An Effective Sentiment Analysis of Social Media Data Using Deep Recurrent Neural Network Models

Vidyabharathi.D[a]*, Marimuthu.M[a], Theetchenya.S[a],Vidhya.G[a], Basker.N[a], Mohanraj.G[a], Dhaynithi.J[a]

[a]Sona College of Technology, Department of Computer Science and Engineering, Salem, Tamil Nadu

*dvbharathi77@gmail.com, mari.btech@gmail.com, theetchenya@gmail.com,gvidhyacse@gmail.com, bas2k9@gmail.com , mohanraj134@gmail.com, dhaya.j@gmail.com

Abstract: Today we can do a lot of analytics and statistics using social media data. The world had been exposed to the COVID-19 pandemic recently. With the rapid increase in epidemics and deaths, people have different feelings about the disease. Collecting and reading boring tweets creates real emotions in these difficult times. The purpose of this study is to provide a specific perspective for understanding the emotions people around the world experience in relation to this condition. To do this, various Tweets related to a specific domain are received through the Twitter platform. When Tweets are collected, they are categorized and read to effectively reflect the Tweet's true sentiment for COVID-19.Natural Language Processing(NLP) is used to analyze the review comments from social media and Recurrent Neural Network (RNN) for sentiment classification .The model identifies the emotional polarity of "conversations" often work with ambiguous tweets, reduced accuracy, and reduced tolerances. Also the RNN model defines the expression of emotions occurring in a subject at any given time.

Keywords: RNN, sentiment analysis, Twitter data.

1. Introduction

Sentiment analysis is an analytical method that captures and analyses the emotions that people directly or indirectly express when expressing themselves on social media platforms, so it has attracted great interest to researchers over time. It got better and better over time [1]. Internet connections from all over the world allow many people around the world to access the social media. Therefore, information generation through social media such as Yahoo and Twitter is increasing.

Why Is Twitter Sentiment Analysis important?

Let's say you just announced a new product and discovered that Twitter shows a different name. Are you tweeting more because your customers are happy with the new features? Or are you really complaining about this situation? It takes time to manually scroll through each of these reviews. You will lose valuable feedback that can instantly improve your customer experience with the latest features (bug issues and user experience).

The overall benefits of Twitter sentiment analysis include:

- Scalability: Automate the business by analysing hundreds or thousands of tweets about a particular brand. As the data grows, can easily extend the sentiment analysis tool to gain valuable insight into our process.
- Real-TimeAnalysis: Twitter sentiment analysis is important for tracking the customers' sudden moods, seeing if complaints are expanding, and taking action before the problem expands. Sentiment analysis is actually a great way to follow a brand on Twitter and get actionable information.
- Consistent Criteria: Avoid inconsistencies due to human error. Customer representatives may not always agree on which metrics to use for each benchmark, which can lead to inaccurate results. Instead, machine learning models use a set of rules to analyze sentiment and ensure that all Twitter data is continuously reported.

With so much data being generated every day, data scientists start extracting useful data from data sets, eventually allowing social media platforms to explore the different perspectives of Internet users. [2] One of the most common topics for classifying invisible Tweets into different categories is the sentiment analysis of Tweets [3,9]. Twitter has become a useful option for sentiment analysis researchers. Despite its impact on the website, Twitter has a downside to the knowledge it generates in the form of tweets. By few years, tweets can be up to 140 characters long and 280 characters long. The new ban applies to almost all languages supported by Twitter. However, in Korean, Chinese, and Japanese, fewer words can get more complex. In most cases, tweets are about unofficial descriptions, including slang, emoticons, acronyms, incomplete sentences, and more. In order to get the most out of the data, experts analyzed the sentiment in detail through multiple Tweets. [5]

COVID-19 is a highly contagious disease caused by the coronavirus. The effects of COVID-19 range from mild to severe respiratory disease. Older people and people with chronic conditions are more likely to get the disease later. The next task is to identify COVID-19 related tweets using various machine learning algorithms. The classification algorithm used in the following article uses a sequential regression engine that predicts naming at the emotional level [6, 7].Some researchers conclude that the regression problem of machine learning algorithms on Twitter data is superior to the problem of sentiment analysis [8]. Additionally, the following approach has done a lot of work to improve the results. The current purpose of this article is to analyze the impact of coronavirus on people's brains by investigating the feelings of expressing opinions on social media sites like Twitter. #Corona and #COVID19 tweets were manually recorded using the Twitter API. This approach mainly involves the following steps: tweet aggregation, tweet preprocessing, word bag generation, a scoring method for predicting tweet polarity, and finally a recurrent neural network (RNN) that generates tweets in the field. This approach mainly involves the following steps: tweet counting, tweet preprocessing, keyword generation, scoring methods for predicting tweet polarity, and finally comparing different machine learning methods to classify tweets in various domains.

1.1 Different ways of sentiment analysis

Many diverse ways are there to process and understand the emotions in the natural language and understand emotions. Dividing the cases into two groups, the first type of traditional dictionary, contrary to the possibilities of deep learning, is not a new approach. This study also uses recurrent neural networks. In the Traditional approach, dictionary has predefined words, and every word has a meaning, whether it represents either positive or negative polarity. It then analyses the sentences, defines each word, and assigns a specific meaning to the sentences based on the dictionary. The overall sum indicates the emotional importance of a particular sentence. Of course, here we solve a few problems like denials, word turns, and word combinations that have unfamiliar feelings. So the subject is changed to intensive education using a fully trained model.

Various Types of Sentimental Analysis

It is essential to have an understanding of the types of sentiment analysis.From this ,the method which is suited to our requirement is selected.

- Fine-grained sentiment

This analysis informs the feedback received from customers. Precise results can be obtained in terms of the polarity of the input. However, the process of understanding this issue can be more time consuming and costly than other types.

- Emotion Detection

It is a sophisticated way of recognizing emotions in text. Vocabulary and ML are used to identify emotions. A dictionary is a list of positive or negative words. This way, you can easily separate words according to your needs. This also has the advantage of allowing companies to understand why their customers feel special. This is more algorithmic and harder to understand than before.

- Aspect-based

Analysing this impact affects the aspect of a product. As a sample, if a television retailer uses this type of concept, it can report on a particular aspect of television - such as brightness, sound, etc. This helps them understand what customers think about some of the features of any product.

- Intent analysis

The customer's intention are given significant importance and concentrated. The company can also predict the customer intentions about the usage of the product. So it is possible to identify and model the intent of a specific customer and use it for marketing or advertising purposes.

Different methods of rule-based and automated methods are used for this diverse kind of sentiment analysis. Rule-based analysis is effective. In contrast, automated sentiment analysis is much profound than rulebased.. ML algorithms are used to understand each customer's reaction. So it is more accurate and flexible here. To understand the public opinion about a product or business Sentiment Analysis(SA) plays an important role. However, there are unique challenges and limitations that can be overcome when used effectively. Annotative text can be difficult to understand, especially in the case of irony and irony, and some algorithms may be complex and may not provide practical results. But analyzing your sentiment is a great way to get neutral customer feedback on a lot of things. You can support your business in a variety of ways, including marketing, advertising, and market research.

1.2 Deep learning – RNN

NN is deliberately triggered many number of times. A portion of each process is passed on to the next process. Specifically, in the next process, the hidden layer of the previous process provides multiple inputs to the same hidden layer. Recursive neural networks are very useful for sequence estimation. Thus, hidden layers can be learned from previous neural network implementations in the first half of the diagram. The values obtained from the hidden layer during the next pass are part of the input for the identical hidden layer. Likewise, the value obtained in the second run of the hidden level is part of the feedback for the third round of the same hidden level. In this way, iterative neural networks gradually convey the value of the entire series rather than individual values.

The advantage is that in addition to RNN, you can process goods at any time. The input size does not grow in size. The quote takes into account the details of the story. Weights are distributed over time. These general objections can usually be slow to compute.

Next the various types of RNN architectures are discussed.

1. Fully Recurrent Neural Network (FRNN)

FRNN was founded in the earlier 80s.It can learn the temporal sequences, either in batch mode or online. FRNN has two levels of inputs and outputs, linear units and nonlinear units, respectively. Entry-level units with adjustable weight are fully connected to the unit at the emission level. Each device has real-time variables and controls. Rejected devices are notified before they become active. This input level wakes up the device. Learning with FRNN means defining an input sequence and activating a different set of output sequences. It provides the input order of the annotations, the sorting of outputs in different time steps, and an abstract view over time.

2. Recursive Neural network

Networks are built into a graphical structure that can be determined statically using the same set of weights within the network. This network is trained in reverse order using automatic differentiation [2]. The linear sequence structure is followed to handle distributed structure representations. One form of RNN is an recursive tensorneural network that uses complex functions of tensors at each node in the network.

3. Hopfield Network

John Hopfield created the network with all connections is identical .The Hopfield network is a series of interconnected neurons that simultaneously update their activation functions of other neurons independently. As soon as a new input is applied, the output is calculated and the input is adjusted according to the input at the feedback level. This process continues until the output stabilizes. You can train the network using either the Hebbian training database or the stork training database. This is used for Addressable Content Memory (CAM).

4. Elman networks

The Elman Network is a three-level grid with additional context blocks. The levels are input layer, middle hidden layer which is connected to the content block and output layer. At every given moment, the input unit is combined with the learning unit and fed-forward. Context units from hidden layer serve as back-connections to the SRN network this allows preceding hidden unit values to be saved. The context units are fed-back from the output layer in Jordan network, but in Elman network fed-back is from hidden layer. It is also called as Simple Recurrent Network (SRN) orJordan networks.

5. Echo State Network (ESN)

The ESN [5] is a network with very unusual hidden layer connectivity. Hidden neurons with fixed connectivity and weights are allocated at random. The weights of output neurons can be taught, allowing the network to (re)produce complex time-based patterns. Hidden-to-output neurons bind the weights that are changed through training. A reservoir is an optimized C++ library for various Echo state networks, and a Liquid state machine is a related idea.

6. Neural History Compressor

It's a heap of unsupervised RNNs [6]. The next layer's input is estimated based on the previous layer's input. When the inputs cannot be estimated, they are passed on to the next higher stage as feedback, with more hidden units attached. As a consequence, unpredicted inputs and compact knowledge representations make up each higher stage of the RNN. A higher level chunker and a lower level automatizer may be thought of as a two-level network. The vanishing gradient problem of automatic differentiation or back propagation in neural networks can be partially solved

7. Long Short Term Memory (LSTM)

It is a deep learning model that can learn a task while avoiding the vanishing gradient problem [7]. Recurrent gates known as "forget" gates are often used to boost LSTM, and learning exercises necessitates recollection of past events.

8. Gated Recurrent Units

Kyunghyun Cho [8] implemented the gating function in RNN (2014). Unlike the LSTM, this mechanism does not have an output gate and has less parameter. On polyphonic music and speech signal modelling, it performs similarly to LSTM

9. Bi-directional RNN

Each member of a finite sequence is predicted by a bi-directional recurrent neural network based on its previous and next condition. It will process sequences from left to right and right to left. Then their performance in concatenated in both directions. When used in conjunction with LSTM [9], this approach is extremely useful.

10. Continuous Time RNN

It models the effects of an incoming spike train on a neuron using ordinary differential equations on a device. Cooperation, perception, and limited cognitive actions are all addressed by CTRNNs in evolutionary robotics [10]. The Shannon sampling theorem is a CTRNN.

11. Hierarchical RNN

The HRNN works in a top down hierarchy to generate responses for a network. One part produces lower network primitives, such as stop-to-down, while the other retains top-down sequencing's global targets.

12. Recurrent Multilayer Perceptron(RMLP)

RMLP consists of fully connected, cascaded, feed forward sub networks [12]. Each subnetwork has multiple layers of nodes. Several layers of this network are interconnected in the feed forward manner. No return connections are allowed except for the last layer that can provide access connections between nodes. Nonlinear dynamic processes are defined and constructed using a recurrent multilayer perceptron (RMLP). This network architecture combines the temporal behaviour of multilayer perceptron with known properties.

13. Multiple Timescales Model

The functional hierarchy in the brain: neural system is simulated using a several timescale RNN model. A functional hierarchy is one in which complex features are broken down into simple features and simple are merged to form a complex one. Functional hierarchy can be divided into two: space hierarchyand time hierarchy. This model's functional hierarchy will self-organize by using several timescales of neural activity [13]. Author [17,18,19] conducted additional studies in information processing of classical music.

14. Neural Turing Machines (NTM)

Paradigm combines fuzzy pattern matching with the processing power of programmable processors. Itis a neural network that has been expanded to provide a controller that communicates with external memory services. It's similar to a Turing Machine with the exception that it can be trained fast using gradient descent. Easy algorithms like sorting, copying, and associative recall can be completed with NTM and LSTM.

15. Differentiable Neural Computer

The DNC is published in 2016 [15] using an auto associative memory model. It is an extension of NTM, with more memory and temporal focus, as well as being abstract and more stable than NTM. DNC has a various subcomponent which is similar to Von-Neuman architecture with gradient descent instruction.

16. Neural Network Push Down Automata (NNPDA)

NN pushdown automata are networks with CFG(Context Free Grammars). It is similar to NTM, but instead of tape, it uses identical stacks. NNPDA aims to provide a precise statistical explanation as well as a theoretical study of relevant issues.

2. Related Works

Citizens all over the world are suffering greatly as a result of the Covid-19 pandemic outbreak, which has a greater impact on mental wellbeing than physical illnesses. Scientists from all over the world are interested in learning more about the causes, symptoms, and issues that people face on a daily basis. Many problems are raised as a result of fear and false information spread by individuals through social media blogs. Few scholars have looked at the same issue and, based on their findings, hypothesised some solutions. Abhilash et al. studied a

common social blog called the Twitter platform in 2019 and used nostalgic analysis to explore its awareness. The outstanding characteristics of the proposed system of weighted correlated effect (WCI) training were derived from tweets linked to the Covid-19 epidemic. The suggested WCI was used to calculate each person's effect score in relation to the Covid-19 pandemic. Characteristics are derived from the perspective of profile data and plotted as a relationship graph in standard structured data[3]. The disadvantage of the suggested solution is that only three features were used for tweets, retweets, and mentions, resulting in a deficiency. Jim Samuel et al. developed Naive Bayes and logistic regression classification algorithms in 2020 to test coronavirus-related fear-sentiment tweets. The previous method outperformed the regression test when it came to tiny tweets.

The suggested classifiers are trained using both exogenous and endogenous textual data. Word count, keywords, hash tags, mentions, and special characters are examples of endogenous features, while exogenous information about source devices with details of their location information is an example of exogenous information. Naive Bayes only achieved 52 percent accuracy for large-size tweets, which can be increased by other effective learning algorithms[4]. Catherine Ordun et al. used three different methods to explain dishonest tweets by coronavirus in 2020: Uniform Multiple Approximation and Projection, Digraph, and Subject Modelling. The major facets of Covid-19 are analysed using the Latent Dirichlet Allocation (LDA) technique as a keyword study, and the topics are merged to visualise the retweeted data with its time-bound nature. According to the report, it took at least 2.87 hours for government employees to retweet about the coronavirus[5].

In 2019, Márcio et al investigated various machine learning models for classification and prediction problems. Using sentimental analysis to classify text data with the aid of a vector machine, decision tree, random forest, and the Bayes theorem[6]. To recognise sentimental feelings, these algorithms are qualified and validated on social media. Kumar Jain et al. developed Naive Bayes in 2019, which supports vector classifiers for detecting individual emotions using multilingual text data[7]. K-Nearest Neighbor (KNN)[8] is the most general algorithm for classifying text data using similarity calculations. The relation between the two data points is measured by estimating their distance and proximity. The KNN uses a clear majority vote to decide the classification of each data point's closest neighbour. For each object, the number of neighbours closest to (K) must be determined either by definition or by an approximation of the number of neighbours within a given radius[9].

The new strategies of deep learning(CNN, RNN, and LSTM) were verified and related in papers[14,15,16] in the sentiment analysis difficulties. By combining subject modelling and the effects of a sentiment analysis is performed on customer specified social media data, Jeong et al.[19] discovered opportunities for product creation. The tracking tool is used to determine the market changing demand in emerging product environment. Pham et al. used layers of knowledge representation to examine travel related reviews and assess feelings for the factors: importance, room, location, cleanliness, and service[11]. The other method[17] combines sentiment and semantic features in the LSTM model focused in emotion detection. In [14,15,16], the deep learning strategies, CNN, RNN, and LSTM, were tested individually on different datasets. However, there hasn't been a comparative analysis of these three methods. Most researchers use the same approach for emotion analysis. The Word2vec tool[18] is used to automatically extract text characteristics from different data sources and then pass them to word embedding. Extensive literature in the proposed systems technology area has also focused on sentiment analysis. The majority of the approaches in this area rely on filtering the information and can be divided into the following categories: content-based, collective filtering (CF), demographic, and hybrid. Social data are used in various ways by using these strategies.

Content-based methods rely on item and user profile characteristics, CF methods on implicit or explicit user expectations, demographic methods on user demographic details, and composite methods on several item and user data that can be extracted or inferred from social networking.In addition, when dealing with both explicit (data provided directly by users) and implicit data, hybrid methods and lifelong learning algorithms are considered as in-depth approaches for recommendation systems (which are inferred from the behaviour and actions of users).Shoham[20] suggested one of the first hybrid recommendation systems, which combines content and shared screening recommendation methods. The proposal's content-driven section involves the development of user profiles based on their web-based subject interests, while the system's collaborative filtering section is based on feedback from other users. While sentiment analysis is not used in this study, it can be seen as a guide to other research that combine both approaches and use sentiment analysis to elicit implicit feedback from consumers. Wang et al.[21] propose a hybrid technique in which sentiment of movie feedback is used to create a uncertain recommendation list resulting from a mixture of shared filtering and content-based approaches. After collaborative filtering in the same application area, Singh et al. recommend using an emotion classifier triggered by movie feedback as a second philtre[22].

3. NATURAL LANGUAGE PROCESSING (NLP)

Natural Language Processing (NLP) is a problem-solving technique that helps you to analyse text. It is used to assist users in accessing and ingesting large amounts of text-based content that already exists. Using NLP, statistics, or ML tools, retrieve, or define, the sentiment of a text part.

3.1 Natural Language Toolkit (NLK)

NLK is a leading platform for developing Python programmes using human language technologies. In sentiment analysis, processes such as POS tagging, parsed tree view, stemming, and named entity recognition are used.

3.2 Tokenizing the Data

The process of making the text easier for the machine to comprehend. This is a technique for breaking down strings into tokens. A token is a text string of characters that functions as a single entity. Depending on how the tokens are designed, they can be made up of sentences, emoticons, hashtags, connections, or even single characters. Splitting text into tokens on the basis of whitespace and punctuation is a basic method of doing so.

3.3 Normalizing the Data

"Ate," "eats," and "eating," for example, are all separate variations of the same word "eat." Words come in a variety of shapes and sizes. Depending on the needs of your research, each of these versions will need to be converted to the same form, "eat." In NLP, normalization is the process of transforming a term to its canonical form.

The term "normalization" refers to the aggregation of words with similar meanings. "Food," "eat," and "eat" can all be viewed as separate words without normalisation, even though you want them to be treated as separate words.

3.4 Removing Noise from the Data

The process of removing noise from data entails removing meaningless terms from the text section. Words like "is," "the," and "a" should not be removed. They are usually useless when processing language unless they are required for a special use case.

3.5 Determining word Density

The most fundamental type of research is determining the word frequency from textual evidence. Since a single tweet is too limited to determine the frequency of words, the measurement of the frequency of words will be performed on all positive tweets.

3.6 Preparing Data for the Model

To determine the author's attitude toward a subject that is being written about, a supervised machine learning algorithm is used. And a training data set is generated, with each dataset having a "sentiment" for training. Text can be classified into a number of emotions using emotion analysis. This tutorial lets you train your model in just three categories: positive, negative, and neutral, for consistency and availability of the training dataset.

3.7 Building and Testing the Model

The percentage of tweets in the research dataset for which the model correctly predicted sentiment during this constructs and test phase. It's then calculated whether the mood is positive, negative, or neutral based on the percentage of the sentiment.
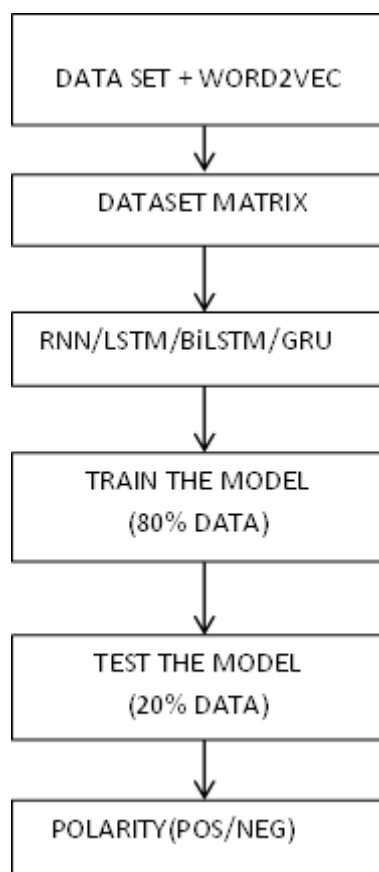
4.METHODOLOGY

The tweets regarding the COVID 19 is analysed to predict the actual sentiment .The correctness of the tweets has to be checked. In this pandemic, the tweet produces greater impact among the public regarding this situation. To get the actual and better feedback, a model has to develop and trained to classify the tweets suitably. When we consider the text classification, the RNN gives good performance. Here we consider some of the familiar architectures RNN, LSTM, BiLSTM and GRU for experimental purpose. The general framework for the recurrent neural network model is shown in figure.1.

The model concentrates on the neutral content reduction which says that there should not be any polarity unidentified tweets from the dataset. The latest findings for the coronavirus subject (up-to-date topic) have a lot of scope to analyse and compare the model, since there is no pre-compiled dataset. The improvements is shown, when fresh data is more considered. We compare accuracy of each model with traditional model and what assumptions can be drawn aboutperformance, precision, and speed in particular cases.

With latest online messages formulas, complicated phrases, and specific subjects, we expect the model studied, created, and taught to produce more accurate results than a conventional or third-party system, which, though successful, has larger error ranges than our more precisely prepared model.

Figure.1 General Framework for the Model/Classifier

```
┌─────────────────────────┐
│  DATA SET + WORD2VEC     │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│    DATASET MATRIX        │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│  RNN/LSTM/BiLSTM/GRU     │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│   TRAIN THE MODEL        │
│     (80% DATA)           │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│    TEST THE MODEL        │
│     (20% DATA)           │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│   POLARITY(POS/NEG)      │
└─────────────────────────┘
```

The model was generated with the help of Keras and Tensorflow software. A sequential model is created by forwarding the instances to the embedding layer, which can be used on text data in RNN. This entails integer encoding of the input data, in which each expression is assigned a unique integer. As the embedding layer is modified with random weights, the research dataset will learn embeddings for theentire terms. Next is denselayer and dropout layert. A dense layer is a conventional neural network layer that connects every input node to every output node. The activations for those arbitrary nodes are set to zero when the dropout layer is used. In this case, overfitting can be stopped.The matplotlib. pyplot module is used for visualization, and our model uses the prediction mechanism to walk through the described dataset. We then classify whether a tweet or expression is positive or negative, and visualize the results.

Table.1. Accuracy and Loss for the Models

Model	Accuracy(in %)	Loss
RNN	94.12	0.2547
LSTM	95.89	0.1891
BiLSTM	97.01	0.1594
GRU	92.74	0.3781

The classification accuracy for the models RNN, LSTM, BiLSTM and GRU give good results for the covid19 tweets. The loss value is also minimum. The table1 shows the accuracy and loss values for each of the model.The outcomes are accurate, but it may be outdated by the time the analysis is completed. The result will no longer be important. Scraping, as a test dataset, can thus be a massivestage forward. As a result, the solution takes less time, and we can even use third-party to speed up the operations.We will find incomparably wide transformationsamong traditionally sustained experiments and tests stayed by different scraping and further dataset compilation selections as the discrepancy in time and precision is discovered. It should also be recalled that, in contrast to conventional testing methods, the twitter polarity accuracywould be excellent, but it would not include such a large sample, i.e. not in such a minimum time.

4.1 Model comparison with third-party sentiment analyser - TextBlob

TextBlob and our own well-trained model have been able to accurately philtre out these terms. The RNN may seem to be more significant, but we can't prove it yet; nevertheless, the RNN doesn't always have a neutral part, giving us more analysis enhancement. The quantity of data will be the key determining factor(Here we consider only the RNN architecture).

The model was trained and tested using an imdb dataset. The fresh scratched dataset is then used for testing purpose. Later it is inferred that the classification of the model is based on the limited details. The model considered gives better result in both the highly positive and weakly negative bits, indicating a simple measure of subject division and the abundance of interactions on the issue. Positive manifestations do, of course, continue to prevail on social networking platforms, and are often influenced by partial results, but it is also fair to expect pessimistic and varied views. TextBlob always deviates in the positive way, with all effects tilted in the same direction, but in this situation, in addition to the negative implications, there is a larger neutral sense.

The TextBlob vs. RNN model comparison demonstrates how well it does on larger and larger research datasets, as well as how effective it is in delivering less incorrect results. The growing gap between the two extremes can be seen in a study of 50 people. Unlike textBlob, where the neutral unit is more relevant, tweets in the RNN (Figure 2) model were not divided into marginally negative and marginally positive groups, but rather into marginally negative and marginally positive groups.

Figure.2 Sample of 50 tweets analyzed by using TextBlob and RNN
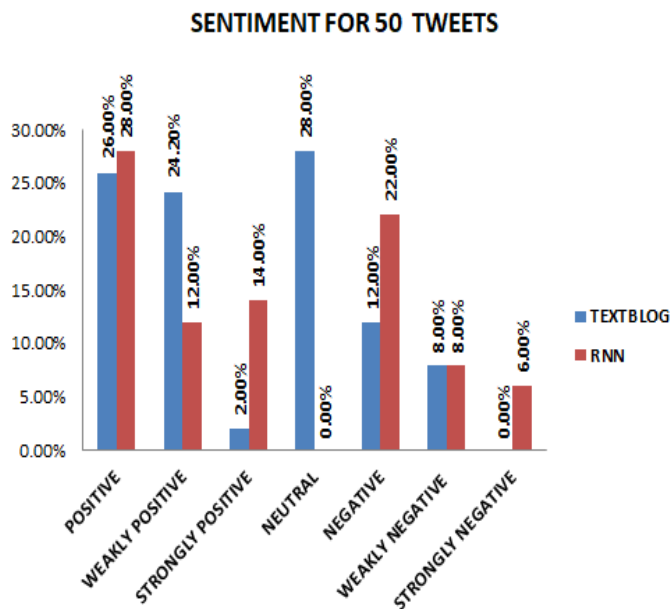
By increasing the number of tweets to 200 and retaining the word 'covid', it is noted that the division remains the same. While a positive trend can be detected, the rise in the number of tweets for both models can be explained in this situation.

The increased number of tweets shows that 'positive speech, motivation, and optimism' are mainly the 'higher degree of positivity in social media, which was anticipated, but there are still large levels of negative messages on the topic in figure 3. The effect is quite close to the trends shown previously by using a bigger dataset and the key 'covid.' Only small variations can be found in Figure 3 by looking at the positive and negative columns. We were

able to see differences of opinion quite clearly because neutral tweets were not included. RNN models assess tweets on a scale of 0 to 1, while textBlob evaluates tweets on a scale of -1 to 1.

When the number of tweets is increased to 500 in same period of time. The reinforcing of the negative component (simple negative not paired with strong and weak negative) can be seen in the RNN model, and the same can be said for the TextBlob result.

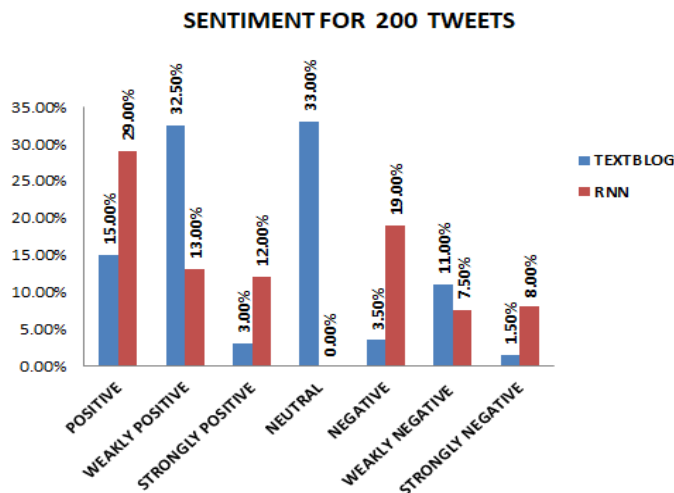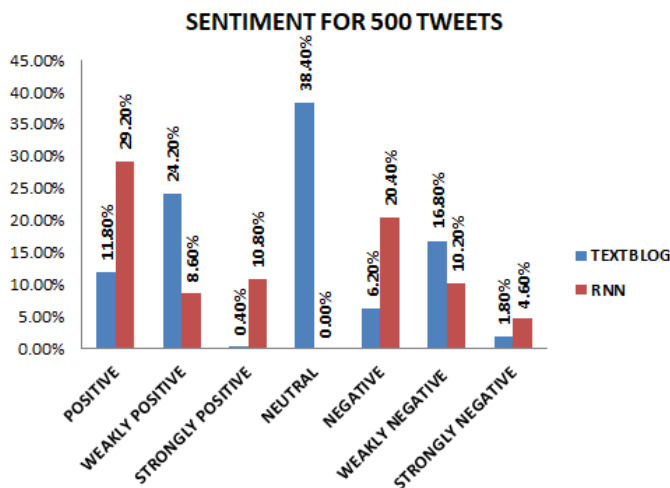**Figure.3** Sample of 200 tweets analysed by using TextBlob and RNN

Figure.4 Sample of 500 tweets analyzed by using TextBlob and RNN

Overall, comparing the categorical values of the two experiments can be used to interpret the positive displacement, but the two models express the separation of this end result in slightly different ways. A positive value of 29.20% should be used in the RNN, as opposed to a negative value of 20.40%, which is a balanced split, and the positive index in the set of 500 tweets is slightly high. According to a TextBlob survey, 24.20 % of people have a weakly positive worth.

The positive percentage is 11%, while the negative percentage is 4%. Responses and assessments of unrelated political commentaries and judgments provoke substantial behaviour from those who debate and analyse the social media implications after the declaration. This significantly increases the number of tweets about the subject. Various international events on these topics have elicited a similar response, particularly when the details have been described. There is a visible distinction between the positivecolumns and negative columns, mostly from neutral, although there are some variations in the circulation between the positive and negative columns.

5. CONCLUSION

Various architecture of RNN are considered to classify the emotions from the tweets with the keyword 'covid'.The tweets are closely examined for the prediction. Instead of using binary positive limits and negative limits, the texts emotion were divided into weakly positive/negative and overwhelmingly positive/negative

categories. It demonstrates that, even with sparse data, the RNN provide better prediction in text classification. The majority of the calculations were made against TextBlob, which worked consistently.

References

1. G. Gautam and D. Yadav, "Sentiment analysis of twitter data using machine learning approaches and emantic analysis",in 7 th Int. Conf. on Contemporary Computing, 2014, pp. 437-442
2. Abhilash Mittal, Sanjay Patidar. Sentiment Analysis on Twitter Data: A Survey: Proceedings of the 7th International Conference on Computer and Communications Management July 2019 pp.9 https://doi.org/10.1145/3348445.3348466
3. Somya Jain, AdwitiyaSinha. Identification of influential users on Twitter: A novel weighted correlated influence measure for Covid-19. Chaos, Solitons and Fractals, 2020
4. Jim Samuel,Nawaz Ali, MokhlesurRahman, Esawi, Yana Samuel. COVID-19 Public Sentiment Insights and
5. Machine Learning for Tweets Classification, MDPI, Information Journal, 2020
6. Catherine Ordun, Sanjay Purushotham, Edward Raff. Exploratory Analysis of Covid-19 Tweets using Topic Modeling, UMAP, and DiGraphs.ArXiv: 2005.03082v1 [cs.SI] 6 May 2020
7. MárcioGuia, Rodrigo Rocha Silva, Jorge Bernardino. Comparison of Naïve Bayes, Support Vector Machine, Decision Trees and Random Forest on Sentiment Analysis. Conference: 11th International Conference on Knowledge Discovery and Information Retrieval, DOI: 10.5220/ 0008364105250531
8. Dmitry Davidov, Ari Rappoport." Enhanced Sentiment Learning Using Twitter Hashtags and Smileys". Coling 2010: Poster Volume pages 241{249, Beijing, August 2010
9. Po-Wei Liang, Bi-Ru Dai, "Opinion Mining on Social MediaData", IEEE 14th International Conference on Mobile Data Management,Milan, Italy, June 3 - 6, 2013, pp 91-96, ISBN: 978-1-494673-6068-5, http://doi.ieeecomputersociety.org/ 10.1109/MDM.2013.
10. Pablo Gamallo, Marcos Garcia, "Citius: A Naive-Bayes Strategyfor Sentiment Analysis on English Tweets", 8th InternationalWorkshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland,Aug 23-24 2014, pp 171-175.
11. Has¸imSak, Andrew Senior, Franc¸oiseBeaufays. Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling.Google,USA {hasim,and rewsenior,fsb@google.com}
12. T. Mikolov, M. Karafiat, L. Burget, J. Cernock y, and S.Khudanpur "Recurrent neural network based language model," in Proceedings of INTERSPEECH, vol. 2010, no.9. International Speech Communication Association, 2010,pp. 1045–1048.
13. M. Sundermeyer, R. Schluter, and H. Ney, "Lstm neural networks for language modeling." in INTERSPEECH,2012, pp. 194–197.
14. RabindraLamsal,"School of Computer and Systems Sciences, JNU, Twitter Database.2020", DoI number:10.21227/781w-ef42.
15. Ain, Q.T.; Ali, M.; Riaz, A.; Noureen, A.; Kamran, M.; Hayat, B.; Rehman, A. Sentiment analysis using deep learning techniques: A review. Int. J. Adv. Comput. Sci. Appl. 2017, 8, 424.
16. Singhal, P.; Bhattacharyya, P. Sentiment Analysis and Deep Learning: A Survey; Center for Indian Language Technology, Indian Institute of Technology: Bombay, Indian, 2016.
17. Rojas-Barahona, L.M. Deep learning for sentiment analysis. Lang. Linguist. Compass 2016, 10, 701–719.
18. Gupta, U.; Chatterjee, A.; Srikanth, R.; Agrawal, P. A sentiment-and-semantics-based approach for emotion detection in textual conversations. arXiv 2017, arXiv:1707.06996.
19. Roshanfekr, B.; Khadivi, S.; Rahmati, M. Sentiment analysis using deep learning on Persian texts. In Proceedings of the 2017 Iranian Conference on Electrical Engineering (ICEE), Tehran, Iran, 2–4 May 2017; pp. 1503–1508.
20. Jeong, B.; Yoon, J.; Lee, J.-M. Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis. Int. J. Inf. Manag. 2019, 48, 280–290.
21. Balabanovic, M.; Shoham, Y. Combining content-based and collaborative recommendation.Commun. ACM 1997, 40, 66–72.
22. Wang, Y.; Wang, M.; Xu, W. A sentiment-enhanced hybrid recommender system for movie recommendation: A big data analytics framework. Wirel.Commun.Mob.Comput. 2018.
23. Singh, V.K.; Mukherjee, M.; Mehta, G.K. Combining collaborative filtering and sentiment classification for improved movie recommendations. In Proceedings of the International Workshop on Multi-disciplinary Trends in Artificial Intelligence, Hyderabad, India, 7–9 December 2011; pp. 38–50.
24. Asraf Yasmin, B., Latha, R., & Manikandan, R. (2019). Implementation of Affective Knowledge for any Geo Location Based on Emotional Intelligence using GPS. International Journal of Innovative Technology and Exploring Engineering, 8(11S), 764–769. https://doi.org/10.35940/ijitee.k1134.09811s19

25. Muruganantham Ponnusamy, Dr. A. Senthilkumar, & Dr.R.Manikandan. (2021). Detection of Selfish Nodes Through Reputation Model In Mobile Adhoc Network - MANET. Turkish Journal of Computer and Mathematics Education, 12(9), 2404–2410. https://turcomat.org/index.php/turkbilmat/article/view/3720