International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 12, No.1 (Jul-2022)

https://dx.doi.org/10.12785/ijcds/120116

Deep Sentiment Approaches for Rigorous Analysis of Social Media Content and Its Investigation

Amit Kumar Sharma[1], Sandeep Chaurasia[2*] and Devesh Kumar Srivastava[3]

[1]*Department of Computer Science and Engineering, Manipal University Jaipur, India, 303007*
[2]*Department of Computer Science and Engineering, Manipal University Jaipur, India, 303007*
[3]*Department of Information Technology, Manipal University Jaipur, India, 303007*
[1]*Department of Computer Science and Engineering, IcfaiTech, The ICFAI University, Jaipur, India, 303031*

Received 7 Jul. 2021, Revised 16 Jun. 2022, Accepted 20 Jun. 2022, Published 1 Jul. 2022

Abstract: Social media has a very important contribution to human lives today. Through social media platforms people can share their information, ideas, knowledge, and activities with connecting people in the form of videos, images, texts, and audios. In the context of sharing information, incorrect information is also shared along with the correct information. In this way, unauthentic (fake news), misleading (rumors), abusing, toxic, extremist contents are also shared through social media platforms. This paper reviews the influences of social media content. In this context, vector representation of the social media sentences, word embedding models has been best applied for better accurate results. Natural language processing (NLP) and text analysis techniques are being used to extract useful information from social media content. The NLP techniques are widely used for correcting the sentences and identifying their meaning also. Currently, machine learning (Decision Tree, Random Forest, SVM, Naïve Bayes) and deep learning (LSTMs, BLSTMs, GRUs, CNNs) models are successfully being implemented to classify social media contents. In the comparative study of different works of literature and results from LSTM and CNN-LSTM deep learning model have been proved that deep learning and the word embedding model provide better accurate results for social media contents categorization.

Keywords: Deep Learning, Machine Learning, NLP, Rumors, Social Media, Text Analysis, Word Embedding

1. Introduction

Social networking platforms like Twitter, Facebook, Instagram, and WhatsApp, etc., are widely used by billions of people to interact with each other. Through these platforms, people show their views on the post by sharing, liking, replying to those posts. Also unauthentic, toxic contents, rumors, fake contents, abusive contents, etc., are also shared on social platforms that eventually affects the human mind and divert them from society [1] [2] [3]. Such contents or post increases negativity and criminal activities in the society. Text, images, videos, and speech contents are widely used on these platforms in different languages. This review is focusing on text and speech contents because of their wide availability in the literature, and most contents are circulated in text and speech form in the social media. Social media platforms provide the facility of language plurality, by using these options users can send their message or post in their languages [4] [5].

Analysis of the contents, such as syntax, sentiment, and semantic analysis with huge sparse data pose a challenge to the researchers. There are several models which are present in the literature that perform analysis of different languages and different forms of data. N-gram language model is used to identify language features like spell correction, speech recognition, next word suggestion, text summarization, etc. [1] [6]. Misinformation is spread on social media to create confusion among people as well as they have harmful consequences. Stopping these rumors and misinformation to spread is the biggest problem that researchers are facing. It takes only a second for the public to share the information but validating the information is necessary else it will have the negative impact. Celebrity death rumors, chain mails, falsities about the Social Network, etc., are the examples of the unauthentic news [4].

In the current scenario, the novel Corona virus which firstly emerged in the Wuhan city of China in December 2019 is continuing to infect people across all countries. This corona virus (Covid19) has become the biggest threat to all the countries. World Health Organization (WHO) and the leaders of the countries are communicating with their public and doing awareness programs through social platforms. In between this, fake news and rumors are also being spread continuously through social platforms and impacting society greatly. For example, a rumor on the 14th

*E-mail address: amitchandnia@gmail.com,chaurasia.sandeep@gmail.com, devesh988@yahoo.com. *Corresponding Author: Sandeep Chaurasia*

and 15th march of 2020 started getting viral by Americans that "martial law is coming". This message was hard to stop or even trace. Due to this misleading information viral on social media, the U.S. politician Sen Marco Rubio, R-Fla tweeted and debunked this rumor. Normally, Indian people get daily messages regarding health topics and are receiving more messages in this Corona pandemic. Such type of messages is not sent after fact-checking and those messages impact people's habits and lifestyle. Some organization works for debunking or destroying such rumors, unauthentic messages, or post by fact-checking. The Indian government also launched a help desk MyGov Corona for awareness about the Corona virus pandemic and tackles social media services of spreading misinformation.

Social media have contributed a lot to society in the modern era and through which society can share their information and news on public platforms. Many social, educational, governments, and private organizations deliver their messages or news to people through social media. As many people have benefited from social media, they also suffered a lot. Different types of content are shared on social networking websites which is very harmful and can divert people in the wrong direction. Many researchers have explained the side effects caused by such content in their researches. In this sequence, unauthentic, misuse, toxic, rumors, negative, fake, domestic violation, illicit drug, abuse, extremist, cyber bullying, etc., contents can affect human minds as well as divert people from society. Following are the social impacts because of social media:

Rumor: Social media have some harmful, unusual, fake, unauthentic contents which are known as Rumors. Rumor is shared to misguide the people. A rumor is propagating to damage the reputation of persons or an organization. Without verification, rumor reaches thousands of people immediately and causes serious damage.

Domestic Violence: It is a very critical issue and harmful for the society. The victim sometimes shares their story or health issues on social media. In current years, domestic violence crisis support organizations are very much active on social media and serve support to them.

Abusing Contents: Abusing content is also posted on social media. The effects of such content are on the psychology of teenagers and also it demoralizes them.

Extremist Contents: Jihadist propaganda is spread on social media by extremist organizations and they spread their propaganda to mislead the people and for recruits.

Cyber Bullying: It includes sharing negative, false and harmful messages on social media about someone else causing humiliation and embarrassment to them.

Toxic Content: The presence of toxic content has become a major problem for many online communities. It includes racism, sexual predation, and other negative behaviors that are not tolerated in society.

Illicit Drug: In current. Social media has become a popular platform to offering new drugs and alcohol to Youngers and teenagers.

Social media content analysis is very important to stop rumors and false information. For this, Natural Language Processing (NLP) is used which helps in linguistic text analysis. The linguistic text analysis can read the text of the different languages. In the NLP process, it is required to convert words into numbers or vectors so; word embedding methods are used to convert words and phrases into vectors. In recent years, word embedding shows a boom in the performance of text analysis tasks. tf-idf, Word2Vec, Doc2Vec, and GloVe models of word embedding are widely used for vector representation of words and phrases. An NLP language method that is n-gram is used to solve the problem of language identification from social media content. In recent years, machine learning algorithms are also used to identify the patterns from the text and help to classify the social media contents. The supervised machine learning methods (Decision Tree, Logistic Regression, SVM, Naïve Bayes, KNN) are being successfully used for the classification of social media content [1] [7] [8]. In this context, deep learning architectures are also being applied in the field of text analysis, NLP, speech recognition, image processing, etc. Apart from this, RNN, LSTM, BLSTM, GRU, CNN are different methods that are used for text and image analysis [9] [10] [11].

The paper is focused on different problems of social media that are usually posted in the form text. These texts are in the sequential form and sequential content have hidden information. The paper has suggested a hybrid CNN+LSTM deep learning classification model with pre-trained Word2Vec word embedding classify the tweets into rumor and non-rumor. This model is useful for every social media problems which are in the sequential text form.

Further paper is organized as: Section 2.0 presents the related works. Section 3.0 presents the methodologies of text analysis. Section 4.0 presents the comparative study of different social media problems. 5.0 present the rumor classification model, experiment evaluation and results analysis of the deep learning text classification model. Section 6.0 presents the analysis and discussions of different social media problems and Section 7.0 concludes the paper and presents the future scope of the research work.

2. Literature Review

There is a lot of content on social media platforms in the form of text, videos, images, etc. These platforms provide service-specific applications which are governed by some organizations. Social media facilitated the growth of online social networks by connecting people of similar interests. People hold both kinds of sentiments; positive and negative. Social media content has the text of semantic and sentiment knowledge [10] [11] [8]. These contents have some harmful, unusual, fake, unauthentic contents which misguide people. In the literature, there are some APIs available that helps in extracting social media post's content for sentiments and sematic analysis, like Twitter API [12] [8] [13], Beautiful Soup [14], and Facebook Graph API [15] [16]. The main focus of current research is social media content analysis and producing authentic content for society. For that researchers are collecting information manually from the news articles, Wikipedia entries, site pages, related magazines [7] [8], and some are using publically available datasets for research purposes, like movie review dataset [11], Twitter Sentiment Corpus [17], Priyo Review [18], Kaggle [1] [19], PHEME [9] etc.

In [20] the author proposed a deep learning model which was based on CNN to detect rumors on Twitter and prove that the existing state of the art methods requires improvement. In this paper, the authors proved that the CNN deep learning architecture obtained more accurate results as compared to existing machine learning algorithms. The model was trained on the publically available PHEME dataset. This research finds that the tanh activation function provides better accurate results as compared with the RELU activation function. Another deep learning model is proposed using RNN classifier for classifying tweets into rumors and non-rumors classes [9], and the results are compared with the machine learning classifiers. In this paper, the authors trained the model with different features. The first model was trained with textual and user characteristics features and traditional machine learning classifiers (SVM, KNN, Gradient boosting, and Random forest). The second model was trained with applied LSTM deep learning architecture on only tweet text, and the third model was trained with tweet text and user metadata features and LSTM deep learning architecture. The Second model has performed better than the machine learning-based classifier. This research also suggested that the machine learning approach is a very time-consuming process and cannot preserve the semantic representation and sequential representation of the sentences. This research also suggested that the deep LSTM model learns the hidden information from the tweet text which was difficult to learn from the hand-crafted feature and machine learning features.

In [21] the authors focused on the classification of both rumors and non-rumors features and noticed that the classified results detect only from the rumored features, so binary classification not might provide beneficial results. To solve this, existing literature deals with only rumored features and suggests a new approach which is one class classification classifier with one class feature. Rumored features were extracted from already available and detected features of rumors on a social network. In the paper, the data is trained on seven class classifiers, namely Autoencoder, Gaussian, K-Means, KNN, SVDD, OCSVM, and PCA which is applied on two major datasets, namely Zubiagaset and Kwonset. The performance of the OCC model was observed by a high level of F1-score. The approach achieved a 74.30% F1-score for the Zubiagaset dataset and 93.98% for Kwonset dataset.

Another research proposes a deep learning model for recognising breaking news, rumors, rather than long-lasting rumors [22]. In this, word vectors were generated from the word embedding model and LSTM-RNN deep learning approach is applied for identifying the rumors. This study also suggested that a deep learning approach with word embedding is performed better than the state of the art method in term of precision, recall, and f1 measures. In [4] a model for rumor detection is proposed which is based on the RNN deep learning approach for learning hidden knowledge from the posts. Experimental results compared with the existing machine learning approach with handcrafted features. In this research, the machine learning model trained with a decision tree, SVM, random forest classifiers, and deep learning model was trained with tanh-RNN, LSTM, GRU-1, GRU-2. GRU-2 with multiple hidden layers and provides better accurate results as compared to different models of deep learning and machine learning model. The experimental results have been measured in terms of precision, recall, F1, and accuracy. A deep learning model for automatic content categorization in multi classes on online posts was also proposed [15], which proves that the deep learning model provides the solution for real-world problems over the traditional machine learning techniques. This research provides the results; comparisons between word vectors generated by domain-specific word embedding and pre-trained word embedding. The performances were evaluated with 5 deep learning models, namely RNNs, LSTMs, GRUs, BLSTMs, and CNNs. Also, performances were evaluated with machine learning approaches, namely SVM, RF, DT, and LR. With GloVe embedding, GRUs and BLSTMs performed the highest with scores of 91.78% and 91.29%, respectively. Further, researchers have also identified domestic violence problems through content [16], shared by victims on social media and some social service organizations search these types of content and find the victim to help them. In this model, data was extracted from Facebook by using the Facebook Graph API and labeled the data into critical and uncritical posts. The model, first accurately evaluated with traditional machine learning models using different ML classifiers with different word setting. LR classifier and tf-idf word embedding with stemming of the words obtain 90.74% accuracy. In the second evaluation of deep learning, Word2Vec+LSTMs obtain 93.08% accuracy and Glove+GRUs obtain 94.26% accuracy. In machine learning, evaluation results were evaluated with NB, SVM, RF, LR, and DT and in deep learning, evaluation the results were evaluated with CNNs, RNNs, LSTMs, GRUs, and BLSTMs classifiers. It is concluded that deep learning models achieved better performance results than traditional machine learning models.

A C-BiLSTM (Convolutional Bi-Directional LSTM) deep learning model for automatic identifying inappropriate (abuse, rude and discourteous) comments on language is proposed [23], and is applied in the real-world language which significantly performs better than both handcrafted feature and pattern-based approaches. [1] proposed a machine learning approach for the detection of abusive content on social media. In this, the skip-grams feature improved the results as compared to previous approaches.

Cyberbullying incidents in social media platforms are also detected by the deep learning model [24]. In this, the model was trained on different publically datasets, namely Wikipedia, Formspring, and Twitter, and experimental evaluation is done with four deep neural network models, namely CNN, LSTM, BLSTM, and BLSTM with attention. For vector representation of the sentences, word embedding models (random, GloVe, and SSWE) was applied. After validating findings from Wikipedia, Twitter, and Formspring, the work was expanded on the new YouTube dataset and investigated the performance of the models in the new social media platform. The experimental result shows that the DNN model has successfully been implemented for all social network platforms and the results also suggest that the performance of the DNN model is better than the machine learning model. A machine learning classification model is proposed for detecting bullying and aggression posts on Twitter [10]. In this, the preprocessing process is applied, like removing stop words, URLs, punctuations, repetitive words, stemming, and labeling. For creating word vectors, the Word2Vec model is applied and it detects the sentiments using the SentiStrength tool. For classifying the results, J48, LADTree, LMT, NBTree, RF, and functional Tree is applied and obtained 90% accuracy to detect bullying and aggression or hate speech comments.

Further, the literature showcases a deep learning-based sentiment analysis model for classifying the tweets into extremist and non-extremist [12]. In this model, after preprocessing of the sentences, the word embedding model is applied for vector generation of the words. The dataset was trained with LSTM+CNN deep learning network and performance were evaluated in term of accuracy, recall, precision, and f-measures. The model accuracy obtained by 92.66% and its performance is better than machine learning and other deep learning classifiers. Sentiment analysis and identification of the toxic online comment by using the SentiWordNet tool is in focus [15]. In this research, researchers have explored various aspects of sentiment detection and their correlation and obtained the results by using a toxicity detection tool. A CNN deep learning architecture is also proposed for opinion mining [11]. This research mainly focused on sentiment analysis of the movie review. In this process pre-trained Word2Vec model was used for vector representation and information gain, the model was trained with CNN deep learning architecture and the accuracy of the model was 97.3% which shows the importance of the deep learning concepts. A deep neural network model is proposed for automatically identifying a subset of webpages and social media content that has extremist content [7]. For dealing with different language challenges, the script Unicode was applied to converts all text into the corresponding ASCII characters. The deep neural network with Doc2Vec was used to classify the text into extremist and non-extremist. z Moreover, a modeling-based approach to identify illicit drug-related contents from social networking sites is also proposed [25]. In this, a dataset was created from the NIDA website with 371 hash-tags. Word embedding (Word2Vec) model was used for vector generation and LDA topic modeling algorithm is applied to identify the illicit drug-related contents. The model has obtained 78.1% accurate illicit drug contents. A language identification model [6] is proposed for short segments of the text contents. In this model, the n-gram model was used to correct the sentences of different languages posts and a common n-gram distance based novel model was used for classifying the results. The Common N-Gram text classification is used with different classifiers, namely Logistic Regression, SVM, Naïve Bayes, and Random Forest.

3. Methodologies of Text Analysis

Social media users are continuously increasing day by day and billions of users are sharing a huge amount of content in different forms like texts, videos, and images, etc., daily. Text is the most common form of content that is used in social media platforms, so it is necessary to find the knowledge from the huge text content. In text analysis processes, a huge amount of unstructured data is collected from different sources by applying different techniques and methods and converts these unstructured data into knowledgeable structured data. The text analysis process helps to explore results and identify patterns, keywords, and attributes of the unstructured text. There are some different methods and processes which have been used in previous researches for text analysis:

A. Process of Text Analysis

1) Natural Language Processing (NLP)

NLP is widely used in the process of text analysis [13], which explores how a computer system becomes an expert system in the context of understanding natural human languages and develops some tools and techniques that can perform to manipulate human languages to the desired task. In the existing literature, lexical and syntactic analysis of the NLP has been applied in the text analysis process, for example, in a bag of words representation technique, only lexical components of the text are considered. Sentimental and semantic analysis of the words is broadly used in the text analysis process. Semantic analysis has been successfully applied in the text and has improved the results [13] [5]. The n-gram model is a probabilistic model which is applied in many NLP and the text analysis process [1] [6]. In n-gram model, the sequence of words is extracted from the sentences or text and probabilities are assigned to them. The n-gram of the size 1 is known as a unigram, the size of 2 is known as bigrams, the size of 3 is known as trigrams, and so on. The n-gram model is used in different tasks like sentence correction, spell correction, word breaking, suggestions while typing, text summarization, etc., and also uses supervised machine learning models for developing the best features.

2) Data Preprocessing

In-text analysis process, there is a requirement of cleaning the data because data are very sparse, unstructured, and noisy. So, in data preprocessing, lexical analysis, tokenization, stop word removing, case folding, special character removing, deleting hyperlinks, normalization, stemming etc., is done on the data [13] [18].

3) Word Embedding for Data Representation

The text analysis deals with huge, raw plain text data, and the machine only understands the numbers. Word embedding is the model that is used to extract features from the text data and convert these texts or word sentences into vectors or numbers so that the machines can understand them. This model is used to project in continuous of the words and deal with the syntactic and semantic similarities between the words of a sentence. The word embedding model has been very effectively applied in NLP, machine learning, and deep learning processes. Word Embedding methods are generally trying to map a word using a dictionary to a vector. Frequency-based word embedding (tf-idf) and predication based word embedding (Word2Vec, GloVe, Doc2Vec) methods have been successfully applied in the research for representing word vector. In frequency-based embedding (tf-idf), the document's text transforms into numeric vectors, this representation is called the Vector Sparse Matrix (VSM) or Bag of Words model. In tf-idf, a particular document weights each word; this word weight makes the importance of the word in the document. The prediction based Word2Vec method is obtained from two models that are continuous bag-of-words (CBOW) and Skip-gram, both models worked with neural network concepts [25] [8] [26].

The Word2Vec CBOW Model uses the surrounding words as inputs to predict a word in a sentence. CBOW is a word embedding model that predicts the target word x_0 from the surrounding contextual words, C i.e. the goal is to maximize $P(x_0 jc)$ throughout the training set. The distance between the current vectors assigned to x_0 and to c is inversely proportional to this probability. The model's goal is to reduce the distance between x_0 and c 's current vectors (and enhance the probability $P(x_0 jc)$. By repeating this process over the whole training set, we may build vectors for words that co-occur and tend to be closer together. The input of CBOW is one-hot encoded vectors V, as shown in Figure 1. This means that for each vector, only one of the V units will be 1, while the rest will be 0. The CBOW model works as a three-layer basic neural network, with two weight functions in the input layer, hidden layer, and output layer. A VxN matrix W can be used to represent the weights between the input and hidden layers (i.e V is the number of words and N is the number of neurons in the hidden layer). The N-dimension vector representation v_x of the related word from the input layer is represented in each row of W. To construct the word vectors, the input layer's weight function matrix (i.e. W input) is used.

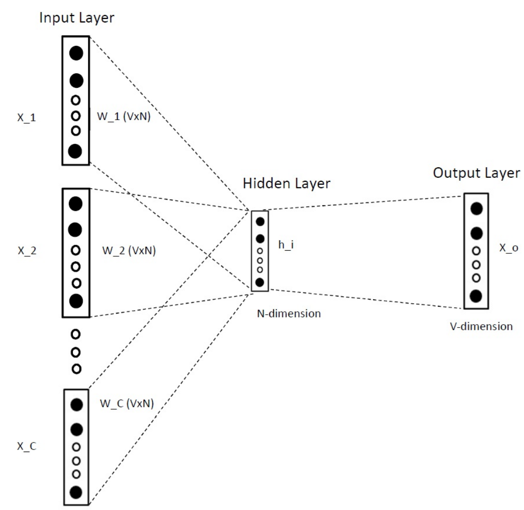

Figure 1. CBOW Model: Predict the target word from all neighbouring words

The skip-gram model is a version of the CBOW model that is reversed. The target word is now at the input layer in the skip-gram model Figure 2, while the other words in the window are at the output layer. We still utilize v_x as the input vector for the input layer's only word, and the hidden layer outputs h are defined in the same way as in CBOW, which means h is just copying a row of the input $>$ hidden weight matrix, W, mapped with the input word x_i.

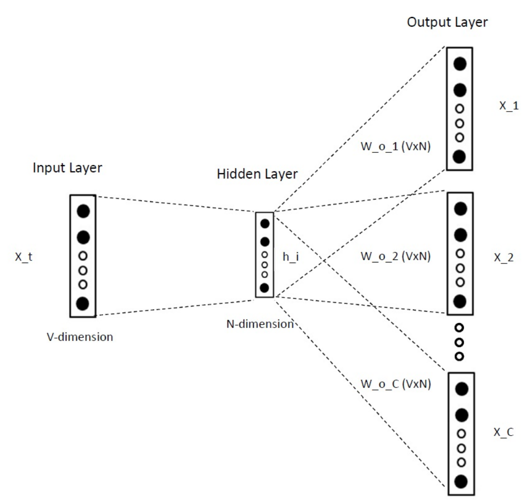

Figure 2. Skip-gram Model: Predict the context words using the main word

The Skip-gram model attempts to anticipate the surrounding keywords by identifying relevant word representations in a sentence or text. The goal of the Skip-gram model is to maximize the average log probability (i.e.$logp(x_1; x_2; ::: x_C|x_t)$ from target words x_t from context, where C is the size of the training window, given a sequence of context words $(x_1; x_2; x_3; ::: ; x_C)$. A larger C indicates that there are more training examples, which can lead to a greater accuracy. However, it also raises the expense of training time.

In the word2vec model, the vector of the words is generated by achieving the process of prediction of surrounding words in a sentence. While in GloVe model, the model learns by constructing a co-occurrence matrix that count how a word frequently appears in a context. In this model, firstly a co-occurrence matrix X is constructed from the training dataset. Where X_{ij} is the frequency of the word I co-occurring with the word j. $X_{ij} = \sum_k^V X_{ik}$ is the total number of occurrences in word i in the dataset. In the second step, the factorization of X gets vectors and reduces noise by identifying relevant words.

Another prediction based word embedding model is Paragraph Vector Doc2Vec (PV-Doc2Vec) model. The idea of the model is inspired by the Word2Vec model. This model is the extension of the Word2Vec model concept. In the CBOW model of the Word2Vec, the trained model predicts a center word by using the context words of a sentence. PV-Doc2Vec, on the other hand, takes a sample of consecutive words from a paragraph at random and predicts a center word from the randomly chosen set of consecutive words using the paragraph id and context words as input. The model is divided into three sections: 1) Paragraph matrix: the matrix where each column represents the vector of a paragraph, 2) Average/Concatenate: it checks whether the paragraph matrix and word matrix are concatenated or averaged, and 3) Classifier: the hidden layer vector (the one that was concatenated/averaged) is sent into the classifier, which predicts the center word in Figure 3.

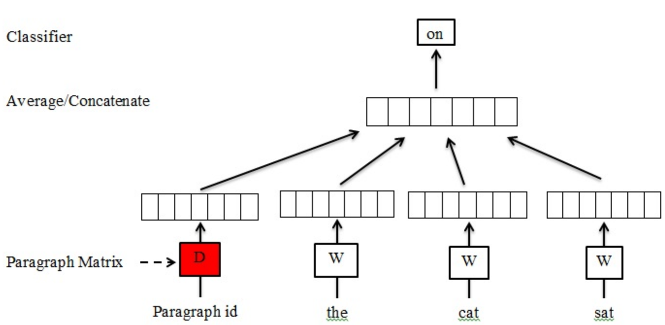

Figure 3. Doc2Vec Model: Distributed bag of words of paragraph vector

Generating a new model with word embedding process takes too much time; to solve such issue, a pre-trained word embedding model is used which is already trained by someone and is publically available for research. Currently, pre-trained Word2Vec and pre-trained GloVe model is widely used for word vector representation [15] [16] [11].

B. Multilingual Text Analysis

Social media user posts their messages in different languages. The NLP task requires a single language platform that can show all text in a single language [6]. Single language platform provides good results of semantic analysis.

C. Speech to Text Conversion

Currently, artificial intelligence, automation and many more applications has been successfully used in the speech recognition process to develop its applications. Speech recognition is the process in which speech or spoken texts are converted into the written text. Google Speech Recognition is one of the easiest methods used for speech text analysis.

D. Machine Learning

Previously, machine learning concepts are broadly used for social media content categorization with more accurate and effective results. The use of NLP methods with machine learning concepts, the generated model helps to identify patterns from the customer's message. Supervised Machine Learning (SML) is suitable for the classification of social media content. Machine learning's social media service aids in improving media quality, assisting brands in reaching target audiences, keeping security, managing, and automating data. Some machine learning approaches that have been effectively applied in social media applications include Logistic Regression (LR), Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), Nave Bayes, and K-Nearest Neighbor (KNN).

E. Deep Learning

Currently, the deep learning approach is used to detect unsupervised and unlabeled data; which is also known as deep neural networks. The method has been successfully implemented to develop different applications of speech recognition, image processing, bioinformatics 1, text filtering, NLP, etc. In a deep learning process, the trained model automatically learns and performs a classification task from the text, image, and audio. A deep learning model trained with a huge number of labeled data and neural network architecture. In the traditional process of neural network architecture, 2-3 hidden layers were used to detect the feature from the data but now, many hidden layers are being used to train the model for directly learning the features from the data. RNN, LSTM, BLSTM, GRU, CNN model of deep learning are used to detect features and classify the data [8] [27].

1) Recurrent Neural Network (RNN)

The feed-forward neural network process, cannot predict the next word in a particular sentence because there is no relation between previous output and current output. To overcome these issues the RNN architecture is used for the prediction of the next word in a particular sentence. A RNN is a deep learning concept, in which neuron connection is established with a direct cycle. It means output depends on previous neurons as well as present inputs Figure 4. The concept of RNN solved various problems of NLP like handwritten recognition and speech recognition etc.

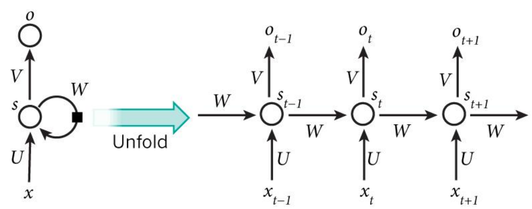

Figure 4. RNN Architecture for Sequential Data

2) *Long Short-Term Memory (LSTM)*

RNN uses a back-propagation algorithm, but it is applied for every timestamp, and back-propagation has a vanishing gradient problem. To solve this problem, a specific RNN architecture is developed that is LSTM, in which a model was designed for learning long term dependencies. In the LSTM process, the activation function is not used for its recurrent components. In the architecture of LSTM, several units of blocks are implemented with four gates which are input, forget, and output gate which uses logistic function to control information flow in the network Figure 5.

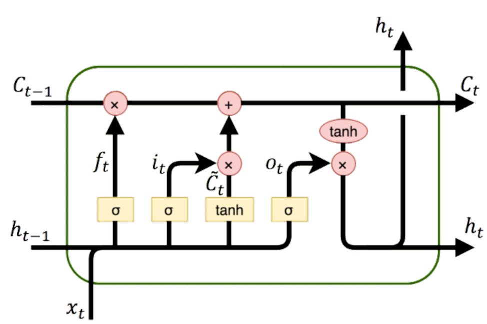

Figure 5. LSTM Architecture for Learning Long Term Sequences

3) *Bidirectional Recurrent Neural Networks (BLSTM)*

The BLSTMs model is trained with two LSTMs model instead of one LSTM in the input sequence. Implementation of this architecture, the first LSTMs input sequence is to be as it is and the second LSTMs input is to be a reverse copy of this input sequence. The BLSTMs provides additional context to the network and also improve the performance of the results Figure 6.

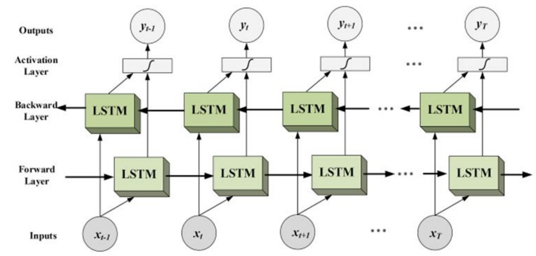

Figure 6. BLSTM Architecture for Learning Sequences using two LSTM

4) *Gated recurrent unit (GRU)*

The GRU architecture is similar as LSTM. In GRU architecture, there is no cell state; a hidden state is used to transfer the information Figure 7. Only two gates were employed in the GRU's secret state: the reset gate and the update gate. GRU's update gate is similar to the LSTM method's forget and input gates in that it determines what information will be delivered and what information will be added. The reset gate of this method is used to decide how much previous information to leave or forget which a GRU is. Currently, it is not clear that which architecture is better, but GRU's tensor operations are little fast to train data than LSTM. The researcher usually tries both the architectures to identify which is better for their information.

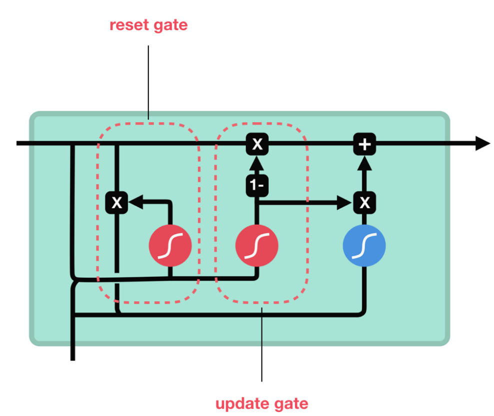

Figure 7. GRU Architecture for Sequential Learning

5) *Convolutional Neural Network (CNN)*

The CNN architecture has been proven to provide outstanding results for speech and image processing. The CNN architecture includes: a) convolutional layer/s, b) pooling layer, and c) a multilayer perceptron variation Figure 8. The result of the convolutional layer is passed to the next layer via the convolution technique. This procedure allows for a much deeper network with many fewer parameters. A CNN model is proposed for short sentence classification [5]. The CNN model uses the Word2Vec model for feature vector representation and conducts a series of experiments and demonstrated that the suggested model performs excellently.

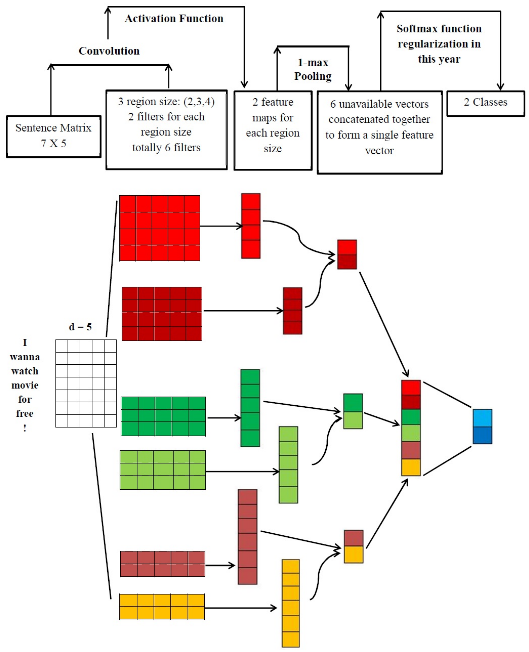

Figure 8. CNN Layered Architecture for Text and Image Processing

4. Comparison of Different Social Media Problems

The comparative study table in Table I showcases different social media problems suggested by the different researchers and also found the objectives of those researches for such problems. Further, the comparative study of social media problems provides the detail of the different datasets, used methods, and techniques to solve those identified problems.

5. Rumor Classification Model

The presented rumor classification model is based on deep learning approach and predicts the result in category of rumors and non-rumors. Word vectors are generated by Word2Vec model that helps to increase the efficiency of the model. The effectiveness of the model is experimentally evaluated on PHEME rumors dataset Figure 9.

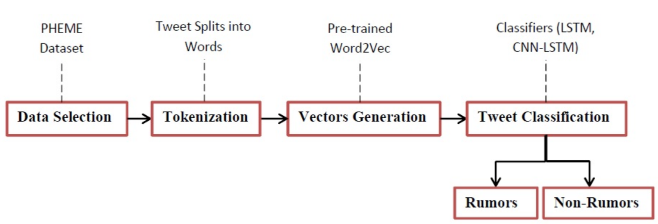

Figure 9. Rumor Deep Classification Model

A. Dataset

In this research, the PHEME dataset have been used for evaluation of effectiveness of the model. The PHEME dataset contains rumors and non-rumors tweets of 9 different events. In this research, 'ottawashooting' events of this dataset have been used for analyzed the model. The ottawashooting event dataset contains total 12284 tweets, in which 5848 tweets was non-rumors and 6436 was rumors.

B. Tokenization

The experiment is required to identifying the words that are important to generate text semantic results. The tweets have the sentences and in this step, there is deleted all the special symbols and split sentences into words. These words will be the inputs for the word embedding model.

C. Vector Generation

Many machine learning and deep learning models have been implemented using word vectors. In this analysis, there is also using pre-trained Word2Vec word embedding model for generating the vectors. Previous Machine learning and deep learning approaches have been used BOW and tf-idf word embedding techniques for generate vectors for huge amount of text. In this research, Word2Vec model is applied for generating vectors and these word vectors have been used as input for the deep classification of tweet sentences. Word2Vec model is obtained from two models that are continuous bag-of-words (CBOW) and Skip-gram, both models worked with neural network concepts [15] [24] [12] [7] [11] [13]. To deal with a similar situation or problem, there are used the pre-trained Word2Vec model instead of building a new model again. Pre-trained Word2Vec model is created by someone that is publicly available for research purpose. This model has trained on GoogleNews dataset which contains 3million words and phrases; each word is represented on 300-dimensional vectors and provides the precise relation of the words. Word2Vec is used to generate word vectors and these word vectors are the input to the deep classifiers.

D. Twitter Classification

In supervised learning, text classifiers are widely used to predict the text in relevant classes. In this research, LSTM and CNN+LSTM deep learning approaches have been used to predict the classes in the rumors and non-rumors. A deep learning model trained with a huge number of labeled data and neural network architecture. In the traditional process of neural network architecture, 2-3 hidden layers were used to detect the feature from the data but now, many hidden layers are being used to train the model for directly learning the features from the data. RNN, LSTM, CNN models of deep learning are used to detect features and classify the data [11] [17].

CNN+LSTM Deep Learning Architecture

The CNN-LSTM architecture is based on a deep learning framework Figure 10. This architecture takes input from the word embedding layer. The next layer of this architecture

TABLE I. Comparison of Different Social Media Problems

Problems	Dataset	Feature Extraction	Methodologies	Results Analysis	Authors Name & Year
Rumors detection on twitter	PHEME	Word embedding	CNN deep learning	91.0% Accuracy	Abdullah Alsaeedi, et all., 2020
Rumor identification	Theme	User verified, sentiment, followers, hashtags, length, count status, and retweets of tweets	LSTM deep learning	0.86 F1-score for no-rumors and 0.72 F1-score for rumor	Jyoti Prakash Singh, et all., 2019
Rumor identification	Zubiagaset and Kwonset	Linguistic & language features, User profile features, Metadata features	Autoencoder, Gaussian, K-Means, KNN, SVDD, OCSVM, and PCA	74.30% F1-score for Zubiagaset with KNN and 93.98% F1-score for Kwonset with KNN	Amir Ebrahimi Fard et all, 2019
Breaking news rumors detection	PHEME	Word2Vec word embedding	LSTM-RNN deep learning	0.791 F1-measure	Sarah A. Alkhodair et all, 2018
Rumors identification	Twitter & Weibo	Refine the keywords, replicate handcrafted features	tanh-RNN, LSTM & GRU deep learning	Twitter dataset: 88.1% Accuracy with the GRU2 method. Weibo dataset: 91.0% Accuracy with GRU2	Jing Ma, et all., 2016
Domestic Violation content categorization	Facebook (Extracted from Graph API)	Pre-trained Word2Vec and Glove word embedding, Domain-specific embedding	CNNs, RNNs, LSTMs, GRUs, and BLSTMs deep learning	GRUs+GloVe 91.78% Accuracy	Sudha Subramani, et. all., 2019
Automatically detection of Domestic Violation on social media	Facebook (Extracted from Graph API)	Google's Pre-trained Word2Vec and Twitter's crawl of Pre-trained GloVe	RNNs, LSTMs, BLSTMs, GRUs, and CNNs deep learning	LSTM + Word2Vec 93.08%, GRU+GloVe 94.26% Accuracy	Sudha Subramani, et. all., 2018
Detection of offensive comments	Kaggle website	N-gram, Count, TF-IDF score, Occurrence of pronouns, Skip-grams	Support vector machine (SVM), Logistic Regression of machine learning	86.92% Accuracy	Vikas S Chavan, et all., 2015
Automatic detection of inappropriate language (abuse contents)	Web search queries	Randomly initialize the DSSM word vectors	Convolutional Bi-Directional LSTM deep learning	F1 Score 0.8720	Harish Yenala et all, 2017

Continue Table I...

Problems	Dataset	Feature Extraction	Methodologies	Results Analysis	Authors Name & Year
Detecting bullying and aggression posts	Twitter data (extracted from twitter streaming API)	Word2Vec	J48, LADTree,LMT, NBTree, RF and Functional Tree of ML	90% Accuracy	Despoina Chatzakou,et all., 2017
Detection of cyber-bullying incidents	Wikipedia, Twitter, Formspring, YouTube	Word embedding (random, GloVe and SSWE)	CNN, LSTM, BLSTM and BLSTM with attention deep learning	F1-score for YouTube dataset: CNN: 0.78, LSTM: 0.14, BLSTM: 0.93, BLSTM with attention: 0.92	Maral Dadvar et all, 2018
Identification of extremist contents	Twitter (extracted from Twitter Streaming API)	Word embedding	LSTM with CNN deep learning model	92.66% Accuracy	Shakil Ahmad, et all., 2019
Automatically identification of extremist contents on social media and webpages	Manually collected from different sources	Doc2vec vectors Deep neural network classifier	93.2% Accuracy	Andrew H. Johnston, et all., 2017	
Toxic content detection	Kaggle, Subversive Kaggle, Wikipedia, Subversive Wikipedia, Reddit, Subversive Reddit	Preprocessing, labeling	SentiWordNet	Accuracy: Kaggle 93.7, Subversive Kaggle 80.1, Wikipedia 85.5, Subversive Wikipedia 82, Reddit 94.3, Subversive Reddit 83.9	Eloi Brassard-Gourdeau et all., 2018
Illicit Drug-related content identification	Extract data from NIDA website	Word embedding	LDA topic modeling	78.1% Accuracy	Tao Ding, et all., 2016
Rumor detection	Kaggle	Word embedding (GloVe)	CNN+BiLSTM	90.93% accuracy	Neetu rani et all., 2021

is the CNN feature extraction layer which consists of 1D convolutional layer, maxpooling layer and ReLu layer. 1D convolutional layer enables to learn low level features from inputs. A maxpooling layer was introduced to minimise computational effort by reducing the dimension of the feature maps by a factor of two. The vanishing gradient problem is solved using the ReLu activation function. To alleviate the over-fitting problem, a dropout1 layer is inserted between the CNN feature extraction layer and the LSTM sequence learning layer in this design. In the training process, this layer comprises a random selection of neurons and deactivating some of them. To create the final output, the sequence learning block's output is connected to a dropout2 layer, which is then followed by a dense layer.

E. Experiment Evaluation

1) Experiment Setup

In this research, ottawashooting event of PHEME dataset have been used to evaluating the model and pre-trained

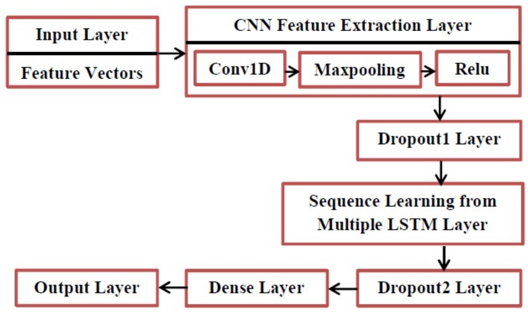

Figure 10. CNN+LSTM Layered Architecture

Word2Vec model have been used for generating word vectors. In this experiment, 90% of tweets were used to train the model and 10% of tweets were used for testing purpose. The model has analyzed on Tensorflow library which running on python 3.7 and model has been trained with Keras neural network library. The model has been trained on 8GB RAM and core i5 processor configured system.

2) Experiment Results

In this research, deep rumor tweet classification model was trained on different parameters Table II III which are extracted from different layers of the model. The results were analyzed from LSTM and CNN+LSTM deep learning models.

TABLE II. Parameters Received from Different Layers of LSTM Model

Layer (type)	Output Shape	Param #
Embedding_1 (Embedding)	(None, 30, 300)	15000000
lstm_1 (LSTM)	(None, 150)	270600
Dropout_1 (Dropout)	(None, 150)	0
Dense_1 (Dense)	(None, 1)	151

TABLE III. Parameters Received from Different Layers of CNN+LSTM Model

Layer (type)	Output Shape	Param #
Embedding_1 (Embedding)	(None, 30, 300)	15000000
Conv1d_1 (Conv1D)	(None. 30, 64)	96064
Max_pooling1d_1 (MaxPooling1)	(None, 15, 64)	0
Dropout_1 (Dropout)	(None, 15, 64)	0
lstm_1 (LSTM)	(None, 150)	129000
Dropout_2 (Dropout)	(None, 150)	0
Dense_1 (Dense)	(None, 1)	151

ottawashooting event dataset and results evaluated from LSTM and CNN+LSTM and deep learning model with the use of pre-trained fine-tuned word vectors with 300 dimensions. After number of 5 epochs loss values, model trained accuracy and execution time are calculated Table IV V.

TABLE IV. LSTM Model Accuracy

Epoch	Execution Time (s)	Loss	Model Trained Accuracy
1	107	0.4965	0.7479
2	107	0.1934	0.9309
3	114	0.0946	0.9696
4	123	0.0592	0.9806
5	118	0.0371	0.9859

TABLE V. CNN+LSTM Model Accuracy

Epoch	Execution Time (s)	Loss	Model Trained Accuracy
1	89	0.5277	0.7166
2	94	0.2211	0.9185
3	93	0.1216	0.9596
4	98	0.0798	0.9721
5	98	0.0600	0.9775

The results have calculated a confusion matrix for measuring the performance of the classification model. A confusion matrix provides the combination of predicted values and actual values Table VI. A confusion matrix helps to measuring the recall, precision, f-score, accuracy and ROC curve. In this experiment, there are measured the recall, precision, f-score and accuracy for the both LSTM and CNN-LSTM deep learning classification model Table VII VIII.

TABLE VI. Confusion Matrix of Test Results from LSTM and CNN-LSTM Models

	LSTM Actual Rumors	LSTM Actual Non-rumors	CNN+LSTM Actual Rumors	CNN+ LSTM Actual Non-rumors
Predicted Rumors	526 (TP)	54 (FP)	524 (TP)	56 (FP)
Predicted Non-rumors	62 (FN)	587 (TN)	40 (FN)	609 (TN)

TABLE VII. Results Analysis from LSTM Model

	Precision	Recall	f-score	Accuracy (%)
Rumors	0.89	0.91	0.90	90.68
Nonrumors	0.92	0.90	0.91	90.44
Average	0.91	0.91	0.91	90.56

TABLE VIII. Results Analysis from CNN+LSTM Model

	Precision	Recall	f-score	Accuracy (%)
Rumors	0.93	0.90	0.92	90.34
Nonrumors	0.92	0.94	0.93	93.83
Average	0.92	0.92	0.92	92.08

3) Results Evaluation

Experiment results Table IV V shows that the model has trained 5 times and calculated the accuracy and loss values. The LSTM model provides the 98.59% accuracy and CNN+LSTM model provides the 97.75% accuracy to train the model. After every epoch the loss value is decreasing and accuracy is increasing it means the model trained very well.

In this analysis, to solve the over-fitting problem, there is used a dropout layer. In deep learning model, the dropout layer is efficient way to solve such problem.

ROC curve diagnose the ability of classifiers. In this results, LSTM and CNN+LSTM both classifiers gives curves near to the top of left corner, it indicate the better performance of the classifiers Figure 11 12.

LSTM model provides the 90.56% testing accuracy and CNN+LSTM model provides the 92.08% testing accuracy. This result shows that CNN+LSTM hybrid model more efficient as compare to LSTM model Figure 13.

LSTM model take more execution time as comparative CNN+LSTM model to build a model Figure 14.

The precision and recall are basically used for most positive information retrieval. The precision is calculated the actual positive results by the out of total predicted positive results and recall is calculated the actual positive results by out of the total actual yes or classes results. In this analysis, CNN+LSTM model provides better ratio of positive results.

F-score is the weighted of precision and recall, if false positives and false negatives cost are very different than f-score looks better. In this analysis, there is no more difference.

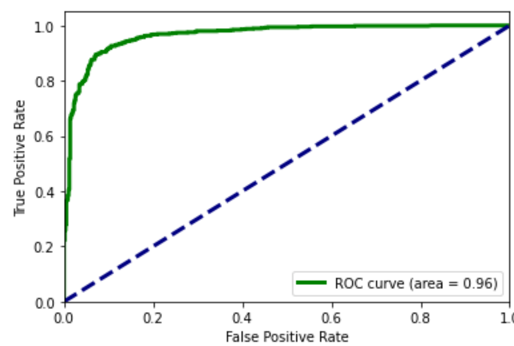

Figure 11. Performance Analysis of LSTM Model using ROC curve

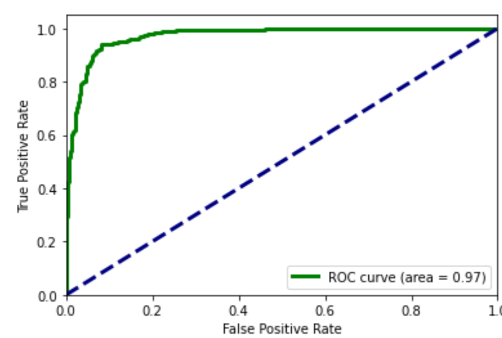

Figure 12. Performance Analysis of CNN+LSTM Model using ROC curve

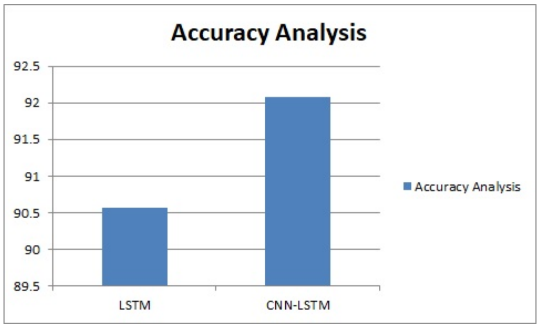

Figure 13. Accuracy analysis of LSTM and CNN+LSTM

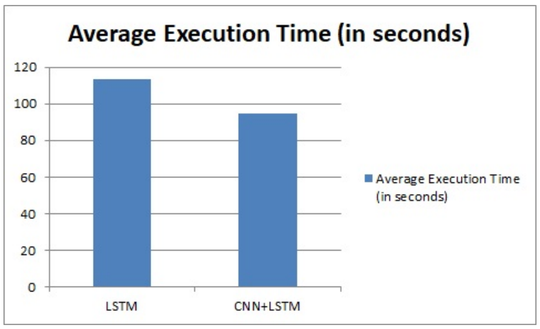

Figure 14. Build Model Execution Time

6. Analysis and Discussion

Manual identification of social media problems is a very tedious process and this process; increase labor and time cost, and degrade the result. To solve such problems, many researchers have been suggested different techniques and models for automatic identification, detection, and stopped them from spreading globally. As per the researchers, the meaning of the information is hidden in the text content and most users share the post in the form of text. For the text analysis, it is required to apply the NLP process to the text. Data has to be indifferent and diverse form, so there is also required to clean them by using common and important methods like stop word removing, case folding, steaming, removes hyperlinks, special character removing, normalization, etc. For processing the data, a vector representation of the words is required. For vector representation of the sentences, researchers suggested some important and useful methods of word vector representation that as word embedding (tf-idf, GloVe, Word2Vec). In between some researchers suggested the process of the multilingual text which converts multilingual text into a single language by using GoogleTrans API. After the word vector representation, it is required to classify the data into different labels that is why researchers have been now focusing on machine learning and deep learning concepts. In this study, there is analyzed the previous researches which were based on text analysis by using text preprocessing, NLP process, and applying machine learning and deep learning for text classification Table I. This study also suggests that how researchers are focusing on the deep learning concept on social media content categorization.

The previous researches have been worked on small size of dataset like Theme [9], Zubiagaset [21], manually created dataset [12] [7] but in this analysis, results have been calculated on large number of sentences and suggested that the model is perfectly working on large dataset and provide better accurate results. The previous research [28] has suggested a hybrid model of CNN+BiLSTM with Glove word embedding to classify the tweets into rumor and non-rumor. This model was applied on publically available dataset from Kaggle and finds 90.93% accuracy. This analysis have been used Word2Vec model for vector generation instead of Glove model and received 92.08% accuracy after 5 number of epoch. Moreover in this analysis the results have compared with LSTM and CNN-LSTM rumor classification model. This research is focused on different social media problems that are related to posted content. Social media users mostly post the information or content in the form of text or sequential sentences, existing state of art methods have limitations to preserve the semantic representation of sequential information and its process is very time consuming to build a model, this research suggested that hybrid deep learning approach have no limitation for sequential text and take less time to build a model of text analysis. The deep learning approaches are helpful to extract hidden information from the sentences or social media text because of multiple hidden layers. Mostly, a sentence has long term dependency, this analysis have been focused on the hybrid CNN+LSTM model of deep learning which provide accurate and fast results. The pre-trained model is also very helpful to generate very fast results. A deep learning network works better with word embedding techniques and it also provides better performance with different features like linguistic and sentiment features, user profile features, and user metadata features of the tweeter post. So, researchers are now more focused and inclined towards deep learning approaches for the identification of social media problems.

7. Conclusions and Future Scope

Social media platforms have become a more important part of our daily life for sharing ideas, opinions, knowledge, and news with people. It is surely a boon in many contexts, but it can be a terrible curse when rumors, offensive contents, abusing content, misinformation, wrong information, and fake news are also shared by the users. Social media spread harmful, unauthentic, unwanted content regularly; due to such type of content it has a dangerous influence on society. Many researchers are now focusing on preventing such types of content on the social media platform and for this; they are using many algorithms and methods to analyze and classify such content and to spread only relevant information. In the existing literature, lexical analysis and syntactic analysis were used for more accurate results and now, sentimental and semantic analysis of NLP helps to extract more features from the sequential sentences. Previously, researchers were more focused on classification with both rumors and non-rumors features and noticed that the classified results detect only from the rumored features, so binary classification might not provide beneficial results. To solve this, a new approach is used i.e., one class classification classifier with one class feature. Rumor features were extracted from the already available and detected features of rumors on a social network. The researchers used the concept of machine learning and deep learning for determining social media platform problems and also compared these two technologies. The results show the effectiveness of the deep learning concept over the machine learning concept. Also, machine learning concepts are very time-consuming approach as compared to deep learning approaches. Deep learning concepts help to solve such issues and are useful with the semantic representation of the sentences, its hidden layer learning concept provides the automatically learn from the text with high performance. Currently, deep learning technique like CNN and LSTM has become more popular for text classification. Natural language processing and word embedding methods play the most important role and help to provide better accurate results. Pre-trained word embedding model is used for generating efficient word vectors which help to the low processing time as well as provide efficient results. The GoogleTrans API is also used for classifying multilingual content. Twitter streaming API, Facebook Graph API, BeautifulSoup API is used for streaming content from social media platforms. The Word2Vec & GloVe model of word embedding is used for representing word vectors in a sentence; these vectors are used as input for deep learning models to provide better accurate results. The comparative study and LSTM deep learning model's evaluation proves that deep learning models are the best for text categorization. In the future hybrid deep learning models will be effective for social media content categorization.

References

[1] V. S. Chavan and S. Shylaja, "Machine learning approach for detection of cyber-aggressive comments by peers on social media network," in *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE, 2015, pp. 2354–2358.

[2] S. H. Yadav and P. M. Manwatkar, "An approach for offensive text detection and prevention in social networks," in *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*. IEEE, 2015, pp. 1–4.

[3] C. Buntain and J. Golbeck, "Automatically identifying fake news in popular twitter threads," in *2017 IEEE International Conference on Smart Cloud (SmartCloud)*. IEEE, 2017, pp. 208–215.

[4] J. Ma, W. Gao, P. Mitra, S. Kwon, B. J. Jansen, K.-F. Wong, and M. Cha, "Detecting rumors from microblogs with recurrent neural networks," *ijcai*, 2016.

[5] Y. A. Putra and M. L. Khodra, "Deep learning and distributional semantic model for indonesian tweet categorization," in *2016 International Conference on Data and Software Engineering (ICoDSE)*. IEEE, 2016, pp. 1–6.

[6] D. Kosmajac and V. Keselj, "Language identification in multilingual, short and noisy texts using common n-grams," in *2017 IEEE International Conference on Big Data (Big Data)*. IEEE, 2017, pp. 2752–2759.

[7] A. H. Johnston and G. M. Weiss, "Identifying sunni extremist propaganda with deep learning," in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, 2017, pp. 1–6.

[8] M. Kanakaraj and R. M. R. Guddeti, "Nlp based sentiment analysis on twitter data using ensemble classifiers," in *2015 3Rd international conference on signal processing, communication and networking (ICSCN)*. IEEE, 2015, pp. 1–5.

[9] J. P. Singh, N. P. Rana, and Y. K. Dwivedi, "Rumour veracity estimation with deep learning for twitter," in *International Working Conference on Transfer and Di usion of IT*. Springer, 2019, pp. 351–363.

[10] D. Chatzakou, N. Kourtellis, J. Blackburn, E. De Cristofaro, G. Stringhini, and A. Vakali, "Mean birds: Detecting aggression and bullying on twitter," in *Proceedings of the 2017 ACM on web science conference*, 2017, pp. 13–22.

[11] A. H. Ombabi, O. Lazzez, W. Ouarda, and A. M. Alimi, "Deep learning framework based on word2vec and cnnfor users interests classification," in *2017 Sudan conference on computer science and information technology (SCCSIT)*. IEEE, 2017, pp. 1–7.

[12] S. Ahmad, M. Z. Asghar, F. M. Alotaibi, and I. Awan, "Detection and classification of social media-based extremist affiliations using sentiment analysis techniques," *Human-centric Computing and Information Sciences*, vol. 9, no. 1, pp. 1–23, 2019.

[13] M. V. G. Aziz, A. S. Prihatmanto, D. Henriyan, and R. Wijaya, "Design and implementation of natural language processing with syntax and semantic analysis for extract traffic conditions from social media data," in *2015 5th IEEE International Conference on System Engineering and Technology (ICSET)*. IEEE, 2015, pp. 43–48.

[14] O. M. C. Rendon, F. Estrada-Solano, and L. Z. Granville, "A mashup-based approach for virtual sdn management," in *2013 IEEE 37th Annual Computer Software and Applications Conference*. IEEE, 2013, pp. 143–152.

[15] S. Subramani, S. Michalska, H. Wang, J. Du, Y. Zhang, and H. Shakeel, "Deep learning for multi-class identification from domestic violence online posts," *IEEE Access*, vol. 7, pp. 46 210–46 224, 2019.

[16] S. Subramani, H. Wang, H. Q. Vu, and G. Li, "Domestic violence crisis identification from facebook posts based on deep learning," *IEEE access*, vol. 6, pp. 54 075–54 085, 2018.

[17] C. Baydogan and B. Alatas, "Sentiment analysis using konstanz information miner in social networks," in *2018 6th International Symposium on Digital Forensic and Security (ISDFS)*. IEEE, 2018, pp. 1–5.

[18] F. T. Hossain, M. I. Hossain, and S. Nawshin, "Machine learning based class level prediction of restaurant reviews," in *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*. IEEE, 2017, pp. 420–423.

[19] É. Brassard-Gourdeau and R. Khoury, "Impact of sentiment de-