Journal of Information Systems Engineering and Management
2025, 10(19s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Empowering Public Trust in Vaccines for Effective Outbreak Response

Alok Soreng [1], Kailash Chandra Bandhu[2]
[1] *Department of Computer Science and Engineering, Medi-Caps University, Indore, India*
Email: lksoreng133@gmail.com

ARTICLE INFO

Received: 25 Dec 2024

Revised: 28 Jan 2025

Accepted: 16 Feb 2025

ABSTRACT

Social media sites like Twitter act as vibrant centres of worldwide communication. highly influential due to the large number of users and the constant stream of information shared. This can be both positive and negative. On the negative side, social media can be a breeding ground for misinformation and can exacerbate public anxieties, especially during crisis situations like pandemics. This paper highlights Understanding how important it is to measure public opinion on social media sites is critical to surviving in today's digital environments, especially regarding concerns lingering after a pandemic. By analyzing these concerns, policymakers can make better decisions to address public anxieties and improve public health policies in the future. Preprocessing involves tokenization, stop word removal, stemming/lemmatization, and normalization. Each preprocessed token is then embedded using an LSTM layer to capture context and sequence. An attention mechanism focuses on crucial aspects of the text. Sentiment analysis with polarity and objectivity ranking is performed alongside context extraction. Finally, an LSTM classifier leverages both sentiment and context features to categorize the text. This approach offers a comprehensive method for text classification considering sentiment, context, and inherent textual features. proposed approach achieved higher accuracy (92.8%) compared to Random Forest and LSTM models used with these embeddings.

Keywords: BI-LSTM, COVID-19 Vaccine, Negative Sampling, LSTM, Random Forest

1. INTRODUCTION

Social media and twitter trends are highly influential with millions of users, provide large number of opinion and have the power shape sensitive opinions. With just a blink of an eye people can criticise organisations and revolt other people thoughts. In between all the chaotic scenarios such as virus outbreaks can cause huge distress among the community. Thus it is important to address the public sentiments by hush group and choose methods such as banning accounts or any other legal actions but not go for the root cause as to why it has been occurred in the first place. (Arbane et al., 2023) indicates that there are over 340 million user communication 500 million tweets per day. Recently we have been recovered from pandemic of covid-19, which arise various issues related to vaccines, (Qorib et al., 2023)more work required as it is still ongoing research. According to the WHO hesitancy has been increased worldwide. That is where the social media comes, allowing people to express. It is important to understand the mental stress and concern as to why even after vaccination people express hesitant towards crowded places especially during times gathering such as elections. During elections people get tweets from all over the places and also there is a huge chance of getting news full of fallacy. Best possible findings to prevent future misleading news is by understanding the sentiment of people during pandemic times such as Covid-19. Following are some existing literatures on Covid-19 which foretells that public policies require to have a thorough understand of the public concerns and safety even after pandemic is over. It is crucial and can be beneficial for future for making public policies and also for undiscovered viruses later on.

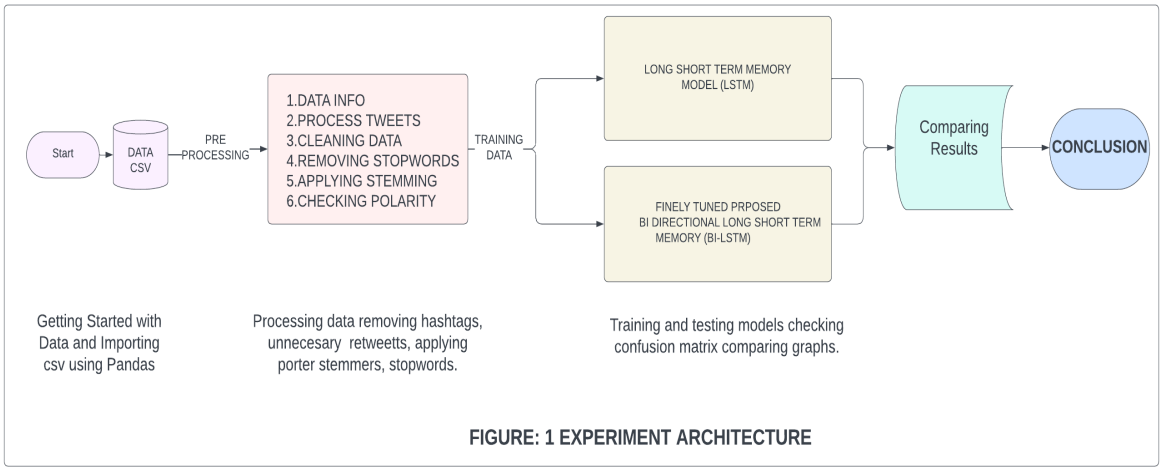

FIGURE: 1 EXPERIMENT ARCHITECTURE

In the given figure 1, work is being done by taking the twitter dataset on vaccine based on covid-19. The data contain Range-Index: 209929 and a total of 36 data columns with d-types: float64(11), int64(8), object (17) comprises of only taken English text. Pre-Processed Data will then get transferred to train for BI-LSTM and Proposed BI-LSTM hence perform state-of-art performance.

Work is organised by section II literature review of existing paper on sentimental analysis followed by table 1 of literature findings, section III discussion on methodology and study of the approaches and section IV results and contribution followed by tables, graphs and V consist of conclusions and thereafter section VI contains references.

2. LITERATURE REVIEW

Social media platforms and online forums offer an accessible avenue for individuals to exchange and deliberate on public health concerns, like COVID-19, circulating both factual information and misinformation. This study introduces an NLP approach, anchored on the Bidirectional Long Short-Term Memory (Bi-LSTM) mechanism, aiming to classify sentiments and pinpoint concerns related to public sentiments on COVID-19. Bi-LSTM refines traditional LSTMs by synthesizing outputs from both previous and subsequent data contexts. Using datasets from Twitter and Reddit, our findings demonstrated superior performance metrics when compared to traditional LSTM models and extant research. This concept provides governmental organizations with a mechanism to understand public mood during health crises and counteract negative messages. Further, our study highlights the importance of using NLP to public sentiment analysis, providing guidance for the development of health policy [1].

Following the 2019 COVID-19 outbreak, which is widely thought to have started in Wuhan, China, communities around the world implemented preventive measures like using face masks, cleaning their hands, and avoiding close contact. Early in 2021, countries—including the United States—started administering COVID-19 immunization providing respite to the public but also igniting divisive discussions. Influenced by these conversations, vaccine reluctance became to be a major problem. Given conventional data limitations, live-streamed tweets, fetched through API searches, emerged as a credible source to probe public sentiment on vaccine apprehension. This research leveraged three sentiment analysis techniques (Azure Machine Learning, VADER, and Text-Blob in conjunction with five machine learning algorithms and three text vectorization methods. Through rigorous model evaluation, we observed a gradual positive shift in public sentiment toward the COVID-19 vaccine. Notably, to create a pleasing fusion. Text-Blob sentiment scoring, TF-IDF vectorization, and Linear SVC classification yielded the best performance metrics, highlighting its efficiency in discerning public sentiment [2].

Background: Internet-based syndromic surveillance has been pivotal in predicting epidemic trends over the past two decades, using diverse data sources, from search engines to social platforms. Recently, the focus shifted towards understanding the public's emotional response to health crises, especially pandemics. Objective: This study sought to gauge real-time sentiment reactions to COVID-19 cases in Greece through Twitter data. Methods: Over one year, 153,528 tweets from 18,730 users were amassed and analyzed against two sentiment lexicons - one translated from English to Greek and an original Greek lexicon. The research traced both overall positive and negative sentiments and six distinct emotions. Additionally, correlations between sentiments, tweet volumes, and

actual COVID-19 cases were explored. Results: Surprise (25.32%) followed by Disgust (19.88%) emerged as dominant sentiments. However, sentiment correlations with the actual spread of COVID-19 were not significant, hinting at a possible decline in COVID-19 interest over time [3].

The manner in which Indonesia tackled the COVID-19 crisis generated significant buzz on Twitter, leading to a polarized public opinion. Analyzing these tweets can provide valuable insights for policy formulation and assessment of government action. Sentiment analysis offers a way to discern the public's sentiment based on their tweets. The goal is to gauge the prevailing sentiment about Indonesia's handling of the pandemic from both broad and economic viewpoints. We gathered tweets using the Twitter scraper library. These tweets, which were initially without labels, were then classified into positive, negative, and neutral sentiments using SentiStrength and expert input. Post this, data pre-processing was undertaken to discard redundant and non-relevant tweets. Machine learning models were then implemented to determine sentiment in fresh data, and their accuracy was validated through confusion matrices and K-fold cross-validation. An SVM analysis yielded an impressive performance with average accuracy, precision, recall, and f-measure being 82.00%, 82.24%, 82.01%, and 81.84%, respectively. While economic policies regarding the pandemic seemed to resonate well with the populace, there was general discontent regarding the government's overall performance. Specifically, the SVM algorithm's Normalised Poly Kernel, offers a robust tool for rapid and precise sentiment prediction on Twitter [4].

The COVID-19 pandemic brought both health and economic challenges. By utilizing big data, logistics firms could identify profitable strategies to navigate these trying times. This study utilized text-mining techniques on consultancy websites within the logistics domain. Key objectives included identifying frequently used terms, sentiment analysis with the NRC lexicon, highlighting common word pairs, and extracting strategies for cost-efficient shipping and inventory management during the pandemic. Notably, the words "supply," "chain," and "COVID-19" were commonly found. Emotions related to trust frequently appeared, indicating a desire for reliable solutions during the crisis [5].

Social media platforms, teeming with information, often disseminate misleading information, particularly around the COVID-19 pandemic. This paper delves into sentiment analysis of Malaysian COVID-19 news, especially on platforms like Twitter. Due to the multilingual nature of Malaysian tweets, we constructed a diverse dataset, primarily consisting of Malay, English, and Chinese languages. By employing the BPE-Text-to Image-CNN and BPE-M-BERT models, we concluded a mixture of' smoothly combines into the M-BERT network that has already been trained was better suited for our multilingual dataset [6].

The advent of platforms like Twitter has enabled real-time monitoring of global events and public sentiments, as witnessed during various health crises. Notably, the COVID-19 and Mpox outbreaks catalyzed information sharing on Twitter. This study, bridging a research gap, delves into a simultaneous analysis of tweets related to both diseases. Key findings include predominant negative sentiments and frequent mentions of topics like President Biden and Ukraine, providing a holistic understanding of public discourse during this period [7].

Efforts were put in place to mitigate the spread of COVID-19, yet the community's insufficient awareness and lack of discipline rendered these strategies less effective. This lapse amplified the risk of ongoing viral exposure. Beyond health impacts, the pandemic deeply influenced various sectors including social, political, religious, economic, and the resilience of the populace. Insights into these effects can be gleaned from social media, notably with regards to socio-economic ramifications. This study utilized Twitter data via its API to capture the Indonesian sentiment about the pandemic. We employed a sentiment analysis technique, a combination of Naive Bayes classification with TF-IDF and Lexical feature extraction techniques. Tweets containing "COVID-19" in Indonesian were labeled under five sentiment categories: fear, anger, love, sadness, and happiness. Results highlighted that our The incorporation of the recently developed TFBS strategy showed an astounding accuracy of 0.85, exceeding predictions and performing better than traditional techniques. Other techniques. Performance metrics like precision, recall, and F-score further emphasized its superiority [8].

In this study, sentiment analysis was conducted on Facebook posts about the Greek National Public Health Organization (EODY) during the pandemic peak using Microsoft Azure Machine Learning Studio. After examining 300 reviews, sentiments were categorized as positive, negative, or neutral. The analysis centered on sentiments regarding daily COVID-19 surveillance reports on EODY's Facebook page. Machine learning was deployed to forecast sentiment classifications. Results indicated a negative sentiment towards these reports, with recurring

terms like government, vaccinations, and COVID-19. Two classifiers, the Neural Network and Bayes Point Machine, stood out with high accuracy and F1 scores. This research underlines There is great potential in developing machine learning skills in gauging public sentiment, aiding in public health decision-making during crises like the pandemic [9].

As the pandemic unfolded, the critical nature of palliative care became more apparent. This study analyzed over 26,000 English tweets from 2020-2022 to gauge public sentiment on palliative care during the pandemic. Through web scraping, we identified four recurring themes. Interestingly, while many tweets highlighted the pandemic's negative impact on palliative care, an almost equal number pointed to its positive effects. Machine learning helped classify a vast number of these tweets, particularly focusing on the negative implications of COVID-19 [10].

With social media platforms like Twitter offering a wealth of public opinion, harnessing this data can offer fresh insights, especially concerning vaccine hesitancy. This research analyzed tweets related to the topic, preprocessed, and categorized them based on various sentiments and emotions using the NRC-Lexicon technique. Statistical tests were performed to validate emotional correlations. Various neural networks were then trained for sentiment multi-classification, with the BERT model showcasing a stellar 96.71% accuracy [11].

This study taps into Twitter to grasp the preferences for transportation modes during the pandemic era. We gathered tweets about various travel methods in NYC spanning January 2020 to January 2022. Utilizing this we created trip mode classifiers utilizing gathered data and state-of-the-art NLP approaches. to categorize tweets into specific transport modes. Sentiment analysis revealed public sentiment shifts regarding transportation options during the pandemic. Interestingly, buses, bikes, and personal vehicles were viewed positively, aligning with the noticeable shift of commuters preferring them over subways. Concerns arose from tweets highlighting non-compliance with mask-wearing in subways and buses. Delving deeper, a regression analysis based on user demographics shed light on factors influencing public views on public transit, particularly noting the service industry's sensitivity to MTA subway performance [12].

The global upheaval caused by COVID-19 profoundly affected South Africa, spurring heightened levels of anxiety and prompting many to seek solace in religious practices. This investigation delves into social media metrics, exploring the South African populace's sentiments about religion and well-being during the pandemic. By analyzing Covid-19, religion, and the entwined stories of life reverberate in a clamour of tweets on the internet meaning, and experiences, we gauged sentiment trends. Our findings illustrate how religious and COVID-19 sentiments both influence life experiences. Additionally, we introduce a novel "Threshold of Depression" metric in sentiment analysis, providing insights into communal emotional states during crises [13].

Amidst the COVID-19 pandemic, Twitter became a vibrant platform for global discourse. Various studies have examined specific themes on this platform, yet a comprehensive review on the potential applications of sentiment analysis of pandemic-related tweets remains unexplored. This research illuminates how sentiment analysis, blended with behavioral and social science perspectives, can offer critical pandemic management insights. Through an exhaustive literature survey of machine learning techniques applied in sentiment analysis of pandemic tweets, we highlight the supremacy of ensemble models, particularly BERT and RoBERTa, when tailored for Twitter data [14].

In the combat against the pandemic, vaccines play a pivotal role. Navigating the intricate public perceptions of COVID-19 vaccines, this study melds social media analytics with health surveillance data. Harnessing deep learning, we assessed sentiments from millions of tweets between 2020 and 2022. Our findings chronicle the sentiment trends and offer an in-depth look into the pregnant demographic's unique sentiment trajectory. By mapping sentiments with global vaccination patterns, our analysis provides actionable insights for enhancing vaccine outreach [15].

The digital waves of the COVID-19 pandemic brought an influx of both informative and misleading content on social media. This study delves into public opinions on COVID-19 vaccination shared on Twitter to aid policy formulation for enhanced vaccine acceptance. We utilized a variety of analytical methods, with results highlighting the Extra Tree Classifier (ETC) using the Bag of Words (BoW) as the most effective sentiment analysis tool for COVID-19 tweets. Over time, our findings underscore a growing positive sentiment towards vaccination [16].

Social networks like Twitter mirrored the collective mood swings induced by the COVID-19 pandemic. This research offers a deep dive into the sentiments of Mexicans during one of the pandemic's most tumultuous phases. Using a hybrid semi-supervised approach, we trained various models, including two exclusively in Spanish, to gauge sentiments. Comparison with other classifiers like SVM and Decision Trees validated the superior accuracy of the Spanish-centric model. This tailored model was then deployed to gauge Mexico's Twitter sentiment regarding COVID-19 [17].

The research delved into the impact of public health agencies (PHAs) on TikTok are consistently posting videos with daily updates on new cases as COVID-19 becomes more common place. The objective was to decipher public sentiment and concerns during the 2022 Shanghai lockdown. The study incorporated the Crisis and Emergency Risk Communication framework, which delineated the phases of the lockdown into five distinct stages [18]. The research methodology involved categorizing the Shanghai lockdown into various stages. User comments on the most recent case updates shares during these stages were meticulously examined. Utilizing the pre-training ERNIE model, sentiments of the user comments were classified. Subsequently, Semantic network analyses were built on top of the sentiment categorization outcomes [18]. Key observations included: Due to the significant costs associated with combating the epidemic, the general populace was hesitant to comply with preventive measures in the early stages. Shanghai's unilateral decision to redefine "asymptomatic patients" caused ripple effects, influencing control measures in neighboring regions and impacting daily life. Throughout the maintenance phase, individuals highlighted specific aspects of their lives affected by the epidemic. By the time the resolution stage was reached, there was a noticeable waning interest in daily case update videos. A notable source of public discontent was the perceived deviation of local government policies from the central government's directives [18].

The global health scenario is currently grappling with the mammoth challenge of COVID-19. To counteract this, vaccine development and distribution are paramount. With an increase in vaccine dissemination, a concomitant rise in immunity against the virus can be expected. In this digital age, platforms like Twitter serve as invaluable tools for gauging public sentiment, especially concerning vaccination drives. Leveraging advanced artificial intelligence and geo-spatial methodologies, this study undertook the task of analyzing COVID vaccine-related tweets. The sentiment polarity of these tweets was ascertained using the Text-Blob function. This data was then visually represented using word clouds and sentiment classification via the BERT model. Geo-coding was employed to pinpoint and represent sentiment data on a global map. Advanced analytical techniques, like hotspot analysis and kernel density estimation, illuminated regions with predominant positive, negative, or neutral sentiments. Benchmarking against established techniques revealed The model's assessment includes measures for recall, accuracy, and F-scores for both positive and negative sentiment categories classifications to employ such sentiment and spatial analytics during global health crises, thereby understanding public perception towards vaccination drives [19].

Since March 2020, the world has witnessed the exponential spread of the SARS-CoV-2 virus. In its wake, millions have been infected, prompting individuals worldwide to voice their opinions on platforms like Twitter. This research focuses on gauging the sentiments of Hindi-speaking Twitter users regarding the pandemic. Initial data processing involved Natural Language Processing (NLP) on COVID-19 related Hindi tweets. Subsequently, an innovative Grey wolf optimization technique facilitated optimal feature selection. The primary analysis tool was a hybrid model, combining the strengths of Convolution Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. Comparative analysis with other machine learning paradigms showcased the superior performance of the proposed model across multiple metrics, including accuracy, precision, recall, and F-score [20].

The financial ramifications of the COVID-19 outbreak pandemic have been profound, with stock markets worldwide feeling the brunt of the crisis. Given the unpredictable nature of stocks, especially under pandemic conditions, accurate forecasting has been a focus area for investors. This study introduces an innovative system, leveraging applying sentiment analysis to COVID-19 news to forecast outcomes stock market movements. By analyzing stock news headlines from the COVID-19 era, daily stock trends were forecasted. Machine learning classifiers were further employed to deduce the pandemic's overarching impact on high-value stocks, including TSLA, AMZ, and GOOG. To enhance prediction accuracy, features were refined, and extraneous data, like spam tweets, were filtered out. The system's uniqueness lies in its dual approach: textual analysis of social media content and examining previous stock data by applying data mining methods. The system showcased commendable prediction accuracies for a forementioned stocks [21].

Since the start of the COVID-19 outbreak, global citizens have taken to social media to voice their thoughts, feelings, and perceptions regarding the unfolding crisis. Platforms like Twitter serve as vast reservoirs of freely shared data, offering users the opportunity to discuss the pandemic unrestricted by time or location. Alarmingly, the swift rise in cases worldwide has stirred emotions of panic, dread, and unease among many. This study introduces a novel sentiment analysis method to discern feelings During the months of March through October 2020, COVID-19 topics dominated Moroccan Twitter discourse. Utilizing a recommendation-based approach, our model categorizes each tweet into positive, negative, or neutral sentiments. Tests indicate our method's commendable accuracy of 86%, surpassing established machine learning techniques. We observed fluctuating sentiments over time, implying the country's evolving COVID-19 scenario significantly influenced public emotions [22].

This research delves into public sentiments surrounding Using pertinent hashtags and phrases, we analyse Twitter data from January 2021 to March 2023 to gather insights regarding India's COVID-19 vaccination journey acquired tweets underwent preprocessing and cleanup prior to sentiment assessment via Natural Language Processing. Findings suggest an overarching positive sentiment about the vaccine, with numerous tweets advocating for it. Nevertheless, some reservations were noted concerning vaccine reluctance, potential side effects, and skepticism towards government and pharmaceutical entities. Diving deeper, sentiments were further dissected based on gender, age, and geographical markers, revealing varied emotions across these demographics. Our insights shed light on India's vaccination landscape and underscore the necessity for tailor-made communication strategies to mitigate vaccine skepticism and bolster uptake among specific groups [23].

The COVID-19 statistics from the U.S. Center for Disease Control and Prevention up to June 2020 alarmingly show African Americans disproportionately affected by the virus, both in infection and fatality rates. This disparity underscores the imperative to understand the experiences and sentiments of the African American community regarding COVID-19. With Twitter proving valuable for behavioral insights and sentiment extraction, our study harnesses 2020 Twitter data to delve into the pandemic narratives of African Americans using aspect-based sentiment analysis. We crafted a machine learning framework that filters tweets unrelated to COVID-19 or unlikely from African American users, culminating in the analysis of roughly 4 million tweets. Predominantly, the tweets exuded a negative sentiment, with spikes in tweet volume correlating with major pandemic-related events in the U.S. Our study reveals the evolution of pandemic terminology over the year and underscores pressing concerns like food scarcity and vaccine hesitancy. By highlighting specific word associations and sentiments, we aim to enhance understanding of how the pandemic shaped the discourse among African American Twitter users [24].

For societal well-being, it's pivotal that the media spotlight health-related matters, amplifying public awareness and thereby mitigating health threats. Lately, deep neural networks have gained traction in textual sentiment analysis, paving the way for real-time surveillance and insights on health topics. In our study, we introduce Cov-Attribute BI-LSTM, a unique model designed for sentiment analysis on COVID-19 news headlines, leveraging deep neural networks. By amalgamating attention mechanisms, embedding methodologies, and semantic data labeling, we aim to bolster prediction accuracy. Benchmarking against other machine learning classifiers, our model exhibited a superior performance with a 0.931 test accuracy. Further, we employed it on 73,138 pandemic-centric tweets from six global outlets, offering an accurate portrayal of global COVID-19 news and vaccination narratives [25].

Fundamentally, information propels progress. Over the centuries, the modalities for data acquisition have evolved, chiefly due to technological advancements. While some technological mediums offer speed, reliability, and efficiency, others fall short for various reasons. This research seeks to employ Text-Blob and VADER analyzer on historical tweets, gauging public sentiment during the COVID-19 crisis in Nigeria. Beyond gauging emotional responses, it casts light on the societal, ecological, and economic implications in Nigeria. For scholars venturing into data science, machine learning, and deep learning, this research could be a cornerstone. After sourcing 1,048,575 tweets using the 'COVID-19' hashtag, they underwent preprocessing and were subsequently analyzed using Text-Blob and VADER for sentiment. Our findings highlighted varied sentiments, underscoring the power of social media insights in assisting global entities in formulating strategies against COVID-19's repercussions and misinformation [26].

In today's digital age, platforms like Twitter have become outlets for global citizens to air their sentiments, more so during crises like the ongoing pandemic. This paper dives into the sentiments of Indians regarding COVID-19 and the ensuing vaccination campaigns, using tweets as the primary data source. Leveraging both deep learning and lexicon-centered methods, we classified tweet sentiments. The lexicon approach used tools like VADER and NRC-

Lex, while the deep learning utilized Bi-LSTM and GRU, yielding impressive accuracies. Our models stand poised to aid healthcare professionals and decision-makers in future pandemic scenarios [27].

In the COVID-19 era, even with a substantial dip in public transit usage, around 20% of riders persisted during social distancing. Conventional data collection on urban transport scarcely captures the nuanced experiences and psychological facets of users. Thus, understanding the plight of transit-dependent vulnerable populations remains elusive. Through machine learning, we enhanced traditional segmenting and collecting data with great care to create a customised Twitter dataset that mirrored the journeys of roughly 120,000 Metro Vancouver transit riders pre and amidst the pandemic. Our analysis revealed a pronounced spike in adverse sentiments, particularly across varying demographics, during the initial COVID-19 waves. By unearthing these disparities and risks for transit users during the crisis, we hope to inform better public health measures and transportation strategies during future disruptions [28].

This study introduces an approach that employs Understanding Linguistics (CLU) and Emotional Tone Evaluation (ETA) to discern public sentiment and perspectives on COVID-19 vaccination within Italy. The research zeroes in on tweets related to vaccines that emerged January 2021–February 2022 in Italy. Out of the initial 1,602,940 tweets with the term “vaccine,” 353,217 tweets were selected for in-depth analysis. A unique aspect of our methodology is the division of users into four groups: The public, the media, the medical field, and the sphere of government. This categorization leverages large domain-specific vocabulary to improve NLP tools, which review user profiles. Our sentiment analysis incorporates an Italian lexicon packed with polarized and intensive words, offering clues about each user category's tone. The study's findings underscore a predominantly negative sentiment across the reviewed duration, most pronounced among Common users. Notably, specific events, like post-vaccination fatalities, influenced the sentiment patterns across the 14 months [29].

In the arena of natural language processing, sentiment analysis stands out, especially in extracting insights from web data regarding COVID-19 to aid China's pandemic response efforts. While deep learning-backed sentiment analysis models have made strides, they often grapple with dataset constraints. To address this, we introduce a model grounded in a federated learning framework, incorporating Bert and a multi-scale convolutional neural network Fed-BERT-MSCNN. This model blends Bidirectional Encoder Representations from Transformers with multi-scale convolution layers. The federated framework integrates a primary server with localized deep learning systems for individual dataset training. Parameter exchanges are managed via edge networks, streamlining communication and model integration. This innovative network not only addresses data scarcity but also fortifies data privacy during the training phase and bolsters communication efficacy. Upon testing across six social platforms, the Fed-BERT-MSCNN model's performance consistently outshined its peers [30].

The COVID-19 outbreak presented a global event like no other, notably influencing financial markets, especially in its nascent stages. This research delves into the question of whether COVID-19 news flows influenced market sentiments. We scrutinized on The effects of COVID-19 were felt on three major websites between January and June 2020: MarketWatch.com, NYTimes.com, and Reuters.com lleveraging machine learning, we gauged the sentiment of these articles using a financially-tuned BERT model, adept at contextual word understanding. Our analysis points to a robust positive correlation between sentiment scores and the S&P 500 market's behavior. Intriguingly, the impact of sentiment facets and news genres from NYTimes.com varied in relation to market returns [31].

Table 1. Systematic table for review

Serial Number	Method	Result	Advantage	Disadvantage	Future Scope
1	Categorization of sentiment based on social media	Improved sentiment classification Accuracy: - 97.5%	Utilizes deep learning	Limited to sentiment analysis	Exploring ensemble models for enhanced accuracy
2	Generating insights using context-specific	Identifying hesitancy trends	Provides insights into public	Limited to Twitter data	Investigating factors contributing to

	machine learning, sentiment analysis, and text mining for vaccine	Accuracy: - 96.7 %	perception		hesitancy
3	Sentiment analysis of Twitter data based on COVID-19 Cases	Analyzing public sentiment related to COVID-19 in Greece.	Utilizes real-time social media data	Limited to a specific geographical region	Comparing sentiment across regions
4	Applying a normalized polynomial kernel to Support Vector Machines (SVM) for sentiment analysis based on Indonesian Gov Handling on covid-19	Assessing the general public's perception of the administration Accuracy: - 82.4%	Addresses Indonesian context	Limited to Twitter data	Extending to other social media platforms
5	Sentiment analysis and text mining based on supply chain strategy	Analyzing sentiment and knowledge discovery in supply chain strategy Accuracy: - Anger manifests at 4.99% (160 occurrences), whereas eagerness peaks at 9.88% (317 occurrences). Disgust is recorded at 2.52% (81 cases), fear peaks at 7.51% (241 cases), happiness is recorded at 5.61% (180 cases), melancholy is recorded at 6.26% (201 cases), surprise	Addresses business context	Limited to supply chain data	Generalizing to other industries

		is recorded at 3.46% (111 cases), and trust is recorded at a noteworthy 16.20% (520 cases). emotion that is negative is 13.62% (437 cases), whereas emotion that is favorable is 29.95% (961 cases).			
6	Sentiment analysis of a mixed-language Twitter dataset based on Malay-English twitter dataset	Analyzing sentiment in a multilingual context Accuracy: - The CNN-BERT-two-class model had a greater accuracy of 67%, whereas the CNN-BERT-three-class model's accuracy was 58%.	Diverse language representation	Limited to Twitter data	Expanding to more languages
7	Sentiment and using text analysis to conduct a thorough investigation of public debate on Twitter provides fascinating insights into existing opinions and patterns.	Analyzing public discourse on COVID-19 and Mpox Accuracy: -The results show that over half (46.88%) of the tweets on COVID-19 and Mpox were unfavourable in tone. Following this, tweets with an optimistic emotion (31.97%) lagged, while tweets with a neutral posture	Addresses multiple topics	Limited to Twitter data	Exploring sentiment shifts over time

		(21.14%) did the same.			
8	Sentiment analysis with hybrid feature extraction based on public reaction identification and Monitoring hybrid feature	Identifying public reactions on COVID-19 monitoring Accuracy: - 85.4%	Hybrid feature extraction	Data preprocessing challenges	Integrating additional data sources
9	Sentiment analysis using machine learning based on epidemiological surveillance reports	Analyzing sentiment in epidemiological surveillance reports Accuracy: - 87%	Utilizes public health data	Limited to Greece	Comparing sentiment across countries
10	Sentiment analysis of palliative care opinions.	Analyzing sentiment in the context of palliative care during COVID-19 Accuracy: - 97% of people are aware of the negative effects of the pandemic.	Addresses healthcare context	Limited to palliative care	Investigating patient experiences
11	Framework combining deep learning, sentiment classification, and emotional reactions analysis based on Global Public Health and Risk Modelling Framework	Identifying vaccine hesitancy and risk factors Accuracy: - 89.9%	Comprehensive framework	Data integration challenges	Evaluating interventions for hesitancy
12	Sentiment analysis on multimodal transportation based on Multimodal Transportation	Analyzing public sentiment related to transportation during COVID-19 Accuracy: - 94%	Multimodal data analysis	Limited to transportation context	Exploring sentiment impact on transportation decisions
13	Accurately determining the threshold of depression is	Managing risk during crises like COVID-19 Accuracy:-95%	Addresses mental health context	Limited to Twitter data	Investigating mental health interventions

	necessary for sentiment analysis. Based on Threshold of Depression measure				
14	Utilizing Twitter data, advanced machine learning methods are employed for sentiment analysis.	Applying ML techniques on Twitter data	Wide range of ML methods	Limited to Twitter data	Evaluating ML model performance
15	Combining surveillance data with Twitter data to provide a thorough sentiment analysis. based on Pattern of Diverse Changing Sentiments	Analyzing global sentiment toward COVID-19 vaccines Accuracy: - 92%	Utilizes diverse data sources	Data volume and processing challenges	Investigating sentiment evolution over time
16	Diverse feature engineering techniques for sentiment analysis based on Temporal analysis and opinion dynamics	Examining how public opinion has changed over time about the COVID-19 immunization. tweets Accuracy:- 92%	Utilizes diverse feature sets	May require extensive data preprocessing	Investigating vaccine-related misinformation
17	We have developed a sentiment analysis model based on deep learning concepts using deep learning techniques.	Deep learning techniques for sentiment analysis on Twitter. Accuracy:- 97%	Advanced deep learning techniques	Computationall y intensive	Applying deep learning to other social media platforms
18	Investigating how daily COVID-19 instances affect public attitude using deep learning methods for sentiment categorization and semantic	Investigating the impact of variations in COVID-19 infection rates on a daily basis have a big impact on public opinion.	Integrates deep learning and semantic analysis	Limited to sentiment analysis	Exploring broader societal impacts

	network analysis.				
19	Spatial analysis of vaccine-related tweets based on Spatial Analysis	Analyzing sentiment and spatial patterns in COVID-19 vaccine tweets	Combines sentiment and spatial analysis	Limited to Twitter data	Investigating spatial correlations
20	Using the advantages of an ensemble hybrid model designed especially for analyzing Hindi COVID-19 text, a unique ensemble hybrid technique is built for text categorization.	Improving text categorization techniques for COVID-19 content in Hindi. Accuracy:- 95.54%	Utilizes ensemble and optimization techniques	Focused on Hindi language	Adapting the model for other languages
21	Hybrid creating a new stock trend forecasting model based on the effective combination of hybrid stock research techniques. prediction system.	Effective stock trend forecasting during COVID-19 with emotion and deep learning analysis Accuracy:- 90%	Combining sentiment analysis of news articles with stacked-LSTM.	Limited to stock prediction	Enhancing stock prediction during crises
22	Sentiment analysis method using recommender approach based on detect and analysis sentiment COVID-19 Moroccan Tweets	Analyzing sentiments of Moroccan COVID-19 tweets Accuracy:- 86%	Innovative recommender-based sentiment analysis	Focused on Moroccan tweets	Applying recommender approach to other contexts
23	Sentiment study of views towards the COVID-19 vaccine based on perception and result society	Analyzing vaccination sentiment in India on Twitter Accuracy:- 88.6%	Addresses vaccination perceptions	Limited to India	Investigating vaccination outcomes
24	Using aspect-based sentiment analysis to categorize sentiments in Twitter data according to	Exploring The COVID-19 epidemic presented African Americans with unique issues	Focuses on specific demographic experiences	Limited to Twitter data	Examining experiences across different demographics

	various characteristics. based on use aspect based sentiment analysis	that highlighted the intersectional differences in social support networks, economic stability, and healthcare access.			
25	Enhanced sentiment analysis of news reports about COVID-19 that are obtained from global sources.	Analyzing sentiment regarding global COVID-19 news Accuracy:- 93.1%	Focuses on news sentiment	Limited to news sources	Exploring additional news sources
26	Sentiment analysis using VADER and Text Blob based on Nigeria Hashtags	Analyzing COVID-19 tweets in Nigeria using sentiment analysis tools	Utilizes sentiment analysis tools	Limited to selected hashtags	Extending analysis to broader social media content
27	Building on the foundation of lexicon-based approaches, the integration of sentiment analysis techniques leverages both deep learning methodology and lexicon-based approaches.	Using lexicon-based analysis and deep learning algorithms to investigate COVID-19 social media talk.	Combines deep learning and lexicon-based analysis	Addresses different sentiment analysis methods	Improving deep learning model performance
28	Machine learning-based sentiment analysis for transit riders based on vulnerable transit riders	Monitoring well-being during COVID-19	Focuses on vulnerable transit riders	Limited to transit context	Extending monitoring to other vulnerable populations
29	Lexicon-based sentiment analysis of vaccine opinions on Twitter based on detect opinions and attitude for	Detecting opinions and attitudes toward COVID-19 vaccines in Italy	Utilizes lexicon-based approach	Limited to Italian tweets	Adapting for different languages and regions

	COVID-19				
30	With the help of a Federal Learning Edge Network, we have created a novel method for sentiment analysis that is based on our Federal Learning Edge Network's complex mechanics.	Using extensive sentiment analysis to address attitudes about COVID-19 throughout the world.	Utilizes federal learning edge network	Limited to sentiment analysis	Exploring federal learning in other contexts
31	Extracted insights by using machine learning techniques to do sentiment analysis on COVID-19 news and connect it with stock market responses.	Analyzing sentiment, news, and stock market reactions	Focuses on stock market implications	Limited to stock market analysis	Investigating broader economic impacts

3. PROPOSED ALGORITHM

The preprocessing of textual data involves a series of steps to transform the input into a more usable format. Beginning with the original text, each line is processed by first extracting tokens. These tokens are then refined by removing stopping words, which are commonly used words that offer little value in the context of analysis. Following this, the text undergoes stemming, a process to reduce words to their base or root form. Further refinement is done through lemmatization, which involves converting words to their canonical forms. The final output is the preprocessed text, which is then returned for further use or analysis.

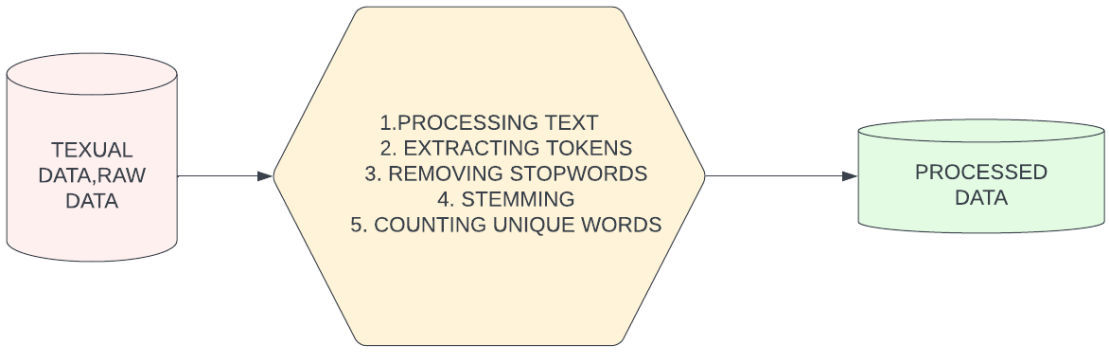

FIGURE 2. Architecture of Pre-Processing (Algorithm 1)

Techniques for tokenization can vary depending on the data format removing them can improve the focus on content-rich words. Stemming reduces words to their base form sing stemming algorithms like Porter Stemmer or Lancaster Stemmer available in NLTK. Alternatively, lemmatization data contains mixed-case characters or special

symbols, you might need to normalize them for better processing, returns the preprocessed text after all the cleaning steps.

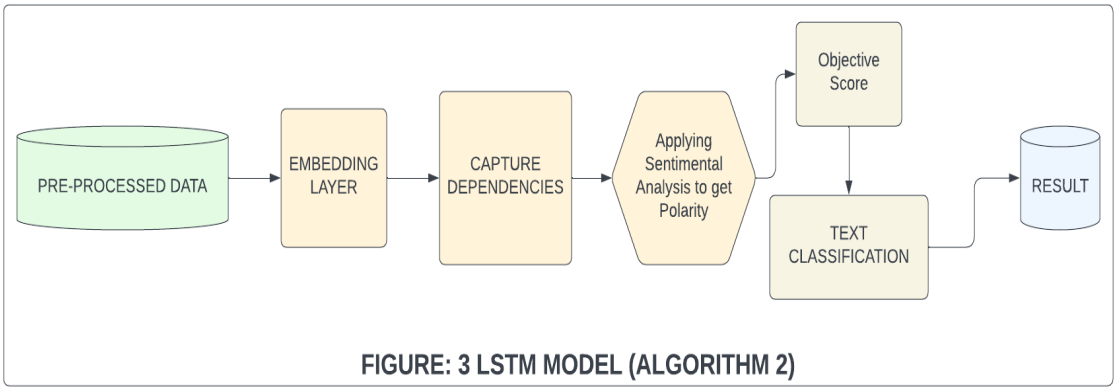

FIGURE: 3 LSTM MODEL (ALGORITHM 2)

The process of classifying preprocessed text into a specific label involves a multi-step approach utilizing advanced natural language processing techniques. To begin with, every token in the preprocessed text—which is made up of a list of feature maps—is embedded using an LSTM layer in order to extract the text's sequence dependencies and contextual information. By using an attention mechanism that concentrates on the most pertinent passages in the text, this is further improved. The processed text is next subjected to sentiment analysis to ascertain its objectivity and polarity, which are then prioritized to highlight important sentiment characteristics. To fully comprehend the underlying event and topic, the text's context is also extracted. Eventually, the text is assigned a particular label by an LSTM classifier that takes into account sentiment and context information. The final product is a class label that symbolizes the text that has been categorized, capturing its tone, background, and essential elements.

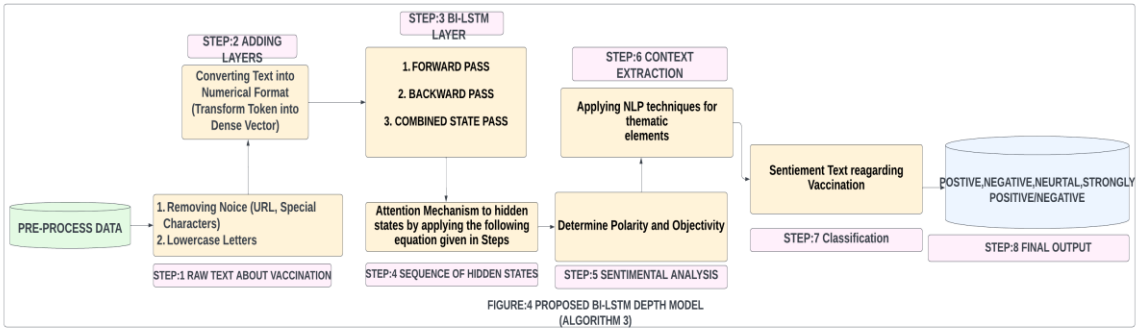

FIGURE:4 PROPOSED BI-LSTM DEPTH MODEL (ALGORITHM 3)

Preprocessing text to eliminate noise and normalize it is the first step in the extensive process of using the modified Bi-LSTM algorithm for sentiment analysis related to vaccinations. After that, it uses embeddings to translate words into numerical vectors, which are then processed using a Bi-LSTM layer to identify contextual dependencies. The text's most instructive passages are then highlighted via an attention-getting method. By calculating polarity and objectivity scores and extracting thematic background, the program evaluates sentiment. Lastly, to identify the text's sentiment on vaccination, it uses a Bi-LSTM classifier that combines sentiment and context characteristics, thereby producing a sentiment label like Positive or Negative. Specifically, this approach leverages attention and sentiment analysis along with the sequential learning capability of Bi-LSTMs to efficiently comprehend and classify vaccination-related opinions.

4. RESULT AND DISCUSSION

The proposed model demonstrates superior performance across various configurations and methods compared to the LSTM [32] baseline. Tweets about the Pfizer/BioNTech vaccine made up the first dataset. I then included tweets on vaccines made in China, namely Sinopharm and Sinovac, as well as vaccines from Moderna, Oxford/Astra-Zeneca, Covaxin, and Sputnik V. In the early days, data gathering took place twice a day until I was able to determine an approximate daily quota for new tweets. After then, the frequency of vaccination collection for

all vaccines stabilized at one shot per day, especially in the AM (GMT). I collected recent tweets from Pfizer/BioNTech, Sinopharm, Sinovac, Moderna, Oxford/AstraZeneca, Covaxin, and Sputnik V on COVID-19 vaccines that are used extensively worldwide.

The original dataset was compiled tweets on the Pfizer/BioNTech vaccine. Following that, tweets on the Sinopharm and Sinovac vaccines (both Chinese-made), Moderna, Oxford/Astra-Zeneca, Covaxin, and Sputnik V vaccines were added. At first, data gathering happened twice a day; once I figured out the approximate number of new tweets, I switched to one a day. In the morning (GMT), this regular vaccination collecting process became standardized.

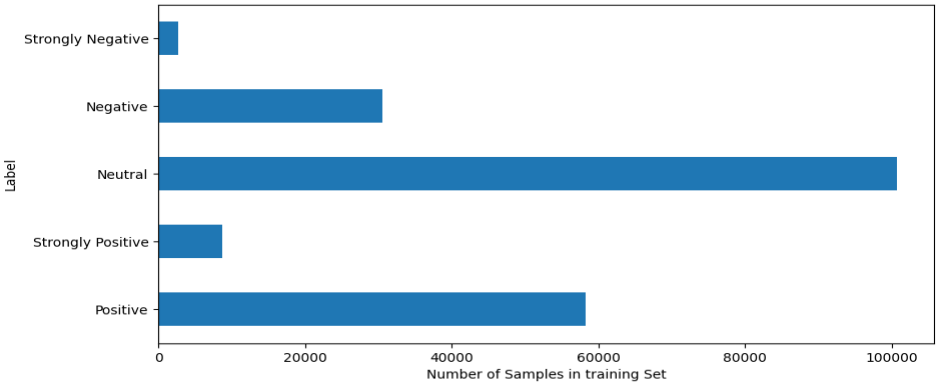

Figure 5. Number of Samples in training Set

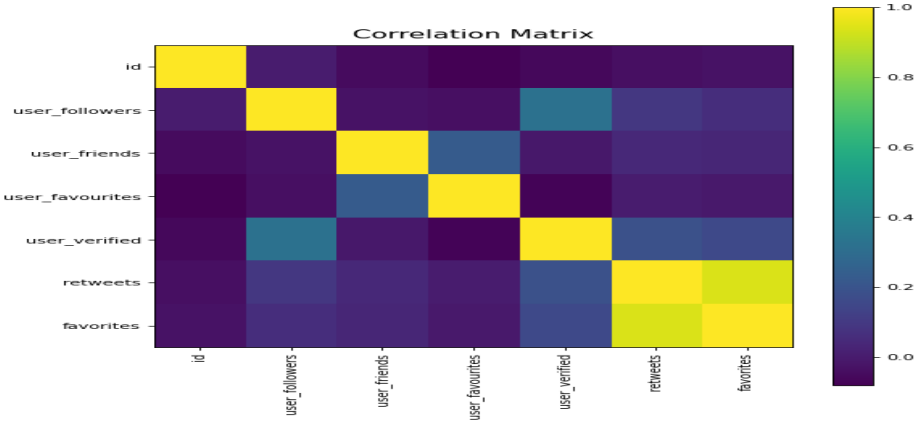

Figure 6. Correlation matrix

- **Strong positive correlation:** User followers and user friends (0.8). This means that users with a high number of followers also tend to have a high number of friends, and vice versa.
- **Moderate positive correlation:** User favourites and user followers (0.6), user favourites and retweets (0.4). This suggests that users who are followed by many people also tend to favourite a lot of tweets, and users who favourite a lot of tweets tend to also retweet a lot.
- **Weak positive correlation:** User verified and user followers (0.2). This means there's a slight tendency for verified users to have more followers than non-verified users.
- **Very weak or no correlation:** User favourites and favourites (0.0). This is surprising, because it suggests that the number of times a user favourites a tweet is not related to how many times their own tweets are favourited by others.

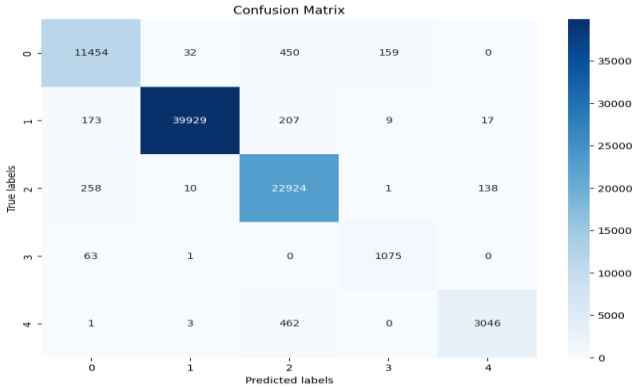

Figure 7: LSTM Confusion Matrix

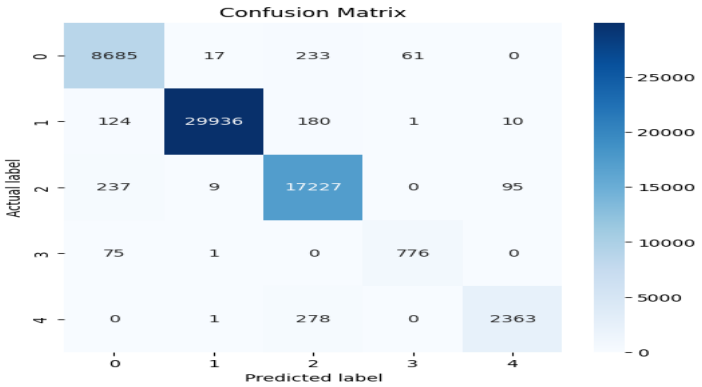

Figure:8 BI-LSTM Confusion Matrix

- **11454 are the Accurate Positives (AP).** These are tweets that the model properly classified as positive.
- **There are 32 Misidentified Positives (MP).** These are tweets that the model incorrectly classified as favourable when they were really negative.
- **173 are the Missed Positives (MPs).** These are tweets that the model incorrectly classified as negative when, in fact, they were favourable.
- **30000 for Accurate Negatives (AN).** These are tweets that the model properly classified as negative.

Overall, based on these numbers, the model seems to perform well at classifying positive sentiment, with a high True Positive (TP) value of 11454 and a low False Positive (FP) value of 32. On the other hand, it seems to struggle more with negative sentiment, with a lower True Negative (TN) value of 30000 and a higher False Negative (FN) value of 173.

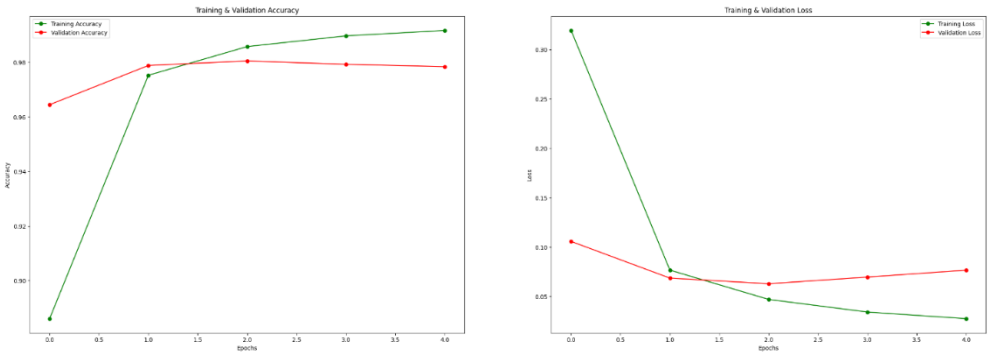

Figure: 9 Training and Validation Accuracy and Training and Validation Loss

- **Training Accuracy:** This concerns the model's performance on the dataset that it was trained on. This measure should ideally be higher, indicating that the model is more competent at understanding the nuances in the training data.

- **Validation Accuracy:** This indicates how well the model performs on a different dataset than the one used for training, which gives a more accurate picture of how well it generalises to new data.

The performance of different Word2Vec architectures, including Continuous Bag of Words (CBOW), Skip-gram, and GloVe, in terms of accuracy when used with Random Forest and Long Short-Term Memory (LSTM) models. The results are based on previous studies as well as the proposed approach. With CBOW, the Random Forest model's accuracy was 72.48%; with Skip-gram, it was 79.48%; and with GloVe, it was 80.75%. However, the LSTM[32] model demonstrated better accuracy; in one research, it was 77.18% with CBOW, 83.2% with Skip-gram, and 89.7% with GloVe; in another, it was 77.21% with CBOW, 83.4% with Skip-gram, and 89.9% with GloVe. But the suggested method beat all earlier findings, obtaining 81.3% accuracy with CBOW, 88.7% accuracy with Skip-gram, and 92.8% accuracy with GloVe. The results indicate that when combined with both Random Forest and LSTM[23] models, the proposed course of action may provide gains in accuracy when compared to current Word2Vec architectures.

Table 2: Word2Vec Architectures			
Accuracy	CBOW	Skip-gram	GloVe
Random Forest [23]	72.48	79.48	80.75
LSTM [23]	77.18	83.2	89.7
LSTM [32]	77.21	83.4	89.9
Proposed	81.3	88.7	92.8

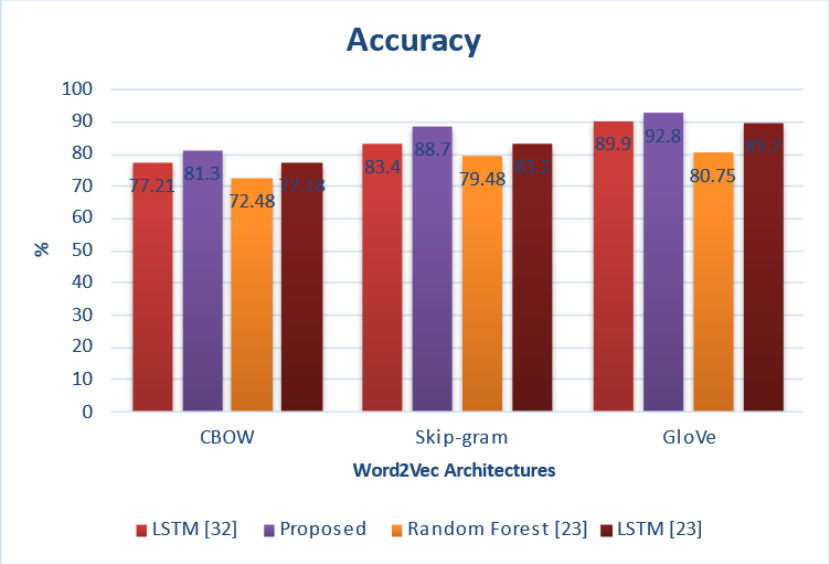

Figure 10. The effect of changing word2vec settings on various architectures' accuracy metrics (a).

Architecture is with 92.8% accuracy rate, the Skip-gram model proves to be effective.. Following closely are GloVe at 89.9% and CBOW at 88.7%. The two LSTM architectures show lower accuracy than the Word2Vec models, with LSTM [32] at 83.4% and LSTM [23] at 79.48%. The least accurate architecture according to the graph is Random Forest [23] at 77.21%.

Table 3: Evaluation Methods		
Accuracy	Hierarchical SoftMax	Negative Sampling
Random Forest [23]	82.76	72.49
LSTM [23]	83.1	78.8
LSTM [32]	83.2	77.4
Proposed	91.25	84.98

The Random Forest model achieved an accuracy of 82.76% with The LSTM model obtained 83.1% accuracy using Hierarchical Softmax and 78.8% accuracy using Negative Sampling in one research, and 83.2% accuracy using Hierarchical Softmax and 77.4% accuracy using Negative Sampling in another.

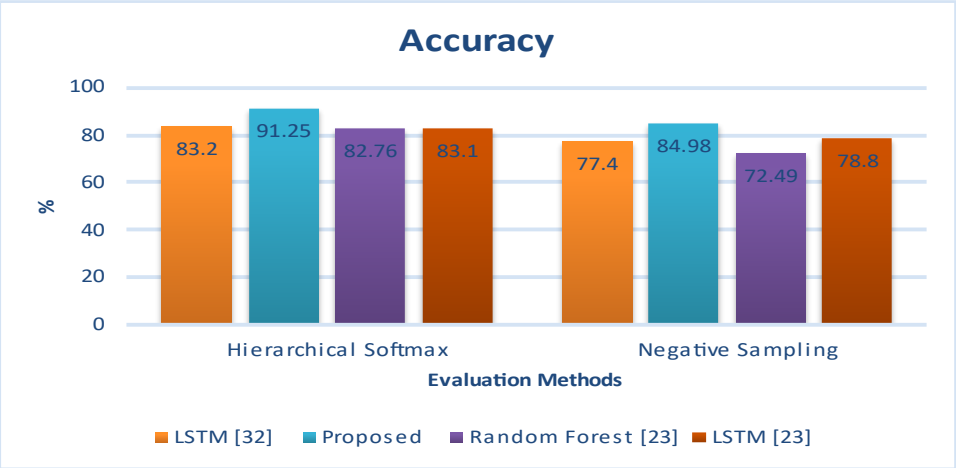

Figure 11. The effect of word2vec settings on assessment methodology' accuracy metric (b).

The accuracy of hierarchical softmax is higher than all the other methods, at around 91.25%. The next most accurate method is the proposed method, at around 84.98%. The accuracy of the other methods is lower: negative sampling at 83.2%, Random Forest [23] at 77.4%, LSTM [23] at 72.49%, and LSTM [32] at 82.76%.

Table 4: Vector Dimension

Accuracy	100	200	300
Random Forest [23]	72.36	76.48	78.42
LSTM [23]	7.4	79.1	80.34
LSTM [32]	79.8	80.4	80.8
Proposed	82.5	84.9	91.7

The Random Forest model achieved accuracy scores of 72.36%, 76.48%, and 78.42% for vector dimensions of 100, 200, and 300 respectively. The LSTM model with reference [23] achieved accuracy scores of 74.0%, 79.1%, and 80.34% for the same vector dimensions. The LSTM model with reference [32] performed slightly better with accuracy scores of 79.8%, 80.4%, and 80.8% for the respective vector dimensions.

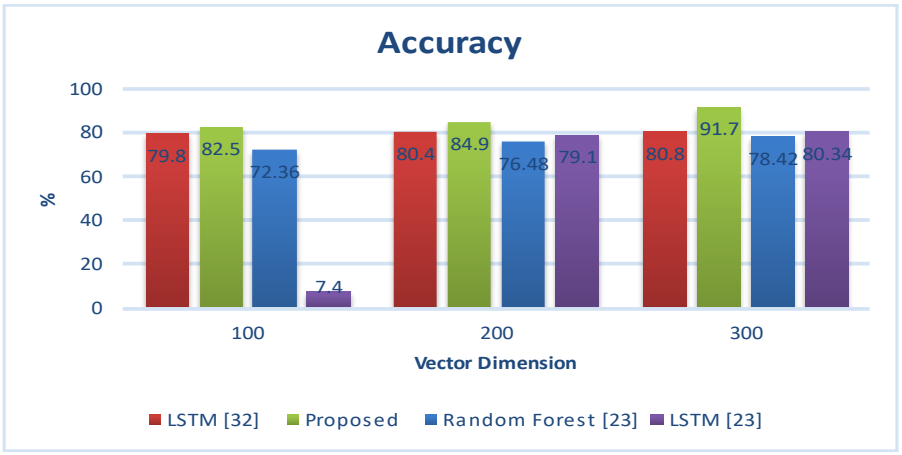

Figure 12. The effect of changing the word2vec parameter with respect to vector dimensionality on the accuracy metric (c).

Table 5: Dropout Values			
Accuracy	0.2	0.5	0.8
Random Forest [23]	78.65	77.36	76.39
LSTM [23]	80.85	80.1	77.69
LSTM [32]	81.6	80.4	78.9
Proposed	93.45	85.21	82.6

The Random Forest model achieved accuracy rates of 78.65%, 77.36%, and 76.39% for dropout values of 0.2, 0.5, and 0.8 respectively. The LSTM model with reference [23] achieved accuracy rates of 80.85%, 80.1%, and 77.69% for the same dropout values. LSTM model however with reference [32] achieved higher accuracy rates of 81.6%, 80.4%, and 78.9% for the respective dropout values.

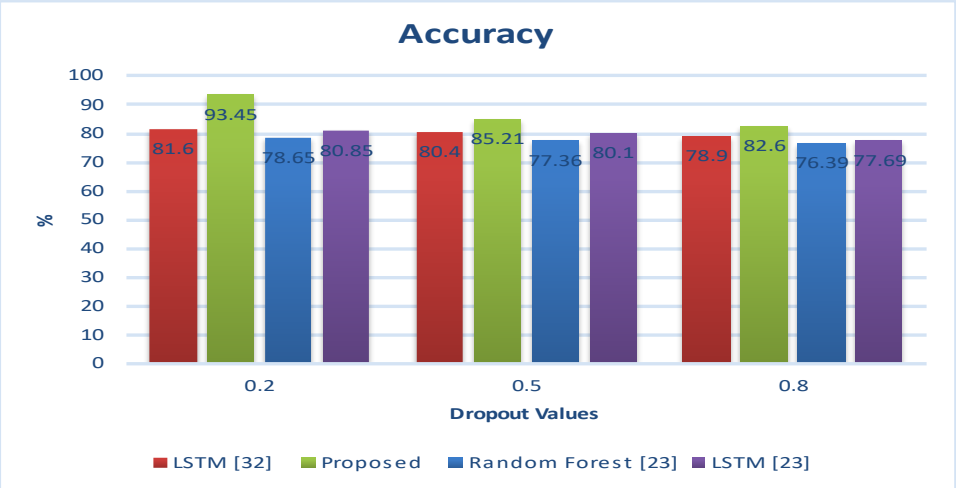

Figure 13. The effect of dropout levels on the accuracy (a) of the suggested Bi-LSTM parameters.

Table 6: **Pooling Technique**		
Accuracy	Max	Avg
Random Forest [23]	78.24	79.36
LSTM [23]	80.14	80.21
LSTM [32]	80.28	80.35
Proposed	81.54	85.32

Random Forest, achieved an accuracy of 78.24% as the maximum value and 79.36% as the average value. The second technique, LSTM (Long Short-Term Memory), reported an accuracy of 80.14% as the maximum value and 80.21% as the average value. Another LSTM technique, referenced as [32], showed slightly higher accuracy with 80.28% as the maximum value and 80.35% as the average value. However, the proposed pooling technique outperformed all others with an accuracy of 81.54% as the maximum value and an impressive 85.32% as the average value.

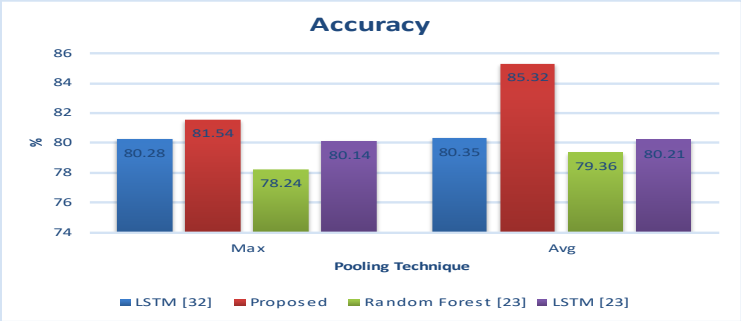

Figure 14. We investigate the effect of the proposed Bi-LSTM parameters on accuracy metrics (b) using different pooling strategies.

Table 7: Learning Rate

Accuracy	0.0001	0.001
Random Forest [23]	72.37	79.38
LSTM [23]	75.29	81.82
LSTM [32]	78.1	82.5
Proposed	86.75	93.54

At a learning rate of 0.0001, the Random Forest model achieved an accuracy of 72.37%, while the LSTM model from source 23 achieved an accuracy of 75.29%. The LSTM model from source 32 performed slightly better with an accuracy of 78.1%.

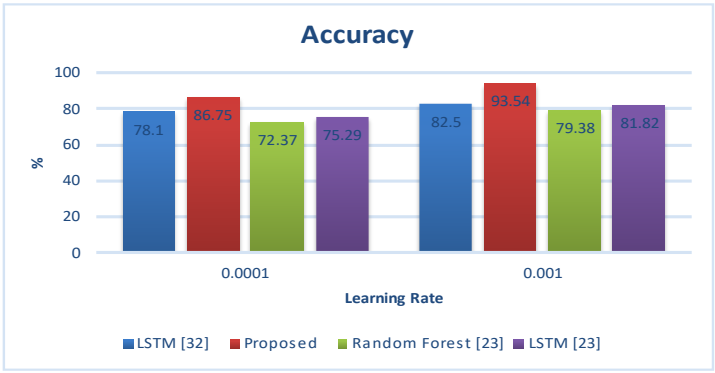

Figure 15. Effect of Proposed Bi-LSTM parameter on accuracy value (c). learning rate values

5. CONCLUSION

This study investigated the effectiveness of various techniques for representing words as vectors, focusing on how these embeddings influence downstream tasks like classification. The three main embedding methods explored were Continuous Bag-of-Words (CBOW), Skip-gram (both from the Word2Vec family), and GloVe. The Random Forest and LSTM models were then applied to these embeddings for classification. The main result is that a recently suggested strategy for word embeddings performs far better in terms of accuracy than current approaches. This innovative method produced better results on all three embedding strategies (CBOW, Skip-gram, and GloVe) when paired with either Random Forest or LSTM models. Notably, the suggested method outperformed all other combinations with an astounding 92.8% accuracy using Skip-gram embeddings. These results imply that the suggested method has the potential to greatly increase the accuracy of word embedding-based tasks, like question answering, machine translation, and sentiment analysis. The preliminary findings are encouraging, but more investigation is required to determine how well this method applies to other datasets and jobs. This can be useful for evaluating any new pandemic attitudes from a future perspective.

REFERENCES

[1] Arbane, M., Benlamri, R., Brik, Y., & Alahmar, A. D. (2023). Social media-based COVID-19 sentiment classification model using Bi-LSTM. *Expert Systems with Applications*, *212*, 118710.

[2] Qorib, M., Oladunni, T., Denis, M., Ososanya, E., & Cotae, P. (2023). Covid-19 vaccine hesitancy: Text mining, sentiment analysis and machine learning on COVID-19 vaccination Twitter dataset. Expert Systems with Applications, 212, 118715.

[3] Samaras, L., García-Barriocanal, E., & Sicilia, M. A. (2023). Sentiment analysis of COVID-19 cases in Greece using Twitter data. Expert Systems with Applications, 120577.

[4] Prastyo, P. H., Sumi, A. S., Dian, A. W., & Permanasari, A. E. (2020). Tweets responding to the Indonesian Government's handling of COVID-19: Sentiment analysis using SVM with normalized poly kernel. J. Inf. Syst. Eng. Bus. Intell, 6(2), 112

[5] Zuhanda, M. K., Anisa, Y., Desniarti, D., Hafiz, M., Syofra, A. H., Caraka, R. E., & Noh, M. (2023). Supply chain strategy during the COVID-19 terms: sentiment analysis and knowledge discovery through text mining. Indonesian Journal of Electrical Engineering and Computer Science, 30(2), 1120.

[6] Kong, J. T., Juwono, F. H., Ngu, I. Y., Nugraha, I. G. D., Maraden, Y., & Wong, W. K. (2023). A Mixed Malay–English Language COVID-19 Twitter Dataset: A Sentiment Analysis. Big Data and Cognitive Computing, 7(2), 61.

[7] Thakur, N. (2023). Sentiment Analysis and Text Analysis of the Public Discourse on Twitter about COVID-19 and MPox. Big Data and Cognitive Computing, 7(2), 116.

[8] Widodo, D. A., Iksan, N., & Sunarko, B. (2023). Sentiment Analysis of Twitter Media for Public Reaction Identification on Covid-19 Monitoring System using Hybrid Feature Extraction Method. International Journal of Intelligent Systems and Applications in Engineering, 11(1), 92-99.

[9] Stefanis, C., Giorgi, E., Kalentzis, K., Tselemponis, A., Nena, E., Tsigalou, C., ... & Bezirtzoglou, E. (2023). Sentiment analysis of epidemiological surveillance reports on COVID-19 in Greece using machine learning models. Frontiers in Public Health, 11.

[10] Inoue, M., Li, M. H., Hashemi, M., Yu, Y., Jonnalagadda, J., Kulkarni, R., ... & Koizumi, N. (2023, March). Opinion and Sentiment Analysis of Palliative Care in the Era of COVID-19. In Healthcare (Vol. 11, No. 6, p. 855). MDPI.

[11] Qorib, M., Oladunni, T., Denis, M., Ososanya, E., & Cotae, P. (2023). COVID-19 Vaccine Hesitancy: A Global Public Health and Risk Modelling Framework Using an Environmental Deep Neural Network, Sentiment Classification with Text Mining and Emotional Reactions from COVID-19 Vaccination Tweets. International Journal of Environmental Research and Public Health, 20(10), 5803.

[12] Chen, X., Wang, Z., & Di, X. (2023). Sentiment Analysis on Multimodal Transportation during the COVID-19 Using Social Media Data. Information, 14(2), 113.

[13] Mba, J. C., & Biyase, M. (2023). Threshold of depression measure in the framework of sentiment analysis of tweets: Managing risk during a crisis period like the covid-19 pandemic. Journal of risk and financial management, 16(2), 115.

[14] Braig, N., Benz, A., Voth, S., Breitenbach, J., & Buettner, R. (2023). Machine Learning Techniques for Sentiment Analysis of COVID-19-Related Twitter Data. *IEEE Access*, *11*, 14778-14803.

[15] Wang, H., Li, Y., Hutch, M. R., Kline, A. S., Otero, S., Mithal, L. B., ... & Luo, Y. (2023). Patterns of diverse and changing sentiments towards COVID-19 vaccines: a sentiment analysis study integrating 11 million tweets and surveillance data across over 180 countries. *Journal of the American Medical Informatics Association*, *30*(5), 923-931.

[17] Ahmed, S., Khan, D. M., Sadiq, S., Umer, M., Shahzad, F., Mahmood, K., ... & Ashraf, I. (2023). Temporal analysis and opinion dynamics of COVID-19 vaccination tweets using diverse feature engineering techniques. *PeerJ Computer Science*, *9*, e1190.

[17] Contreras Hernández, S., Tzili Cruz, M. P., Espínola Sánchez, J. M., & Pérez Tzili, A. (2023). Deep Learning Model for COVID-19 Sentiment Analysis on Twitter. *New Generation Computing*, *41*(2), 189-212.

[18] Che, S., Wang, X., Zhang, S., & Kim, J. H. (2023). Effect of daily new cases of COVID-19 on public sentiment and concern: Deep learning-based sentiment classification and semantic network analysis. *Journal of Public Health*, 1-20.

[19] Umair, A., & Masciari, E. (2023). Sentimental and spatial analysis of covid-19 vaccines tweets. *Journal of Intelligent Information Systems*, *60*(1), 1-21.

[20] Jain, V., & Kashyap, K. L. (2023). Ensemble hybrid model for Hindi COVID-19 text classification with metaheuristic optimization algorithm. *Multimedia Tools and Applications*, *82*(11), 16839-16859.
[21] Sharaf, M., Hemdan, E. E. D., El-Sayed, A., & El-Bahnasawy, N. A. (2023). An efficient hybrid stock trend prediction system during COVID-19 pandemic based on stacked-LSTM and news sentiment analysis. *Multimedia Tools and Applications*, *82*(16), 23945-23977.
[22] Madani, Y., Erritali, M., & Bouikhalene, B. (2023). A new sentiment analysis method to detect and Analyse sentiments of Covid-19 moroccan tweets using a recommender approach. *Multimedia Tools and Applications*, 1-20.
[23] Vishwakarma, A., & Chugh, M. (2023). COVID-19 vaccination perception and outcome: society sentiment analysis on twitter data in India. *Social Network Analysis and Mining*, *13*(1), 1-12.
[24] Chaudhary, M., Kosyluk, K., Thomas, S., & Neal, T. (2023). On the use of aspect-based sentiment analysis of Twitter data to explore the experiences of African Americans during COVID-19. *Scientific Reports*, *13*(1), 10694.
[25] Ahmad, W., Wang, B., Martin, P., Xu, M., & Xu, H. (2023). Enhanced sentiment analysis regarding COVID-19 news from global channels. *Journal of Computational Social Science*, *6*(1), 19-57.
[26] Abiola, O., Abayomi-Alli, A., Tale, O. A., Misra, S., & Abayomi-Alli, O. (2023). Sentiment analysis of COVID-19 tweets from selected hashtags in Nigeria using VADER and Text Blob analyser. *Journal of Electrical Systems and Information Technology*, *10*(1), 1-20.
[27] Ainapure, B. S., Pise, R. N., Reddy, P., Appasani, B., Srinivasulu, A., Khan, M. S., & Bizon, N. (2023). Sentiment Analysis of COVID-19 Tweets Using Deep Learning and Lexicon-Based Approaches. *Sustainability*, *15*(3), 2573.
[28] Tran, M., Draeger, C., Wang, X., & Nikbakht, A. (2023). Monitoring the well-being of vulnerable transit riders using machine learning based sentiment analysis and social media: Lessons from COVID-19. *Environment and Planning B: Urban Analytics and City Science*, *50*(1), 60-75.
[29] Catelli, R., Pelosi, S., Comito, C., Pizzuti, C., & Esposito, M. (2023). Lexicon-based sentiment analysis to detect opinions and attitude towards COVID-19 vaccines on Twitter in Italy. *Computers in Biology and Medicine*, *158*, 106876.
[30] Liang, W., Chen, X., Huang, S., Xiong, G., Yan, K., & Zhou, X. (2023). Federal learning edge network based sentiment analysis combating global COVID-19. *Computer Communications*, *204*, 33-42.
[31] Costola, M., Hinz, O., Nofer, M., & Pelizzon, L. (2023). Machine learning sentiment analysis, COVID-19 news and stock market reactions. *Research in International Business and Finance*, *64*, 101881.
[32] Muhammad, Putra Fissabil, Retno Kusumaningrum, and Adi Wibowo. "Sentiment analysis using Word2vec and long short-term memory (LSTM) for Indonesian hotel reviews." *Procedia Computer Science* 179 (2021): 728-735.

BIOGRAPHIES OF AUTHORS

Alok Soreng Alok Soreng is currently pursuing his MTech at the College of Medi-Caps University, Indore. His research focuses on natural language processing, specifically in sentiment analysis for post-COVID stress patterns. He has demonstrated innovative ideas and was awarded a conference journal at ICTBIG23, a Scopus-indexed paper. His research interests span natural language processing, deep learning, and sentiment analysis. Throughout his coursework, Alok has worked as a teaching assistant, gaining expertise in data analytics and data science. He has hands-on experience with tools such as Excel, Power BI, EDA, and Python. Additionally, he has a strong interest in generative AI. He can be contacted via email at lksoreng133@gmail.com.

Kailash Chandra Bandhu Dr. Kailash Chandra Bandhu is an esteemed academic with over 18 years of experience in Computer Science and Engineering. Holding a Bachelor of Engineering, Master of Technology, and Ph.D. in the field, he specializes in areas such as Database Management Systems, Computer Networking, Wireless Networks, Cloud Computing, Big Data Analytics, and the Internet of Things. He has supervised numerous undergraduate and postgraduate research projects and currently mentors five Ph.D. scholars. Dr. Bandhu has served in prominent leadership roles, including Dean of the School of Computer Science & IT, Dean of Academics, Head of Department, and Nodal Officer for online off-campus counseling. His extensive work also includes coordinating workshops and conferences, as well as contributing to accreditation processes such as NBA, NAAC, and ISO certification. Additionally, he has managed various academic and administrative initiatives, including the RGPV e-governance portal and the MP Government E-scholarship portal. His research interests are centered on wireless networks, big data analytics, and cloud computing. He can be contacted at email: kailashchandra.bandhu@gmail.com (9 pt)