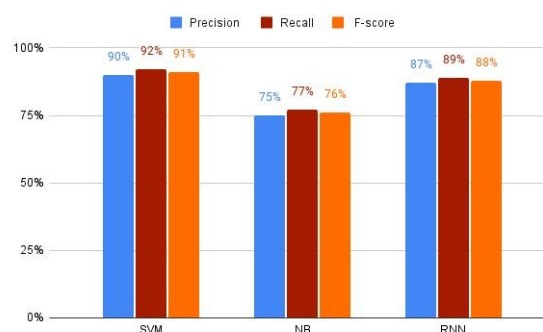


Fig 3 Performance Analysis of Classification

values of 75% and 77%, although computationally efficient. This suggests that NB might have trouble correctly identifying sentiment classes, which could result in the loss of crucial sentiment data. RNN produced great results with a precision of 87% and a recall of 89% by utilizing its capacity to capture contextual information and dependencies in sequential data. The F-score of 88% confirms RNN's excellent performance in big data SA jobs. According to our research, SVM is a solid option for applications where high precision is essential, including SA in vital industries like banking or healthcare. However, NB might still be a good choice if computational effectiveness is a top goal and a minor accuracy loss is acceptable. RNN is an appealing choice for SA jobs requiring context-aware analysis of massive amounts of textual data because of its competitive precision, recall, and F-score values.

Table 4 displays the roc and accuracy performance.SVM demonstrated a remarkable ROC value of 93%, which suggests that it is adept at accurately differentiating between positive and

Table 2 ROC and Accuracy Analysis of TSA Algorithm

Classifier	ROC	Accuracy
SVM	93%	95%
NB	79%	82%
RNN	90%	92%

Negative attitudes. Further confirmation of SVM's strong performance in accurately categorizing sentiments in the text data is provided by the equivalent accuracy of 95%.NB achieved an accuracy of 82% while earning a modest ROC of 79%, showing that it performs reasonably well in SA tasks. However, unlike SVM, NB's performance could be impacted by its probabilistic structure and simplifying assumptions.RNN produced a noteworthy ROC of 90% and an accuracy of 92% by capturing sequential dependencies in data. These findings demonstrate RNN's

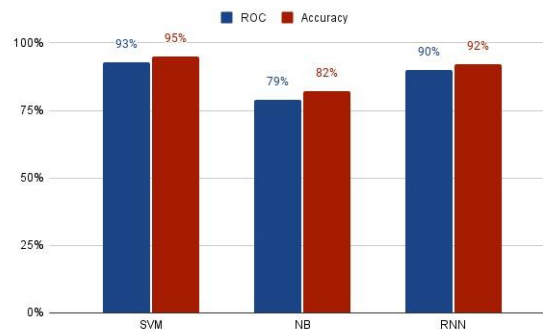


Fig 4 ROC and Accuracy Analysis of TSA Algorithm

capability to accurately and contextually interpret sentiment in text data, making it an appealing option for TSA tasks. As a result, choosing the best classifier for particular sentiment analysis tasks is made easier with knowledge of the performance characteristics of SVM, NB, and RNN in TSA. Researchers can further the field of TSA and its applications across numerous sectors and disciplines by utilizing these results.

5. Conclusion

Sentiment analysis is now applied in various applications and will likely be applied in several other contexts. Its applications will spread to additional domains, inspiring more and more regional research. This study aims to manifest big data SA's significance and efficiency in extracting user sentiments on a large scale and in real time. Employing Apache Spark, a robust big data tool, the proposed model efficiently processes extended data volumes, enabling effective SA of big data. Throughout our study, we reviewed cutting-edge explanations in sentiment classification, analyzing the essentials for a scalable approach capable of efficiently handling large data sets. By using

Hadoop, we acquired successful classification of Twitter data, and our approach showed outstanding performance. We anticipate potential developments as technology develops that will deal with the difficulties faced at the moment, paving the way for even more thorough and accurate SA of massive data sets. Ultimately, this study gives up fresh perspectives for analyzing user attitudes on a big scale, which is advantageous to several sectors, including marketing, customer feedback analysis, and social media monitoring.

References

- [1] R. Addo-Tenkorang and P. T. Helo, "Big data applications in operations/supply-chain management: A literature review," *Computers & Industrial Engineering* 101, 528–543 (2016).
- [2] N. Elgendy and A. Elragal, "Big data analytics: a literature review paper," in *Advances in Data Mining. Applications and Theoretical Aspects: 14th Industrial Conference, ICDM 2014, St. Petersburg, Russia, July 16-20, 2014. Proceedings 14*, 214–227, Springer (2014).
- [3] D. Kurian, S. Vishnupriya, R. Ramesh, *et al.*, "Big data sentiment analysis using hadoop," *International Journal for Innovative Research in Science and Technology* 1(11), 92–96 (2015).
- [4] A. Zaslavsky, C. Perera, and D. Georgakopoulos, "Sensing as a service and big data," *arXiv preprint arXiv:1301.0159* (2013).
- [5] S. Madakam, V. Lake, V. Lake, *et al.*, "Internet of things (iot): A literature review," *Journal of Computer and Communications* 3(05), 164 (2015).
- [6] A. Dwivedi, R. Pant, S. Pandey, *et al.*, "Internet of things' (iot's) impact on decision oriented applications of big data sentiment analysis," in *2018 3rd International Conference On Internet of Things: Smart Innovation and Usages (IoT-SIU)*, 1–10, IEEE (2018).
- [7] P. Zikopoulos and C. Eaton, *Understanding big data: Analytics for enterprise class hadoop and streaming data*, McGraw-Hill Osborne Media (2011).
- [8] K. F. Sudhakar and H. SYED, "A comparative study between flipkart and amazon india," *Anveshana's International Journal of Re-search in Regional Studies, Law, Social Sciences, Journalism and Management Practices* 1(8), 36–46 (2016).
- [9] A. P. Rodrigues, A. Rao, and N. N. Chiplunkar, "Sentiment analysis of real time twitter data using big data approach," in *2017 2nd International conference on computational systems and information technology for sustainable solution (CSITSS)*, 1–6, IEEE (2017).
- [10] A. Abdallah, M. A. Maarof, and A. Zainal, "Fraud detection system: A survey," *Journal of Network and Computer Applications* 68, 90–113 (2016).
- [11] El Alaoui and Y. Gahi, "The impact of big data quality on sentiment analysis approaches," *Procedia Computer Science* 160, 803–810 (2019).
- [12] I. Manyika, M. Chui, B. Brown, *et al.*, *Big data: The next frontier for innovation, competition, and productivity*, McKinsey Global Institute (2011).
- [13] A. Yasin, Y. Ben-Asher, and A. Mendelson, "Deep-dive analysis of the data analytics workload in cloudsuite," in *2014 IEEE International Symposium on Workload Characterization (IISWC)*, 202–211, IEEE (2014).
- [14] R. L. Villars, C. W. Olofson, and M. Eastwood, "Big data: What it is and why you should care," *White paper, IDC 14*, 1–14 (2011).
- [15] M. Khoso, "How much data is produced every day?[information post]. northeastern university," (2016).
- [16] M. Thelwall, "Sentiment analysis for small and big data," *The SAGE handbook of online research methods*, 344–355 (2016).
- [17] M. Shinde-Pawar, "Formation of smart sentiment analysis technique for big data," *Methods* 2(12) (2014).
- [18] R. K. Behera, M. Jena, S. K. Rath, *et al.*, "Co-lstm: Convolutional lstm model for sentiment analysis in social big data," *Information Processing & Management* 58(1), 102435 (2021).
- [19] M. I. Al-mashhadani, K. M. Hussein, E. T. Khudir, *et al.*, "Sentiment analysis using optimized feature sets in different facebook/twitter dataset domains using big data," *Iraqi Journal For Computer Science and Mathematics* 3(1), 64–70 (2022).

- [20] Haryanto, Y. Ruldeviyani, F. Rohman, *et al.*, "Facebook analysis of community sentiment on 2019 Indonesian presidential candidates from Facebook opinion data," *Procedia Computer Science* 161, 715–722 (2019).
- [21] A. Rajput, "Natural language processing, sentiment analysis, and clinical analytics," in *Innovation in health informatics*, 79–97, Elsevier (2020).
- [22] S. Salloum, R. Dautov, X. Chen, *et al.*, "Big data analytics on Apache Spark," *International Journal of Data Science and Analytics* 1, 145–164 (2016).
- [23] S. D. Shetty, "Sentiment analysis, tweet analysis and visualization on big data using Apache Spark and Hadoop," in *IOP Conference Series: Materials Science and Engineering*, 1099(1), 012002, IOP Publishing (2021).
- [24] V. B. Bobade, "Survey paper on big data and Hadoop," *Int. Res. J. Eng. Technol* 3(1), 861–863 (2016).
- [25] T. Hussain, A. Sanga, and S. Mongia, "Big data Hadoop tools and technologies: A review," in *Proceedings of International Conference on Advancements in Computing & Management (ICACM)*, (2019).
- [26] Al-Barznji and A. Atanassov, "Big data sentiment analysis using machine learning algorithms," in *Proceedings of 26th International Symposium "Control of Energy, Industrial and Ecological Systems, Bankia, Bulgaria"*, (2018).
- [27] H. M. Ahmed, M. Javed Awan, N. S. Khan, *et al.*, "Sentiment analysis of online food reviews using big data analytics," *Hafiz Muhammad Ahmed, Mazhar Javed Awan, Nabeel Sabir Khan, Awais Yasin, Hafiz Muhammad Faisal Shehzad (2021) Sentiment Analysis of Online Food Reviews using Big Data Analytics. Elementary Education Online* 20(2), 827–836 (2021).
- [28] I. Awajan, M. Mohamad, and A. Al-Quran, "Sentiment analysis technique and neutrosophic set theory for mining and ranking big data from online reviews," *IEEE Access* 9, 47338–47353 (2021).
- [29] K. Demertzis, L. Iliadis, and V.-D. Anezakis, "Molestra: a multi-task learning approach for real-time big data analytics," in *2018 Innovations in Intelligent Systems and Applications (INISTA)*, 1–8, IEEE (2018).
- [30] A. H. A. Rahnama, "Distributed real-time sentiment analysis for big data social streams," in *2014 International conference on control, decision and information technologies (CoDIT)*, 789–794, IEEE (2014).
- [31] Dwaram, J. R., & Madapuri, R. K. (2022). Crop yield forecasting by long short-term memory network with Adam optimizer and Huber loss function in Andhra Pradesh, India. In *Concurrency and Computation: Practice and Experience* (Vol. 34, Issue 27). Wiley. <https://doi.org/10.1002/cpe.7310>
- [32] Kumar, M. R., & Rani, Z. P. (2019). Energy-efficient Scheduling of Cloud Data Center Servers. In *International Journal of Innovative Technology and Exploring Engineering* (Vol. 8, Issue 11, pp. 1769–1772). Blue Eyes Intelligence Engineering and Sciences Engineering and Sciences Publication - BEIESP. <https://doi.org/10.35940/ijitee.k1759.0981119>
- [33] Kumar, V. K. A., Kumar, M. R., Shribala, N., Singh, N., Gunjan, V. K., Siddiquee, K. N., & Arif, M. (2022). Dynamic Wavelength Scheduling by Multiobjectives in OBS Networks. In N. Jan (Ed.), *Journal of Mathematics* (Vol. 2022, pp. 1–10). Hindawi Limited. <https://doi.org/10.1155/2022/3806018>
- [34] Rudra Kumar, M., & Gunjan, V. K. (2022). Peer Level Credit Rating: An Extended Plugin for Credit Scoring Framework. In *Lecture Notes in Electrical Engineering* (pp. 1227–1237). Springer Nature Singapore. https://doi.org/10.1007/978-981-16-7985-8_128
- [35] Prasad*, M. N., Babu, A. R., Krishna, A. V., & Kumar, Dr. M. R. (2019). Reciprocal Repository for Decisive Data Access in Disruption Tolerant Networks. In *International Journal of Innovative Technology and Exploring Engineering* (Vol. 9, Issue 1, pp. 4430–4434). Blue Eyes Intelligence Engineering and Sciences Engineering and Sciences Publication - BEIESP. <https://doi.org/10.35940/ijitee.a5249.119119>