International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Sentiment Analysis in Social Media Using Deep Learning Techniques

Dr. Suwarna Gothane, Dr. G. Vinoda Reddy, Dr. K. Praveen Kumar, Dr. D. Baswaraj, Gumma Parvathi Devi, Sruthi Thanugundala, Ravindra Changala

Submitted: 27/01/2024 **Revised:** 05/03/2024 **Accepted:** 13/03/2024

Abstract: The process of analyzing feelings, views, and emotions expressed in social media content a process commonly referred to as opinion mining or sentiment analysis has grown in importance. Because social media platforms are growing at an exponential rate, a large amount of user-generated data that offers insightful information on trends, consumer behavior, and public opinion is readily available. Sentiment analysis on social networks is hampered by the inherent qualities of the Twitter language as well as the briefness and lack of context of messages on these platforms. In this study, we provide a deep learning model to detect the degree of polarity in Twitter postings using long short term memory and convolutional layers. By up to 18%, our model greatly increased the accuracy of previous methods, with an accuracy of 91%. The abundance of user generated information on social media has made sentiment analysis more crucial than ever. Sentiment analysis and comprehension of text data may now be accomplished with the help of Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks. Using RNN and LSTM models, this study offers a thorough review of sentiment analysis in social media. It covers the design, methods of training, difficulties encountered, and uses of RNNs and LSTMs for sentiment analysis of social media material.

Key words: *Twitter, sentiment analysis, deep learning, social networks, classification.*

I.INTRODUCTION

Natural language processing tasks like textual emotion identification are essential and have the potential to greatly advance a variety of fields, including artificial intelligence and computer-human interaction. Human reactions to circumstances give rise to physiological ideas known as emotions. In order to properly read emotions, it is important to analyze these feelings without modulating voice or facial expression. This calls for a supervisory approach. Notwithstanding these difficulties, it's critical to recognize human emotions as people increasingly use abusive text on social networking platforms like Facebook, Twitter, and the like to express their feelings. Our proposal in this study is to categorize a large number of tweets sentimentally. In order to categorize an expression's feelings as either good or negative, we employ deep learning algorithms. Excitement, fun, happiness, love, neutral, relief, and surprise are among the other categories for good emotions, while anger, boredom, emptiness, hatred, sadness, and concern are among the negative ones. employing three distinct datasets, we conducted experiments and assessed the approach employing long short-term memory and recurrent neural networks to demonstrate how to get high emotion categorization accuracy. Thanks to their capacity to automatically extract features from unprocessed data, deep learning algorithms have become highly effective tools for sentiment analysis. In addition to

1Assistant Professor, Department of CSE, Pimpri Chinchwad College of Engineering, Pune, Maharashtra India.

2Professor of CSE (AI & ML), CMR Technical Campus, Hyderabad, Telangana, India.

3Assistant Professor, Department of IT, KITS, Warangal, Telangana.

4Professor, Department of Computer Science Engineering, Vasavi College of Engineering, Hyderabad, India.

5Assistant Professor, Department Of CSE(AIML), CMR Technical Campus, Hyderabad, India.

6Assistant Professor, Department of CSE, Princeton Institute of Engineering and Technology for Women, Hyderabad, India.

7Associate Professor, Department of CSE, Guru Nanak Institutions Technical Campus, Hyderabad, India.

** Corresponding Author Email: rtit2009@gmail.com*

addressing several approaches, difficulties, and applications in this field, this study offers an overview of sentiment analysis in social media utilizing deep learning techniques. Sentiment and opinion analysis is crucial, as evidenced by the Internet's exponential expansion of information, especially inside social networks.

With its ability to enable millions of users to upload and share content quickly and easily, social networks have emerged as the most exemplary Web 2.0 applications. This has allowed information to stream continuously. The most recent statistics from 2023 indicates that over 91% of Internet users currently use social media monthly, accounting for 76 of the global population, according to various study numbers. As of November 2023, at least 310 million active users were accounted for by 16 social media networks. Facebook has the largest user base, with approximately 2.93 billion monthly active users. YouTube and WhatsApp are next with 2.571 billion and 2 billion users, respectively. With around 441 million active users worldwide, Twitter is a unique platform that blends the elements of social networking, blogging, and instant messaging. It stands out for its huge potential for monetization as well as for business. Indeed, according to Internet users in the 16 to 64 age range, Twitter appears to be the ninth most popular social network [1]. We shall stick with the previous terminology even if this social network has changed its name recently to make this text easier to read and comprehend.

On the social networking site Twitter, users may share messages in the format known as "tweets." People may express their thoughts and feelings on a wide range of topics, disciplines, and themes on this platform. It is a compilation of user opinions and attitudes about a range of subjects, including blogs and conventional online articles. When compared to older blogging and social media sites, Twitter has a greater amount of relevant data. Twitter's response rate is far faster than that of other blogging platforms. Many parties, including marketers and consumers, use sentiment analysis extensively to comprehend market trends and get insights regarding merchandise [1].

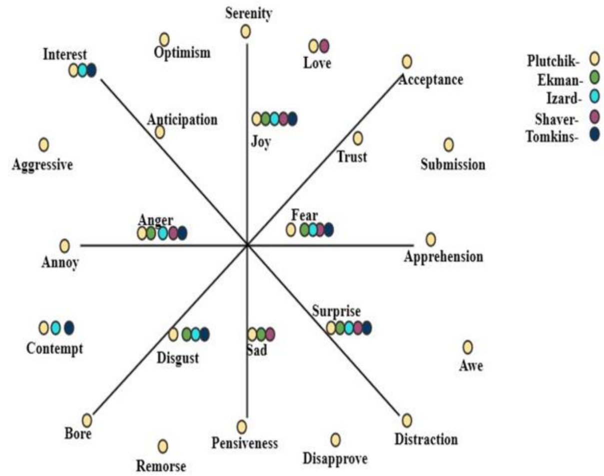

Figure 1. various emotional states

Deep learning has shown excellent results in emotion detection when it comes to sentiment analysis on Twitter, where many techniques are used. Using deep learning models like LSTM (Long Short Term Memory) and RNN (Recurrent Neural Network), this article focuses on identifying human mood in Twitter conversations. The semantic word vectors from each lexical word in the input tweets were extracted using the word2vec technique. To train and predict whether the tweet's emotion categorization labels were positive or negative, the collected features were combined and passed via an LSTM and RNN. Based on the data trial, the LSTM approach offers superior accuracy in sentiment classification prediction.

II. BACKGROUND WORKS

The process of extracting data and identifying patterns from semi-structured or unstructured text is known as text mining, and it is used to extract knowledge from texts. One of the most effective machine learning techniques is deep learning, which uses textual features to predict opinions from a variety of texts based on the recognised features. The opinion that is gathered to determine the opinions and feelings of the general public is utilised as a basis for making different business decisions. The sentiment analysis and deep learning techniques are summarised in this article.

Applying techniques for text-based information and natural language processing to identify and extract subjective knowledge of the text is known as sentiment analytics. While word embedding techniques and semantic texts have garnered greater attention, natural language processing and text classifications can handle small amounts of corpus data. Sophisticated prediction outcomes may be obtained by utilising deep learning, a potent technique that learns several informational layers or attributes. Deep learning techniques at the phrase, document, and aspect levels are applied in various sentiment analytics applications. The primary issues that greatly impact sentiment score, pooling, and polarity identification during the sentiment assessment stage are the focus of this review study. Recurrent and convolution neural networks are the two most often used deep learning techniques. Finally, utilising deep learning models, a comparison analysis is conducted with an extensive literature scan.

Technique Used and Results	Merits	Limitations
Adjective–noun pairs (ANP) and convolutional neural networks (CNN) trained based on Caffe. The proposed DeepSentibank outperformed SentiBank 1.1 by 62.3%.	The suggested model enhances annotation accuracy as well as ANP retrieval performance.	To reduce overfitting, network structure must be adjusted.
Fine-tuned CaffeNet CNN architecture was employed. Obtained an accuracy of 0.83 on Twitter dataset for sentiment prediction.	The model uses fewer training parameters and provides a robust model for sentiment visualization.	To accommodate the presence of noisy labels, the architecture must be rebuilt.
Low-level visual features based on psychology were applied. Accuracy rates of more than 70% were obtained for classifying a variety of emotions.	Able to generate an emotional histogram displaying the distribution of emotions across multiple categories.	To improve the outcomes, more and better features are required.
RGB histograms, GIST descriptors, and mid-level features were employed as visual descriptors. On the Flickr dataset, the proposed model achieved an accuracy of 74.77%.	Multi-view (text + visual) embedding space is effective in classifying visual sentiment.	Investigation is needed regarding features to improve the system performance.
Unsupervised sentiment analysis framework (USEA) was proposed. Achieved an accuracy score of 59.94% for visual sentiment prediction on Flickr dataset.	The "semanticgap" between low- and high-level visual aspects was successfully resolved.	More social media sources, such as geo-location and user history, must be examined in order to boost performance even more.

Table 1. Summary of various existing works.

Extensive research has been conducted using completely automated systems that extract features from datasets without requiring human involvement [3, 4]. Positive, negative, and neutral aspects were identified in the ensuing investigations[5]. DAL scores, n-grams, and other novel features were used in the phrasal-level sentiment analysis for the categorization. Features that were employed included the polarity [6, 7] of syntactic components. But there was a catch to this system: in order to capture the exact emotion forecast, it needed an accurate expression boundary. Due to the variations in word production utilising DAL, which is not a component of speech, it also cannot manage polysemy.

Author and Citation	Data Type
Saad and Yang [1]	Twitter data
Fang *et al*. [2]	Review of consumer products and services on hotel
Afzaal *et al*. [3]	Restaurant and hotel data
Feizollah *et al*. [4]	Twitter keywords related to halal tourism and halal cosmetics
Mukhtar *et al*. [5]	Urdu blogs in multiple domains
Kumar *et al*. [6]	Textual and visual semiotic modalities of social data
Abdi *et al*. [7]	DUC 2002, and Movie Review Data
Ray and Chakrabarti [8]	Nikon Camera Data, and laptop domain data
Zhao *et al*. [9]	Social Media data
Park *et al*. [10]	Amazon reviews and Yelp reviews
Vashishtha and Susan [11]	multiple public twitter data
Yousif *et al*. [12]	Citation sentiment, and Citation purpose
Hassonah *et al*. [13]	Twitter Social Network data
Xu *et al*. [14]	Amazon product and Movie review data
Smadi *et al*. [15]	Arabic hotel's review
Maqsood *et al*. [16]	stock markets review
Abdi *et al*. [17]	Movie Review
Park *et al*. [18]	laptop and restaurant reviews from SemEval 2014
Bardhan *et al*. [19]	Text data from corpous
Araque *et al*. [20]	microblogging and movie reviews domain

Table 2. Various data types used for the work.

Convolutional neural networks (CNNs) combined with independent networks for nouns and adjectives allowed them to get superior outcomes. The sentiment ratings were obtained by the authors of by using a convolutional neural network (CNN) to process certain local areas and complete pictures that include sentiment information. The authors of presented an architecture to extract picture sentiments from the Twitter and ART photo datasets by combining a visual attention model with a CNN. A support vector machine (SVM) classifier was trained to ascertain the sentiment polarity after researchers from retrieved objective text descriptions from the photos in another study on image sentiment analysis. An accuracy of 73.96 was attained by integrating text and visual variables using a dataset consisting of 47235 photos. Using datasets from Twitter and Flickr photos, the authors of developed a long short-term memory model (LSTM) with deep learning capabilities for picture sentiment analysis. On Flickr and Twitter, they obtained accuracy rates of 0.91 and 0.83, respectively. In order to improve the outcomes of picture sentiment analysis, the authors of combined the semantic and visual data.

LSTM has proven to be more effective than traditional machine learning classification algorithms in obtaining higher levels of emotion categorization accuracy [16]. There will be some explosion and diminishing of result precision throughout the back propagation phases if the classification is carried out using a basic RNN. An improved RNN called LSTM was used to solve this issue because of its more intricate internal structure, which contains memory cells that allow architecture to retrieve data that has been held in the memory cells for a very long time [17]. Sentimental and content analysis have been used in tandem in a number of studies to uncover and analyse emotional messages in humans [18]. As many abusers have been using social media over the past few years, emotional computing has made its way into the field of machine learning. Additionally, many applications involving human-computer interface require these skills.

For the categorization, a greedy parsing method known as Transition based dependence is utilized in addition to the conventional parsers that were based on standard searching. The results of the classification will be more accurate with this parser. Yet, in contrast to a standard searching parser, the outcome deteriorates as a result of error propagation. The fact that the categorization can only be done on a minimal amount of data is another restriction. Social media is now an essential tool used by businesses to sell their goods in response to consumer demands and has grown in popularity. Therefore, assessing the accurate perception of public opinion is crucial. A variety of deep learning methods are available to identify the precise polarizing opinion. This study analyzed many deep learning categorization schemes according to their accuracy.

III **PROPOSED** WORK

Lexical Based Sentiment Analysis

With respect to techniques utilizing Twitter's lexical resources, there exist several opinion lexicons containing phrases associated with an emotion or an opinion value (polarity). On Twitter, however, lexicons are few. In terms of polarity lexicons, two noteworthy instances are the iSOL, which is taken from one of the most important English lexicons for polarity categorization, and the opinion lexicon developed by [12]. displays the Affective Norms for English Words (ANEW) translation on Twitter as well. In an independent investigation, the authors developed a lemma-level sentiment lexicon for several languages, including Twitter. Positive and negative phrases can be found in the well-known lexicon SentiWordNet.

Dataset	**No of instance**	Positive			Negative			
Dataset1	100000	56467			43533			
Dataset2	11102	angry	Boredom	Empty	Hate	Sadness		worry
		611	1678	1826	1022	1507		4458
Dataset3	11261	Enthusiasm	Fun	Happiness	Love	Neutral	Relief	surprise
		760	1775	5208	3844	962	1526	2186

Table 3. Data sets descriptions

.

Some methods consider topic content in addition to sentiment. Consequently, Anta et al. [14] have suggested a categorization framework for Twitter tweets by assessing the application of stemmers and lemmatisers, n-grams, word kinds, negations, valence shifters, link processing, search engines, unique Twitter semantics (hashtags), and various classification techniques; the best results are approximately 58% for topics and 42% for Twitter polarity detection. In addition, the authors in [15] used a naïve-Bayes classifier to identify the various degrees of polarity in Twitter tweets.

They also used unigrams of lemmas and multiwords based on PoS tag patterns to determine the extent of negativity. For four polarity levels, the accuracy of the system was 66%, while for six polarity levels, it was 55%. Ultimately, the authors of [26] provide a Transformer-based method that achieves accuracies of 80% to 90% in identifying negation in a corpus of Twitter product evaluations. However, this system has not been used on Twitter or in social network environments. In contrast to negation signals, alternative negation techniques are based on annotated corpora with negation. As a result, there are several corpora pertaining to clinical records, including the IULA Twitter Clinical Record, the UHU-HUVR, and the IxaMed-GS corpus.

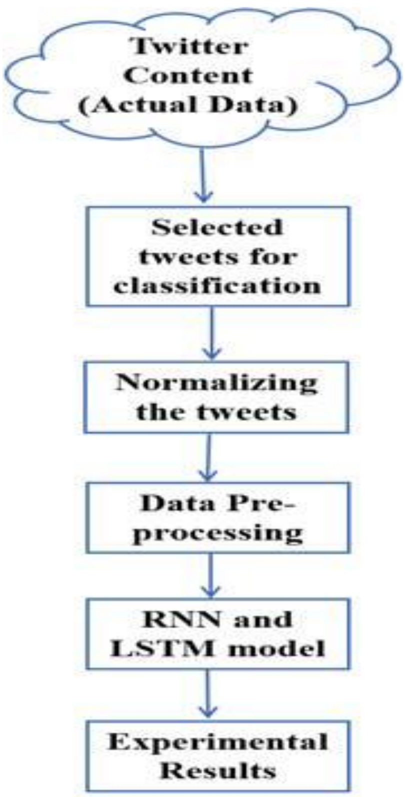

Figure 2. Working architecture.

The characteristics offered by convolutional and recurrent layers are combined in our deep learning model. In addition to proving their mettle in perceptual-related tasks like computer vision, convolutional neural networks (convnets) have also replaced other traditional machine learning techniques in other domains, like natural language processing. Convolutional networks' lack of memory is a key characteristic. Without preserving any state between inputs, each input that is displayed to them is processed separately. In contrast, reading a tweet requires processing each word individually and memorization of the parts that have already been read; this results in a coherent portrayal of the message contained in the tweet. Furthermore, the same concept of operation is used by recurrent neural networks (RNNs), which analyze sequences by iterating through their constituent components and preserving a state that contains information about what it has seen thus far. Consequently, gated recurrent unit (GRU) layers and long short-term memory (LSTM) layers are made to handle this issue [19].

By identifying text patterns, the convolutional layer suppresses several intermediary steps in the RNN layer's processing, therefore streamlining the work of the subsequent RNN layer. A fixed number of neurons in the convolutional layer will be maximized using an experimental process that will be explained in the next sections. Next, the MaxPooling1D layer reduces the dimensionality needed for the RNN stage that follows.

As the width and height dimensions grow, the number of feature-map coefficients to be processed for the next layer usually decreases. In the next stage, a recurrent neural network (RNN) is utilized, and as was previously said, there are two options: an LSTM layer or a GRU layer. For this reason, in our experiments, we will be looking at two different versions of the hybrid model in order to evaluate the performance of both hybrid architectures. A fixed number of neurons in the convolutional layer will be maximized using an experimental process that will be explained in the next sections. Next, the MaxPooling1D layer reduces the dimensionality needed for the RNN stage that follows.

As the width and height dimensions grow, the number of feature-map coefficients to be processed for the next layer usually decreases. In the next stage, a recurrent neural network (RNN) is utilized, and as was previously said, there are two options: an LSTM layer or a GRU layer. For this reason, in our experiments, we will be looking at two different versions of the hybrid model in order to evaluate the performance of both hybrid architectures. An LSTM layer will be the foundation for the first hybrid model, and a GRU layer for the second. An analysis will be conducted on the optimal parameters of the dropout rate, recurrent dropout rate, and neuron count that define the layer in the LSTM scenario.

IV. EXPERIMENTAL DISCUSSION

The suggested approach has been assessed using the twitter data set, which has been gathered and applied to various data sets for both positive and negative emotions. For the purpose of pre-processing the data, we normalised the tweets. In order to get the tweets in a normalised format and prevent undesired symbols, pre-processing actions are required. To achieve higher classification accuracy, the combined RNN and LSTM model is finally used. A emotion classification's most common steps are shown in Figure 3.

We used three data sets for our research and experiments in the publication. An already-existing dataset with one lakh tuples was used to create the first data collection. Additionally, both positive and negative categorization data are included in the dataset. There are 23938 tuples in the second data set, which pertains to the positive categorization. Enthusiasm, Fun, Love, Happiness, Neutral, Relief, and Surprise are the seven class names into which the positive data are divided. There are 26889 tuples in the last and third data set, which is for the negative categorization. Six categories Angry, Boredom, Hatred, Empty, Sadness, and Worry are used to categorise the negative data. Table 1 shows how the datasets' sentiment labels are grouped.

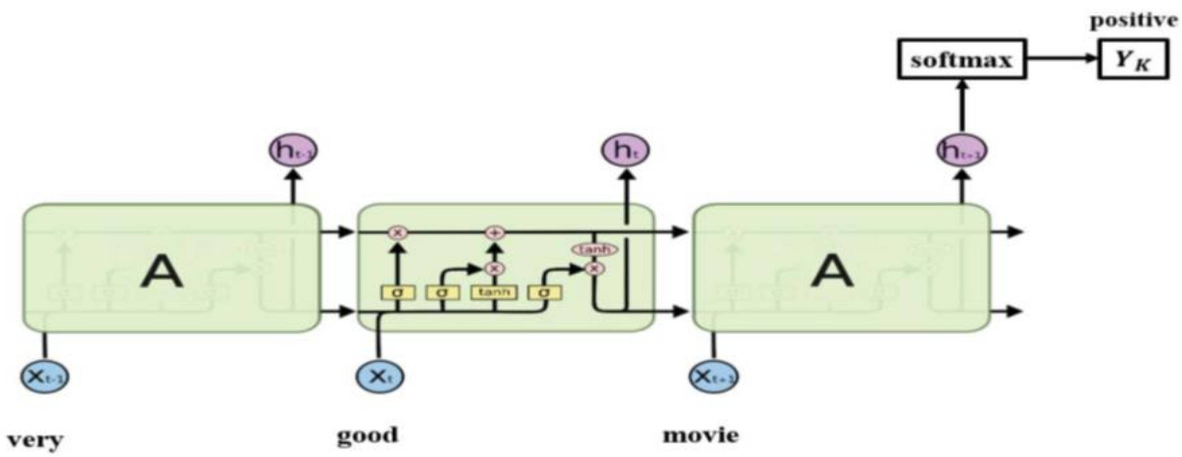

Figure 3. LSTM working model.

When analysing text data, natural language processing uses TF-IDF (Term Frequency Inverse Document Frequency) to find significant or uncommon terms. To make the data comprehensible for machine learning models, term frequency transforms string-formatted words into numerical-formatted data. The frequency of terms We utilise TF to determine the frequency of incidence of the terms we use in the classes. The feature in our data collection is words. The frequency of every word in the dataset that is being determined by term frequency.

$$\text{TF(I,j)} = \frac{term\ i\ frequency\ in\ document\ j}{total\ words\ in\ doc\ j}$$

where the phrase frequency in document j is represented by i. A table containing the frequency of words used in the document will be stored in memory. To determine which terms are more significant or least frequent in the document, utilise IDF (Inverse Document Frequency). IDF facilitates our extraction of important terms from the text.

First things first, we need to figure out which properties to feed the cell state. This choice of accepting or forgetting the characteristics is made by the LSTM architecture's forget layer. Two values will be produced by taking into account the values of h_{t-1} and x_t. In other words, the output values for each cell state can be either 0 or 1. Acceptance of the feature is indicated by a value of 1, while avoiding the information is indicated by a value of 0. The task at hand involves the acceptance or rejection of feature words derived from data sets in order to categorise emotions. The new emotion term has to be approved or refused in each cell state according on the tweets' emotions.

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f)$$

Finding additional information that has to be kept in the cell state is the next step in the categorization process. Which data should be updated will be determined by the sigmoid layer in collaboration with the input gate layer. Tanh layer creates the vector for the freshly updated data ct, which is then stored in the cell. We then update the status of the cell. In other words, sentimental analysis updates the state with fresh emotions that were lost in the preceding stage.

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i)$$

$$\check{C}_t = \tanh(W_C.[h_{t-1}, x_t] + b_C)$$

In conclusion, we apply the newly updated state CT to the cell state c_{t-1}. We multiply the state by foot after the previous state has been forgotten. We multiply the newly received value by i_t*c_t to obtain the new candidate value. This completes the process of updating the state and forgetting it.

Based on the filtered cell state is the final result. The output that we will generate is provided by the sigmoid layer when it is executed. Only by running it through the tanh layer and multiplying the result by the sigmoid layer's output can we produce the required portion. Our ability to deliver the desired product is facilitated by this.

Text is translated into numbers by the machine. Word embedding or word vectorization is the process of translating text or words to realvalued vectors. The process involves segmenting a text into sentences, which are then subdivided into words. This feature extraction approach then creates a feature map or matrix. With each row in the matrix representing a sentence or document and each feature column representing a dictionary term, the feature map's cell values therefore frequently show how many times a word appears in a phrase or document. "Bag of Words" (BOW) is one of the easiest methods for extracting features. It builds a fixed-length vector of the count, with each item being a word from a pre-established dictionary.

Model	Classification	Accuracy (%)
LSTM	Positive/Negative	88.47
	Positive subclasses	89.13
	Negative subclasses	91.3
RNN	Positive/Negative	83.21
	Positive subclasses	81.24
	Negative subclasses	87.02
LSTM	Positive/Negative	70.66
CNN	Positive/Negative	65.33
CNN	Positive/Negative	87.62

Table 4. Accuracy summary of models used in the work.

In a sentence, the count starts at 0 for terms that are not found in the specified dictionary and goes up or down by 1 depending on how many times the word appears in the phrase. Because of this, the vector's length always corresponds to the total number of words in the dictionary. Although this technique has significant drawbacks, it has the advantage of being easy to implement. It results in a sparse matrix, loses the word order of the phrase, and doesn't preserve the sentence's meaning, claim Bandhakavi et al. (2017) and Abdi et al. (2019). In the pre-defined dictionary I hope you are loving reading, for instance, the text "are you enjoying reading" would be rendered as (0,0,1,1,1,1) in the example. However, these representations can improve with pre-processing the text and the use of n-gram and TF-IDF.

Layer Size		Dropout Parameter		Kernel	Cross-Validation
Convolutional	LSTM	Convolutional	LSTM	Size	Accuracy
192	96	0.2	0.3	8	74.13%
160	128	0.2	0.2	8	74.10%
192	128	0.2	0.4	3	74.03%
128	96	0.4	0.4	5	73.99%
160	64	0.3	0.2	8	73.90%
192	64	0.3	0.3	3	73.81%
128	32	0.4	0.2	5	73.77%
96	64	0.3	0.2	3	73.66%
64	32	0.3	0.3	8	73.61%

Table 5. Summary of the test performed by LSTM.

Nonetheless, there are two methods for loading the embeddings while developing deep neural network models for the word processing task. Initially, the embeddings are learned during the network's training phase using data from the corpus. Our model employs the alternative strategy of using pre-trained word embeddings. In this case, the weights are matched to a specified dictionary that has been developed specifically to address this type of issue, rather than being updated with the corpus itself.

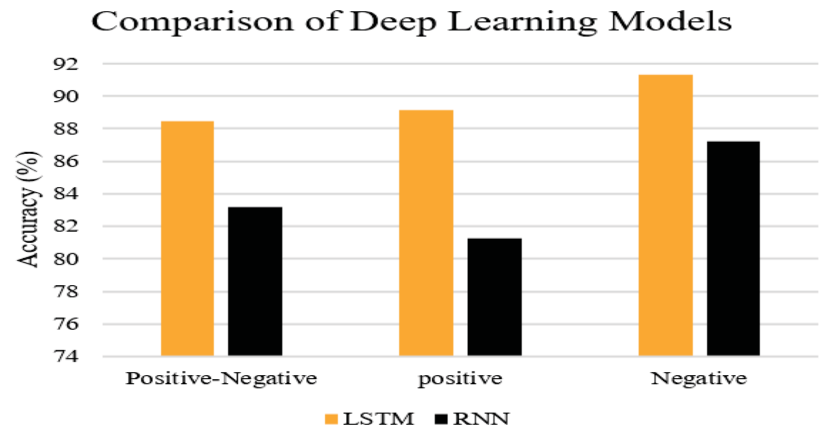

Figure 4. Accuracy obtained for sentiments with LSTM and RNN.

As part of our experimental setup, we employed the Kaggle dataset. The accuracy of the categorization of positive and negative tweets served as the experiment's assessment criterion. We looked at the accuracy attained for different datasets used by Ming-Hsiang Su et al., 2018 and Zhao Jianqiang et al., 2017 using the identical models, RNN and LSTM, for comparison's sake. The LSTM and RNN accuracy metrics across several categories are displayed in Table II. According to the experiment findings, the LSTM model outperforms the RNN and CNN classifiers in terms of accuracy. With the help of the Twitter dataset, the LSTM model was successfully trained. For the three categorization scenarios positive-negative, negative, and positive we have three datasets. We used accuracy as a performance metric to analyse the system.

V. CONCLUSION

In order to glean insightful information from the massive volumes of user-generated content available on social media, sentiment analysis via deep learning algorithms is essential. The accuracy and resilience of sentiment analysis models are getting better despite current obstacles because to developments in deep learning architectures. In the age of social media, sentiment analysis is expected to continue to be an essential tool for gaining insight into consumer behaviour, public opinion, and societal trends with further study and development. RNN and LSTM models have shown to be effective instruments for social media sentiment analysis, allowing companies, academics, and other entities to glean insightful information from user-generated material. The accuracy and resilience of sentiment analysis models are still being improved by developments in model designs and training methods, despite current obstacles. As long as research and development continue, RNN and LSTM models will be crucial for comprehending consumer behaviour, public opinion, and social trends in the ever-changing social media world. There aren't many deep learning-based methods for sentiment analysis in social networks accessible on Twitter right now, and the ones that are might benefit from more optimization. In fact, we propose a hybrid deep learning model combining convolutional and LSTM layers to extract the best from both approaches. We may use a tagged corpus with this model to determine the polarity of thoughts on Twitter. At around 81% and 92% accuracy for two and three polarities, respectively, our method beats previous models. By using our method, Twitter sentiment analysis performance is enhanced to 81% (up to 20 percentage points) more accurate than previous machine learning algorithms that have been developed with accuracy ranging from 54 to 59%. Besides, our model outperforms deep learning techniques, which achieve up to 55% accuracy, by 20 percentage points.

REFERENCES

[1] N. Yadav, O. Kudale, S. Gupta, A. Rao, and A. Shitole, "Twitter sentiment analysis using machine learning for product evaluation," in 2020 International Conference on Inventive Computation Technologies (ICICT), pp. 181–185, IEEE, 2020.

[2] D. Ramachandran and R. Parvathi, "Analysis of twitter specific preprocessing technique for tweets," Procedia Computer Science, vol. 165, pp. 245–251, 2019.

[3] N. F. Alshammari and A. A. AlMansour, "State-of-the-art review on twitter sentiment analysis," in 2019 2nd International Conference on Computer

Applications & Information Security (ICCAIS), pp. 1–8, IEEE, 2019.

[4] S. Niklander and G. Niklander, "Combining sentimental and content analysis for recognizing and interpreting human affects," in International Conference on Human-Computer Interaction, pp. 465–468, Springer, 2017.

[5]. Ravindra Changala, "UI/UX Design for Online Learning approach by Predictive Student Experience", 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA 2023), DVD Part Number: CFP23J88-DVD; ISBN: 979-8-3503-4059-4.

[6] B. Bhavitha, A. P. Rodrigues, and N. N. Chiplunkar, "Comparative study of machine learning techniques in sentimental analysis," in 2017 International Conference on Inventive Communication and Computational Technologies (ICICCT), pp. 216–221, IEEE, 2017.

[7] A. Sharma and U. Ghose, "Sentimental analysis of twitter data with respect to general elections in india," Procedia Computer Science,vol. 173, pp. 325–334, 2020.

[8] M.-H. Su, C.-H. Wu, K.-Y. Huang, and Q.-B. Hong, "Lstm-based text emotion recognition using semantic and emotional word vectors," in 2018 First Asian Conference on Affective Computing and Intelligent Interaction (ACII Asia), pp. 1–6, IEEE, 2018.

[9] Z. Jianqiang, G. Xiaolin, and Z. Xuejun, "Deep convolution neural networks for twitter sentiment analysis," IEEE Access, vol. 6, pp. 23253–23260, 2018.

[10] Bansal, B.; Srivastava, S. On predicting elections with hybrid topic based sentiment analysis of tweets. Procedia Comput. Sci. 2018, 135, 346–353.

[11] El Alaoui, I.; Gahi, Y.; Messoussi, R.; Chaabi, Y.; Todoskoff, A.; Kobi, A. A novel adaptable approach for sentiment analysis on big social data. J. Big Data 2018, 5, 12.

[12] Ravindra Changala ,"Development of Predictive Model for Medical Domains to Predict Chronic Diseases (Diabetes) Using Machine Learning Algorithms And Classification Techniques", ARPN Journal of Engineering and Applied Sciences, Volume 14, Issue 6, 2019.

[13] Ravindra Changala, "Evaluation and Analysis of Discovered Patterns Using Pattern Classification Methods in Text Mining" in ARPN Journal of Engineering and Applied Sciences, Volume 13, Issue 11, Pages 3706-3717 with ISSN:1819-6608 in June 2018.

[14] Ravindra Changala "A Survey on Development of Pattern Evolving Model for Discovery of Patterns in Text Mining Using Data Mining Techniques" in Journal of Theoretical and Applied Information Technology, August 2017. Vol.95. No.16, ISSN: 1817-3195, pp.3974-3987.

[15] Drus, Z.; Khalid, H. Sentiment Analysis in Social Media and Its Application: Systematic Literature Review. Procedia Comput. Sci.2019, 161, 707–714.

[16] Zhao, H.; Liu, Z.; Yao, X.; Yang, Q. A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach. Inf. Processing Manag. 2021, 58, 102656.

[17] Dashtipour, K.; Gogate, M.; Adeel, A.; Larijani, H.; Hussain, A. Sentiment Analysis of Persian Movie Reviews Using Deep Learning. Entropy 2021, 23, 596.

[18] Farisi, A.A.; Sibaroni, Y.; Faraby, S.A. Sentiment analysis on hotel reviews using Multinomial Naïve Bayes classifier. J. Phys. Conf.Ser. 2019, 1192, 012024.

[19] Melton, C.A.; Olusanya, O.A.; Ammar, N.; Shaban-Nejad, A. Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence. J. Infect. Public Health 2021, 14, S1876034121002288.

[20] Mishra, N.; Jha, C.K. Classification of Opinion Mining Techniques. Int. J. Comput. Appl. 2012, 56, 1–6.

[21] Kim, M.; Lee, S.M.; Choi, S.; Kim, S.Y. Impact of visual information on online consumer review behavior: Evidence from a hotel booking website. J. Retail. Consum. Serv. 2021, 60, 102494.

[22] Xiao, Z.;Wang, L.; Du, J.Y. Improving the Performance of Sentiment Classification on

Imbalanced DatasetsWith Transfer Learning. IEEE Access 2019, 7, 28281–28290.

[23] Ravindra Changala, Framework for Virtualized Network Functions (VNFs) in Cloud of Things Based on Network Traffic Services, International Journal on Recent and Innovation Trends in Computing and Communication, ISSN: 2321-8169 Volume 11, Issue 11s, August 2023.

[24] Ravindra Changala, Block Chain and Machine Learning Models to Evaluate Faults in the Smart Manufacturing System, International Journal of Scientific Research in Science and Technology, Volume 10, Issue 5, ISSN: 2395-6011, Page Number 247-255, September-October-2023.

[25] W. Wunnasri, T. Theeramunkong, and C. Haruechaiyasak, "Solving unbalanced data for Thai sentiment analysis," Proceedings of the 2013 10th International Joint Conference on Computer Science and Software Engineering, JCSSE 2013, pp. 200–205, 2013.

[26] N. Chirawichitchai, "Emotion classification of Thai text based using term weighting and machine learning techniques," in Computer Science and Software Engineering (JCSSE), 2014 11th International Joint Conference on, 2014, pp. 91–96.

[27] P. Sarakit, T. Theeramunkong, C. Haruechaiyasak, and M. Okumura, "Classifying emotion in Thai youtube comments," 2015 6th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES), pp. 1–5, 2015.

[28] Sudprasert, S., Kawtrakul, A., "Thai word segmentation based on Global and Local Unsupervised learning." In NCSEC, Chonburi, Thailand, 2003.

[29] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," The Journal of Machine Learning Research, vol. 12, pp. 2121–2159, 2011.

[30] Kemp, S. Digital 2022: Global Overview Report; Techreport; Kepios Pte. Ltd.: Singapore, Singapore, 2022.

[31] Mäntylä, M.V.; Graziotin, D.; Kuutila, M. The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. Comput. Sci. Rev. 2018, 27, 16–32.

[32] Martínez-Cámara, E.; Martín-Valdivia, M.T.; Urena-López, L.A.; Montejo-Ráez, A.R. Sentiment analysis in Twitter. Nat. Lang.Eng. 2014, 20, 1–28.

[33] Cervantes, I. El español en el Mundo 2021. Anuario del Instituto Cervantes; El español en el Mundo, Instituto Cervantes: Madrid,Spain, 2021. 2021.

[34] Ravindra Changala, A Dominant Feature Selection Method for Deep Learning Based Traffic Classification Using a Genetic Algorithm, International Journal of Scientific Research in Computer Science, Engineering and Information Technology, ISSN : 2456-3307, Volume 8, Issue 6, November-December-2022, Page Number : 173-181.

[35] Perez-Rosas, V.; Banea, C.; Mihalcea, R. Learning Sentiment Lexicons in Spanish. In Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12), Istanbul, Turkey, 23–25 May 2012.

[36] Betancourt, E.R.; Chacón, P.S.; Murillo, E.C. Deep Neural Network Comparison for Spanish Tweets Polarity Classification. In Proceedings of the 2019 XLV Latin American Computing Conference (CLEI), Panama, Panama, 30 September–4 October 2019; pp.1–6.

[37] Srivastava, R.; Bharti, P.; Verma, P. Comparative Analysis of Lexicon and Machine Learning Approach for Sentiment Analysis.Int. J. Adv. Comput. Sci. Appl. 2022, 13, 71–77.

[38] Köhn, A. What is in an Embedding? Analyzing Word Embeddings through Multilingual Evaluation. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Lisbon, Portugal, 19–21 September 2015; pp. 2067–2073.

[39] Ravindra Changala, "Integration of IoT and DNN Model to Support the Precision Crop", International Journal of Intelligent Systems and Applications in Engineering, Volume 12, Issue 16s), February 2024.

[40] García-Cumbreras, M. Á.; García-Vega, M.; Gutiérrez, Y.; Martínez Cámara, E.; Piad-Morffis, A.; Villena-Román, J. TASS 2018:The Strength of Deep Learning in Language Understanding Tasks. Proces. Leng. Nat. 2019, 62, 77–84.