JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Model for Twitter Mood Visualization and Prediction Based on Deep Learning Approach Deep Tweets Analyzer Model for Twitter.

Arif Md. Sattar[1*], Dr. Sanjay Kr. Tiwari[2], Mritunjay Kr. Ranjan[3]
[1]P.G. Department Mathematics & Computer Science, Magadh University, Bodh Gaya, India.
amsattargaya@gmail.com
[2]P.G. Department Mathematics, Magadh University, Bodh Gaya, India. E-mail: tiwari.dr.sanajy@gmail.com
[3]School of Computer Science and Engineering, Sandip University, Nashik, India. E-mail: mritunjaykranjan@gamil.com

Abstract:

It is possible that in many of today's applications for big data analytics, it will require to evaluate feeds from social media in addition to visualising the views of users. This will give a workable alternative source that we may use to determine new measurements for our digital lives. The open-ended nature of social engagement with other users on Twitter makes media analysis on Twitter much simpler when compared to study of other social media. This is due to the fact that the majority of these forms of communication are private, which results in a distinctive mode of engagement. As a result, the emphasis of this study is narrowed down to the design and implementation of a deep model for visualisation of Twitter opinions (moods) that is based on a deep learning network. It is concerned with Natural Language Processing (NLP)-based sentiment analysis using the Deep Learning framework for the visualisation and categorization of opinions mined from Twitter. The approach that is applied is based on applying natural language processing (NLP) sentiment analysis to a large number of tweets in order to display the projected mood score of the tweet and, as a result, to capitalise on public tweeting for the purpose of knowledge discovery. In addition to that, this will help in the identification of bogus news. The relevant mechanism is comprised of a series of sequential stages, including the stage of dataset collection, the stage of pre-processing, the stage of natural language processing, the stage of sentiment analysis, and the stage of prediction and classification using a deep learning model. The Sentiment Analysis of United States Airlines The dataset from Twitter, which is already available via Data for Everyone, was leveraged for this study. The device being shown is capable of monitoring Twitter streams coming from both the media and the general population. It is able to view and extract significant data from tweets in real time, and then store that data in a Deep model so that it can be analysed afterwards. It is handy for a broad range of applications, including big data analytics solutions, anticipating the behaviour of e-commerce customers, optimising marketing strategy, and gaining a competitive edge in the market, in addition to visualisation in other data mining applications.

Key words: Deep learning, data mining and web mining, visualization in social networks, NLP and sentiment analysis, machine learning.

1.Introduction

There is a growing consensus that social media platforms such as Twitter, Facebook, Snapchat, Instagram, Telegram, and YouTube, amongst others, should be seen as an important and indispensable component of contemporary life [1]. The term "social media analysis" refers to the process of concentrating a learning experience on one or more social media services for the purpose of analysing, deconstructing, and then highlighting the relevant quantitative and qualitative systematic observations that are available in order to evaluate certain large-scale trending topics with particular indications. 'Data analysis', on the other hand, refers to the process of examining, cleaning, manipulating, and modelling data with the intention of gaining access to helpful information, offering suggestions for conclusions, and providing assistance for decision-making. Data mining is a multi- and interdisciplinary topic within the discipline of computer science that refers to the process of extracting new information from massive datasets that have already been created. Discovering patterns in huge data sets via the use of techniques from machine learning, statistics, and database management systems is the discipline known as data mining [2]. The process of extracting useful information from the World Wide Web by locating relevant patterns is referred to as "web mining," and it is an application of data mining methods. Web mining is thus capable of being segmented into three distinct subcategories: web use mining, web content mining, and web structure mining [2, 3]. In the context of social networks, "visualisation" refers to the practise of presenting conceptual data in order to heighten a person's realisation and reveal hidden linkages in conjunction with the data. The visualisation of online information has thus become essential for end users in order for them to get their desired information in an easy, quick, and accurate manner from the enormously large web [4]. The field of computer science and artificial intelligence known as natural-language processing (NLP) focuses on the ways in which computers and human languages interact with one another. In specifically, this refers to the process of programming computers such that they can effectively handle vast volumes of data relating to natural languages. Speech recognition, natural-language comprehension, and natural-language creation are three areas that regularly present difficulties in the field of natural-language processing. [2], [5]. Sentiment analysis, also known as sentimental analysis or opinion mining, is a phrase that is usually used to arrive at one of two conclusions: either people like something or detest something, or the product is either excellent or poor, or someone is either for or against something. The term "sentiment analysis" refers to the application of natural language processing (NLP), text analysis, and statistical methods in order to recognise the "emotional attitude" relevant to a text and classify it as either enthusiastic (positive), unenthusiastic (negative), or impartial (don't-care/neutral). [4]-[6]. The process of machine learning is a kind of artificial intelligence that attempts to imitate or reproduce the way in which human brains work in order to endow computers with intelligence. Knowledge-based systems and expert systems, which are both concerned with the process of knowledge discovery and the implementation of data mining, make extensive use of the concept. Methods that have undergone extensive revision and are used for machine learning often make use of artificial neural networks (ANN). The support vector machine, sometimes known as SVM, is a relatively recent development in the fields of statistical machine learning and data mining [3, 5]. The purpose of this research is to shed light on the aspects of NLP-based sentiment analysis for opinion mining on Twitter that are the most relevant. In order to evaluate and show the opinions expressed on Twitter, it employs a data mining methodology that focuses on the extraction of text features and categorization.

2.Related Works:

Works That Are Related The canon of published works in literature encompasses a diverse range of writing styles and subjects. [1]. presented their work regarding "Ushio," which is a system that analyses public trends in twitter and news media trends. They concentrated on these concerns by developing the Ushio system, which analysed Twitter streams, including tweets that were updated by several news organisations as well as tweets that appeared in the public timeline [7] presented a thorough analysis of the key developing technologies for big data by elaborating on the crucial aspects of these technologies and explaining how they function. They demonstrated a performance measurement of the "Apache Hive" query that was used to execute tweets from Twitter in order to create Map Reduce the amount of time spent by the CPU as well as the overall amount of time needed to complete the task [2]. Hadoop Cloud has been deployed by [8] for the purpose of intelligently analysing and storing large amounts of data. They offer a system that does sentiment analysis on tweets in a cloud-based environment [9]. have created a system for automating the identification of false news on Twitter. This approach involves learning to predict accuracy ratings using two Twitter datasets that are focused on trustworthiness. The first is called CREDBANK and it is a crowd-sourced dataset of accuracy evaluations for events that occurred on Twitter. The second is called PHEME and it is a dataset of possible rumours that occurred on Twitter along with journalistic judgments of how accurate they were [10]. presented some interesting and useful work in the field of sentiment analysis by means of determining the polarity of the speaker in Twitter data. They established some methods for analysing the data's sentiment by supplying some twitter data as input and collecting the relevant ratings as output [5]. Using data from Twitter, [6] attempted to investigate the association between geographical location and emotional state as part of their research. They conducted an investigation of the factors that contribute to the incidence of criminal activities in the area by using data from Geo-Twitter. Data obtained from Location-based Social Networks (LBSNs) have been used in modelling and gaining an understanding of human-mobile behaviour patterns associated with local criminal activities [5]. have done a very excellent job of defining the function that text pre-processing plays in the process of sentiment analysis for mining internet opinion. They showed that the accuracy of sentiment analysis in this domain utilising support vector machines (SVM) may be greatly improved by applying suitable feature selection and representation via their experimental findings, which were shown in [4]. presented their research on the topic of network-based visualisation of users' opinion mining and sentiment analysis on Twitter. They built a free and open source system that can take the opinion of users in raw text format and generate an easy-to-interpret visualisation of opinion mining and sentiment analysis results on a social network.

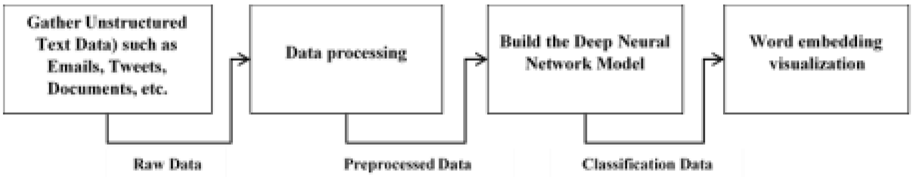
This system can take the opinion of users in raw text format. The LingPipe Library, which is a public machine learning library, was used to categorise the feelings of users' opinions into positive, negative, and neutral categories [3]. presented their work (which was based on convolutional and recurrent neural networks) about social emotion mining approaches to anticipate people' responses to Facebook postings. They suggested and assessed several approaches for predicting these responses to user postings that were made on the public sites of corporations and companies (such as grocery chains) [11]. Lasty, [2] have developed "two" efficient methods to mine large amounts of data. The first method uses Apache Hadoop Map Reduce, while the second method is a visualization-based method called Visual Web Mining (VWM). They came to the conclusion that VWM is efficient for displaying and gleaning the front end insight of large data, but Apache Hadoop and other comparable technologies are efficient for supporting back-end issues such as storage and processing.

3.Background Concepts

Theoretical Foundations Machine Learning (ML) is the study of the behaviour of computer systems and how they may learn from the data they are given. In a sense, the study of computer algorithms that improve themselves automatically via experience is what we mean when we talk about machine learning [12]. In recent years, it has grown pervasive in a number of domains that are competing with or even exceeding the degree of knowledge offered by humans [13], [14]. There are many different general applications in which machine learning is used to predict or classify a problem. Some examples of these applications include the detection of message spam, face detection, object recognition, pattern recognition, speech recognition, text classification, and translation [15]. The learning job that falls within the purview of machine learning (ML) may take on a variety of guises, including supervised, unsupervised, or semi-supervised learning. In addition, supervised learning is shown if the computer is used as an example of an input together with the intended outputs, which are provided by a "teacher" in this scenario. One of the most often cited examples of supervised learning is the process of screening the input messages to determine whether they are spam communications or ham messages. When filtering email, a supervised learning algorithm is given email messages that have been tagged as examples of "spam" or "not spam." These labels are then used to train the system. Because of this, the term "machine learning" refers to a computer system (software) that is able to accurately forecast and identify fresh input messages as either spam or non spam [16]. Over the last several years, deep learning strategies have shown excellent performance in a variety of computer vision and pattern recognition applications. Deep learning allows for the automatic learning of feature sets for specific challenges rather than the hand-crafted design it traditionally required. One of the most common forms of deep learning models [17]-[18] used in image processing is called a convolutional neural network (CNN). The conventional news aggregator known as CNN is an example of a feed-forward artificial neural network. It is a network for learning that enables several levels of representation and abstraction at the same time. The input layer, the convolution layer, the down sampling layer, and the output layer are the components that make up a CNN, which is a multi-layer perceptron. It's possible that the convolution and down sampling levels of a CNN design are made up of many layers each. The CNN architecture is favoured because of its straightforward construction, less reliance on the number of training parameters, and flexibility. Fully Convolutional Network, also known as FCN, and Region-based Convolutional Network are both examples of expansions of CNN (R-CNN). Fully Convolutional Network, abbreviated as "FCN" [19, 20]. complete convolution layers are used in place of the fully linked layers in CNNs. Additionally, class labels are assigned to each each pixel in the picture rather than having a single label assigned to each image block. The first step in text mining using Deep Learning is to process the many documents that have been obtained. In other words, the extraction of information or features from a document via the application of Deep Learning-based methods and tools, which are primarily utilised for the purpose of processing such features or information. [21]. The text analysis phase, also known as the pre-processing step, is the most important part of the text mining process when Deep Learning techniques are used. In this phase, a number of approaches are used on a recurrent basis in order to extract some essential information from the process papers that are located in this stage [22]. Text Deep Learning techniques extract some information from the semi-structured and structured datasets such as e-mails, text files, HTML files, and so on [21]. Deep Learning approaches or tools arrange the document or data structure from the databases just once. Nevertheless, making use of Deep Learning technologies and strategies is the greatest alternative for organising and managing the data that is found online [15]. Figure 1 is an illustration of a high-level generic method to text mining that makes use of Deep Learning.

3.1. Opinion Mining Based Classification Approaches

Classification Methods That Are Derived From Opinion Mining In general, opinion mining approaches as it has been shown in a high-level abstraction and general approach in Fig. 1, the opinion mining using machine learning consists of many tasks and functions such as opinion clustering, opinion concept/entity extraction, opinion summarization, and opinion classification. Fig. 1: A high-level abstraction and general approach to opinion mining. Text categorization is one of the most significant ways for text mining when utilising machine learning technologies, and it's also one of the simplest straightforward. Text classification is synonymous with text categorization and goes by both names. Text categorization is the process of automatically defining unlabeled texts into a set of previously established categories [23]. The categorization of options should be approached as a challenge in machine learning, and the mathematical model should be explicitly established. In this scenario, in addition to the mathematical model, we need also pre-



define the feature extraction model. This model is primarily used for the purpose of obtaining the feature vector that

Figure 1. A high level of text mining general approach using deep learning approach

will subsequently be utilised to feed the machine learning model [24].

3.2. Convolutional Neural Network Based Opinion Classification Approach

An Approach to Opinion Classification Constructed Using Convolutional Neural Networks The Convolutional Neural Network, or CNN, is a function that maps one set of input data, indicated by x, to an additional output vector, which is denoted by y. Then, the function g is the intense combination of a succession of simpler functions fl, which we call computational blocks, or layers that are represented by the equation g = f1... fL. These computational blocks are what we mean when we say that g is the result of this combination. Assume that the input to the network is x0 = x and that the outputs to the network are x1, x2, etc., xL, etc. Where each output is thoroughly calculated from the previous output xl 1 by applying the function fl with the parameters of wl as is specified in (1) [25]. This process is repeated until all outputs have been computed.

$$xl = fl(xl - 1; wl) \ldots\ldots\ldots\ldots\ldots(1)$$

Since the data x has a spatial structure, Hl and Wl are spatial coordinates, and Dl is the depth of channels. The data that is flowing through the network forms a feature field that is denoted by the equation xl = RHl | Wl | Dl. Because the functions fl and fl behave as local and translation invariant operators, the network is referred regarded as a convolutional network. In order to differentiate between the various classes, CNNs are used to generate a vector of probabilities, which is then represented by for each and every picture that is tested [25]. This is shown in (2).

$$\hat{y} = f(x) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

CNN performance of true label y of data x is assessed by a loss function that imposes a penalty to classification mistakes [25]. If y is the true label of the data, then the loss function measures CNN performance of true label y of data x.The output of the preceding layer is sent to the convolutional layer, where it is then convolved using a series of learnable filters [26], as illustrated in Figure 2, where the weights describe the convolution filter. Each filter is being moved horizontally and vertically over the input volume, which results in the production of a two-dimensional activation map for that filter.. The depth of the filters is same to that of the input [27], [28]. The depth, stride, and zero-padding are the three main parameters that may be used to modify the output in order to get the desired level of output size.

The number of filters that are used to process the input data determines the "depth" of the convolutional layer. Despite the fact that the stride number of convolutional filters makes it possible for the filter to leap when sliding over the data size dimensions The next step is the zero-padding, which involves placing zeros around the input's boundaries in order to maintain its size.. The processing example with I is shown in Figure 2.

t

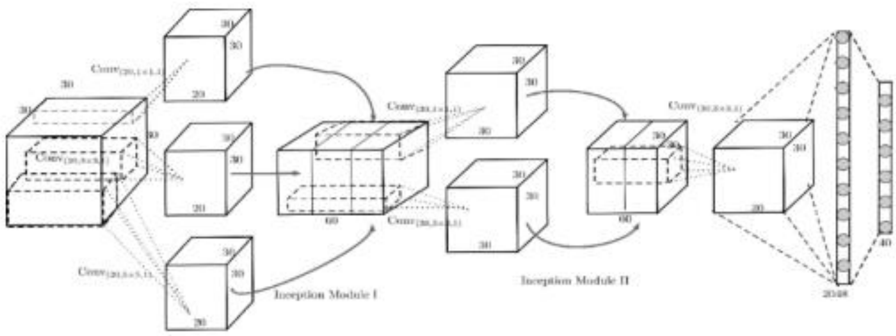


Figure 2 Convolution neural network (deep Learning) Structure [28]

The pooling layer facilitates multi-scale analysis while also reducing the volume of their input. The most common types of pooling operations are known as max-pooling and average-pooling (Fig. 2). These operations determine the greatest value or the average value contained inside of a relatively tiny spatial block [27]. It is recommended to use a pooling method with filters that have a size of 2 by 2 and a stride of 2 [29]. Figure 2 depicts the maximum pooling procedure with filters arranged in a 2 x 2 grid. In addition, it demonstrates an example of a max 2 2 pooling layer and explains how this kind of layer may be used to lower the spatial dimension of a representation, hence cutting down on the amount of parameters and computations required by a network. A fully-connected layer is one that is linked to each and every neuron in the layer below it (Fig. 2) [29]. The categorization is normally carried out by the fully linked layers, which are often utilised as the last layer of the network. The Convolutional Neural Network (CNN) is illustrated in full, including both of its layers, in the example shown in Figure 2.

3.3. A Mathematical Model for the Classification of Opinion Mining

The mathematical formulation of the opinion mining and classification model is dependent on the display of documents in order to gain the capacity to appropriately carry out a categorization activity. In addition, text categorization might be defined as follows, according to (3) [24]:

x equals (x1, x2,..., xn), where n is the total number of classified documents that were utilised to classify the whole data job [24]. Because the classification job for the text classification is an important component of the supervised learning that we have previously discussed, a machine learning classification model for text classification is provided as an example of a training set. Input samples are included in the training set, along with any related labels or desired output. The training examples, which are presented in the form of attribute vectors, are, in essence, a subset of Rn [24], which indicates that the input is a subset of Rn. Consider, for instance, an input vector X equaling the components x1, x2,..., and xn, where X is a member of an n-dimensional vector space Rn and where x1, x2,..., and xn are the vector's components. If this is the case, then is allocated to the positive class whereas the function of,f(x) 0 and to the negative class if f(x) 0 as it is stated in (4) [24]. In other words, this particular instance of is assigned to the positive class.

Output = {Postive, if $f(x) \geq 0$
Negative, if $f(x) < 0$
}…………………………………………… (4)

In this particular instance, the function f(x) may be interpreted as a choice function since every vector has a target characteristic. To put it another way, the target label will be disproved by the expression Y = *1, +1+, where I = 1,..., n and -1 and +1 respectively represent the negative and positive classes [24]. The mapping function implying as X Y is learnt using a typical model for machine learning. This function may be represented by a collection of alternative learned mappings, such as the one shown in (5) [30] OX f(X,) (5), where is a set of parameters that are connected with the function f(x). For example, if you provide the machine learning model with an input value of X and a parameter value of, where is a value that should be selected in some manner, the model will always produce about the same result [30]. The creation of the classification model may be inferred from the number of classes included in the model. For instance, in the event that there are two classes, the objective here is to build a binary classifier, which is built using the training examples. On the other hand, the predictor-target value pairings that are used for training the machine have a low chance of producing incorrect classifications (also known as classification errors) when applied to a testing sample that is used for validating the machine [31]. Take, for instance, the document classification issue X, which is given in the form of a feature vector. This document's user may designate a category, which is represented by the Y component of this feature vector [30, 32]. This feature vector is made up of the frequency of different terms that make up the distribution. A method termed as "Supervised Learning" may be used to instruct the CNN structure's fully linked layer. In this method, input samples are supplied to the network together with a set of intended answers at the output layer. This is done in order to achieve the desired results. It is determined what the disparities (errors) are between the intended answer and the actual response for each node in the output layer. The learning algorithm takes into account a variety of factors before making adjustments to the weights of nodes in the network. The goal is to construct an input-output model that has accurate mapping, with the intention of correctly predicting outputs based on inputs that have not yet been seen [33]. Learning, also known as training, is a process that involves making adjustments to the parameters of a network in order to get the gap between the actual output and the planned output as close to zero as is practically feasible [34]. The weights start out in a random order, and then they are adjusted in the direction that results in the least amount of inaccuracy. The fluctuation in weights is shown in (6)

$$\Delta w = -\eta \frac{\partial E}{\partial w} \qquad (6)$$

where the learning rate, which represents the relative magnitude of the change in weights, is shown by the symbol. The output result for neuron j is provided in equation (7), and it is based on the input units xi = (x1, x2,..., xN):

$$net_j = \sum_{i=1}^{n} w_{ji} x_i + w_{j0} = \sum_{i=0}^{n} w_{ji} x_i = \langle w^j, x \rangle \qquad (7)$$

where I represents an index for the input layer, j represents an index for the unit on the hidden layer, and n represents the total number of input units. wji is shorthand for the input-hidden weights that are present at the hidden unit j. In equation (8), we derive the output of neuron j by applying the activation function f to the net output netj. This results in the following:

$$y_i = f(net_j) \qquad (8)$$

For the hidden units, the outputs that are obtained from the equation (10) and sent to neuron k of the output layer are as follows:

$$net_k = \sum_{j=1}^{n} v_{ki} y_j + v_{k0} = \langle v^k, y \rangle \qquad (9)$$

where the index unit of the output layer is denoted by the subscript k, and the number of hidden units is denoted by the subscript p.

$$z_k = f(net_k) \qquad (10)$$

4.Propose System:

Because users are more likely to be open and honest about their perspectives on Twitter, it makes for a fantastic place to begin doing research on social media beliefs. This is completely different from Facebook, where people choose to keep their social relationships private. In this paper, we present a method that is depicted in Figure 3 and is based on deep learning. The core components of the proposal system are as follows: The first step is the preparation of the Twitter dataset using several Natural Language Processing (NLP) technologies. The second ally is the dataset visualisation on Twitter. Finally, a Datamining strategy for opinion mining on Twitter that is built on a prediction and classification using Deep Learning.

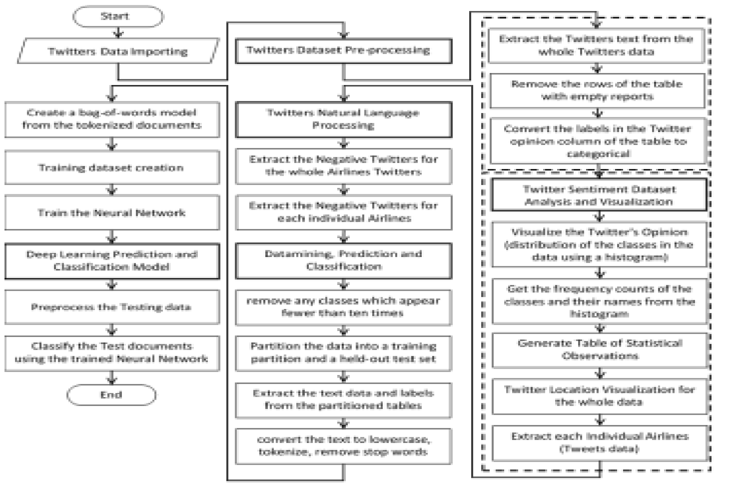


Figure 3. Opinion-mining based deep learning propose approach.

4.1. Twitters Dataset Pre-processing Based Natural Language Processing

In this stage of the process, several Natural Language Processing (NLP) techniques are used for the purpose of pre-processing the tweets in order to extract and forecast the tweets' opinion mining method. The data from Twitter is processed in the beginning so that the row text may be extracted from the whole dataset. Following the extraction of the text from Twitter from the whole of Twitter's data, we proceed to eliminate rows from the database that include reports that are blank. Following this, we proceed to categorise the Twitter opinion across the board by converting the labels in the Twitter opinion column of the table to categorical terms. This phase includes the execution of three separate preceding stages. The next step is to exclude from the bag-of-words model any terms that do not occur a total of more than two times. After that, we delete any documents from the bag-of-words model that do not include any words and also remove any entries in the labels that belong to such documents.

4.2. Twitter Sentiment Dataset Analysis and Visualization.

In this stage, we make use of a histogram to create a visual representation of Twitter's Opinion based on the distribution of the classes included in the data. The most frequent kind of analysis that may be done on a large number of tweets is called sentiment analysis. The terms that are found in a tweet are used to compile a score for sentiment analysis. The use of sentiment analysis offers a practical method for gauging the mood of the general audience who tweets. However, we can see that tweets referencing both brands have a strong tilt toward the negative, despite the fact that the sentiment distributions for the two companies are essentially equal to one another.

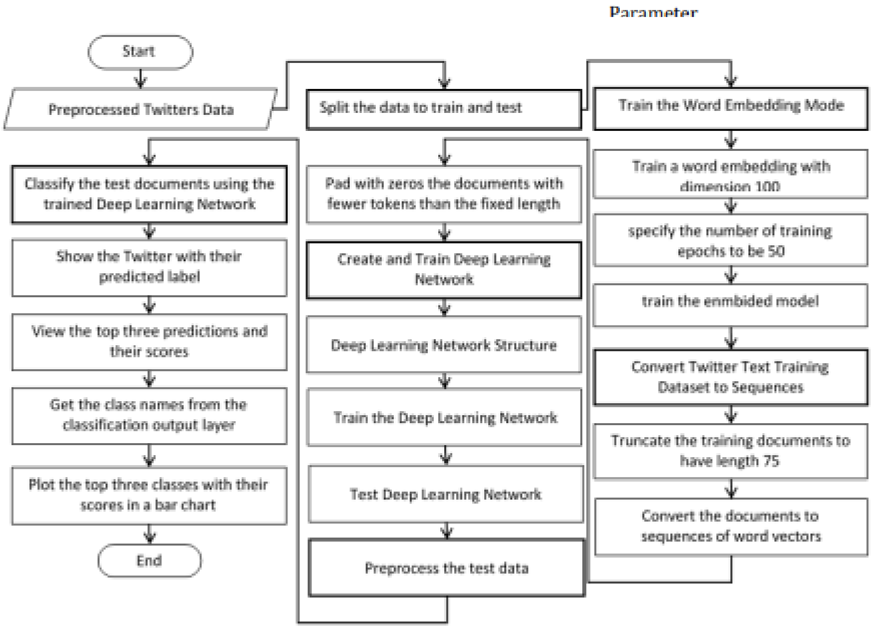
4.3. Prediction and Classification Utilizing Deep Learning Systems

In this stage of the process, datamining methods are used for opinion mining production and classification on Twitter. In order to complete this objective, we make use of a number of datamining strategies, including dimensionality reduction, feature selection, and supervised learning inside a classification framework. In order to achieve it, a number of major stages are carried out, including the following: Prepare the training data in advance. First, remove all of the punctuation marks, then change the text to lowercase, and last, tokenize it. Stemming or removing words is not recommended since it may result in an inadequate match for the word embedding. The words of a language may be mapped to numeric vectors using word embeddings. These embeddings are able to capture semantic characteristics of the words, making it such that words with similar meanings have vectors that are also comparable. In addition to that, they use vector arithmetic to describe the connections between the words. For instance, the connection "king is to queen as man is to woman" may be expressed by the equation "king - man + woman = queen." Another example would be "man is to woman as king is to queen." The documents were fed into the deep learning network, and the network then converted the documents into sequences of word vectors. In this scenario, while training the network, the software makes mini-batches of sequences of the same length by padding, truncating, or separating the input data. These sequences are then used to train the network. Input sequences may be padded or truncated using the training Options function's provided options; however, these options are not ideal for use with sequences consisting of word vectors. Instead, you are going to need to manually pad and truncate the sequences. It is possible that the training will become more effective if you left-pad and truncate the sequences of word vectors. In the initial phase of the conversion process, you will choose a goal length. After that, you will left-pad documents that are shorter than the target length and truncate documents that are longer than the target length. A shorter goal duration is preferable for producing the greatest results; however, this should not be done at the expense of throwing away significant quantities of data. Viewing a histogram of the training document lengths may help discover an appropriate goal length for the document. The majority of the training papers have a token count that is lower than 75. Shorten the training manuals so that they are just 75 words long. The anonymous function inputted to documents accepts input in the form of a string array and returns the top 75 items of that array. Perform a word vectorization on each of the texts. Utilize the example function doc2sequence, which is shown at the conclusion of this example, to convert the training documents into a cell array of sequences. The word vectors are represented by the columns of each sequence. If you have Parallel Computing Toolbox TM installed on your computer, the function will do a parallel iteration of the documents. In such case, the function will iteratively go through all of the documents one after the other, which may take a few minutes. Documents containing fewer tokens than the set length should have zeros padded in between them. Left-padding the sequences is required in order to pad the sequences with word vectors for the Deep Learning network model. You are need to right-pad the sequences manually since the sequence padding option for the Deep Learning network model, by default, right-pads the sequences. In order to properly construct a Deep Learning network using sequences of word vectors, you need to left-pad the sequences. You are need to right-pad the sequences manually since the sequence padding option for the Deep Learning network model, by default, right-pads the sequences.

4.4. Create and Train the Deep Learning Network

Include a sequence input layer in the network and make the size of the input space equal to the dimension of the word embedding in order to feed sequence data into the network. The next step is to include a Deep Learning Network layer, and you should set the output size to 180. Make sure that the output mode is set to 'last' before you attempt to utilise the Deep Learning Network layer to solve a sequence-to-label classification issue. The next step is to add a fully linked layer, a SoftMax layer, and a classification layer. The size of each of these layers should be proportional to the number of classes. In Table 1, you may find a description of the Deep Learning structural model.

Table 1. Architecture of the CNN Training/Testing Model for the Final Classification Approach



The general flowchart of the Deep Learning Model that has been used in our approach is illustrated below in Fig. 4:

5.Experimental Results

Experimental Results Evaluation of the findings that were achieved using the proposed system may be found in this section. which is compatible with the Windows 10 operating system, is used throughout the course of this study. The assessment was broken down into its two primary areas, which led to the collection of a variety of outcomes. The first one is a statistical and graphical representation of the opinions expressed on Twitter. The second one is the development of opinions and the categorization of them on Twitter. In this research, we evaluate the validity and precision of the findings obtained from the tests by making a direct comparison between the various algorithms for dimensionality reduction. a confusion matrix for the detection system is specified as a m x m matrix, where m specifies the number of classes. This differs from our method, which uses a n x n matrix. The actual classifications as well as any predictions made by a detection system are recorded in the information contained inside a confusion matrix. It is usual practise to assess the performance of such systems using the data included in the matrix. The examples that belong to each anticipated class are represented along the columns of the matrix, while the instances that belong to each actual class are shown along the rows. The confusion matrix illustrates the classes that have been categorised appropriately as well as the classes that have been incorrectly classified. For the purpose of evaluating these factors, a confusion

Figure 4 Deep learning Twitter opinion mining Classification approach.

matrix is utilised [35]. 5.1. Airlines The Observation and Visualization of Statistical Data The initial outcome of this experiment was to create a visualisation of the whole Twitter population's opinion based on the classes (mood) distribution for the entire dataset. Based on the observation utilising the statistical measurement for each mood class, Fig. 5 and Table 2 display the sentiment analysis (opinion histogram) and the statistical measurement for the whole dataset of tweets collected from Airlines Twitter accounts, respectively.

5.1. The Whole Twitters Dataset Visualization:

Figure 6 world cloud visualization of each twitter's mood location

In this experimental result, we displayed the (Twitters location) for the whole dataset as an indicator for the twitter opinion locations based on the word cloud visualisation. This was done so that we could see where people's opinions on Twitter are located. The word cloud depiction of each Twitter user's location and mood is shown in Figure 6. Separate out each individual tweet from the airline: Within the scope of this investigation, we furthermore glean the statistical information about each specific airline.

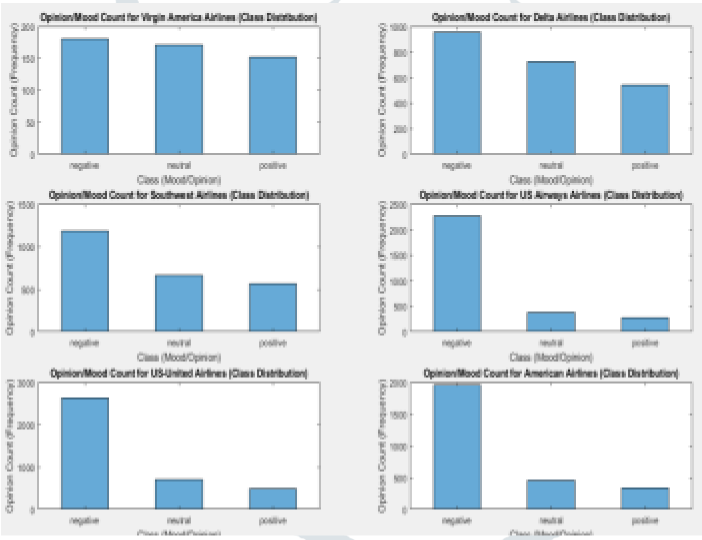


Figure 7

Fig. 7. A sentiment analysis (opinion histogram) of the current state of mind on Twitter about each specific airline. The sentiment analysis (opinion histogram) of the mood on Twitter is shown in Figure 7. This analysis is based on the opinions expressed on Twitter for each specific airline.

5.2. Reasons to Be Pessimistic

An Examination of the Entire Airline System As a consequence of this investigation, we were able to identify and depict the factors that contribute to entire Airlines' problems. Table 3 presents the results of an

Table 2.

No	Negative Reason	Counted No.
1	Late Flight	1665
2	Damaged Luggage	74
3	Longlines	178
[illegible]	[illegible]	[illegible]
5	Can not Tell	1190
6	Flight Attendant Complaints	481
7	Cancelled Flight	847
8	Flight Booking Problems	529
9	Bad Flight	580
10	Lost Luggage	724

examination of the negative sentiments and observations about the whole dataset. We extracted and made a graphic representation of the unfavourable causes for each specific airline. Table 5 presents the results of the statistical observation and the sentiment analysis conducted on all of the airlines included in the overall dataset.

Table 3. Negative Reason Observation for Whole Airlines

Training Parameter	Setting Value	
Momentum	0.9000	
Initial Learnning Rate	0.0100	
Learnning Rate Schedule Setting	DropRateFactor	0.2
	DropPeriod	5
L2 Regularization	1.0000e-04	

Table 4. Training Setting Parameter

MaxEpochs	20
MiniBatchSize	32
Verbose	1
Verbose Frequency	50
Validation Data Frequency	50
Validation Patience	5
Shuffle	1
Sequence Padding Value	0

Table 5. Statistical Observation of Each Airlines

Airlines Name	Statistical Observation		Airlines Name	Statistical Observation	
Virgin America	Negative	181	US_Airways	Negative	2263
	Neutral	171		Neutral	381
	positive	152		positive	269
Delta	Negative	955	United	Negative	2633
	Neutral	723		Neutral	697
	positive	544		positive	492
Southwest	Negative	1186	American	Negative	1960
	Neutral	664		Neutral	463
	positive	570		positive	336

5.3. Opinion Mining and Training Model Deeply Rooted in Twitter

There are many different parameters that may be used to conduct an analysis of the performance of the opinion mining detection and classification system that uses the Deep Learning Approach. Train Network will utilise a GPU automatically if one is available; however, you will need Parallel Computing ToolboxTM and a CUDA® capable GPU with compute capability 3.0 or above in order to use one. Otherwise, it utilises the central processing unit. Make advantage of the 'Execution Environment' name-value pair parameter found inside the training Options menu in order to explicitly set the execution environment. Training on a central processing unit (CPU) might take a lot more time than training on a graphics processing unit (GPU). To begin the process of training the Deep Learning network, we must first input the documents into a structure for the Deep Learning Network known as an LSTM network. This is accomplished by transforming the texts into sequences of word vectors. Then, while the Deep Learning Network (LSTM Network) is being trained, the model makes mini-batches of sequences of the same length by padding, truncating, or dividing the input data. This is done so that the sequences all have the same length. Input sequences may be padded or truncated using the training options function's provided options; however, these options are not optimal for use with sequences consisting of word vectors. Instead of that, there is the need of manually padding and truncating the sequences. Table 4 presents the training setup parameters that are used in the Deep Learning model that we have developed.

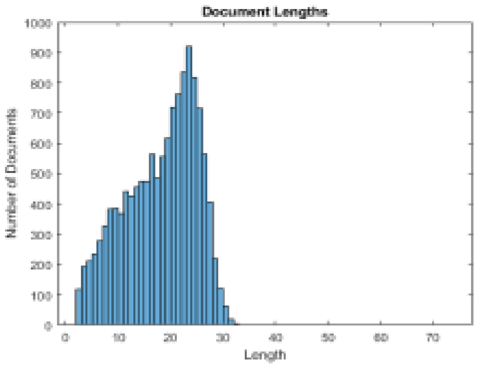


Figure 8

Fig.8.Choosing the data (Twitter text) goal length is the first stage in the process of training the Deep Twitter Opinion Mining prediction. After this step, you will truncate documents that are larger than the target length and left-pad documents that are less than the target length. A shorter goal duration is preferable for producing the greatest results; however, this should not be done at the expense of throwing away significant quantities of data. Viewing a histogram of the training document lengths, such as the one shown in Fig. 8, may help you select an appropriate target length.

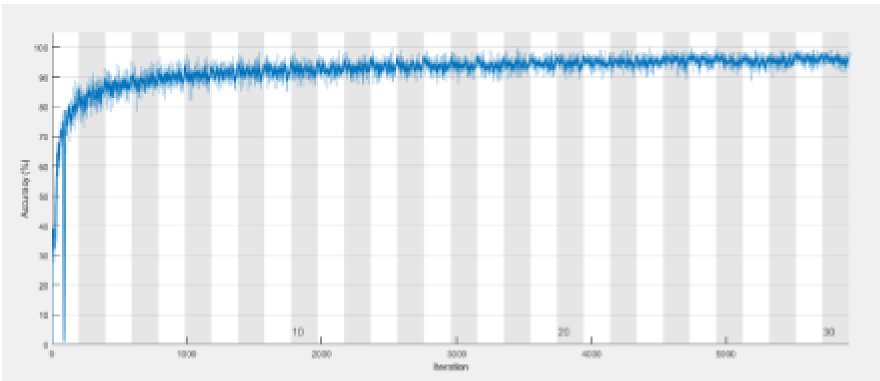


Figure 9

Fig. 9. The comprehensive performance results of the strategy of using deep opinion mining (training accuracy). Figure 8 and Figure 9 depict, respectively, the overall training progress of the Deep Twitter opinion-mining strategy that makes use of the Deep Learning approach and the loss function plot.

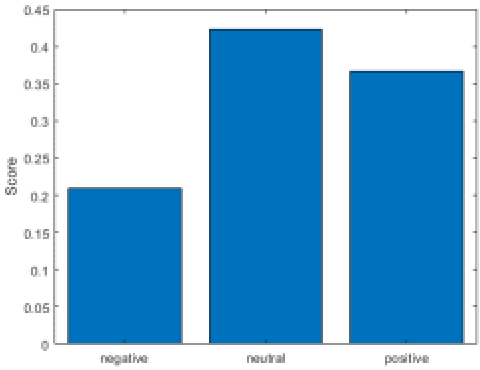


Figure 10

Fig. 10. Predict visualization of the deep twitter opinion mining prediction model

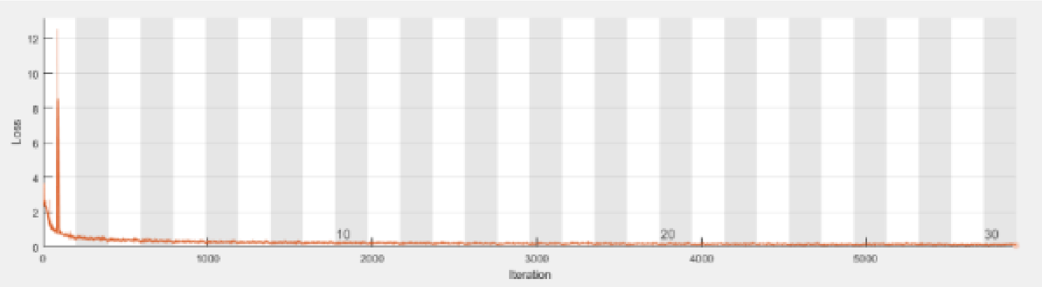


Figure 21

Fig. 11. Overall performance results of the opinion mining approach (lost function).

5.4. Model for Opinion Mining, Visualization, and Prediction Based on Deep Twitter Data

We utilise the taught parameters to forecast each Twitter mood using the Deep Learning test model. This allows us to predict and categorise the mood on Twitter based on the trained Deep Learning model for the testing dataset, which is comprised of Twitter's text data. In order to show the Twitter Opinion prediction, also known as the (Mood), we visualise the top three forecasts, together with their scores, for each Twitter that was used in the testing data. The prediction scores should be sorted, and then the top three scores should be selected and visualised. The initial tweets from the testing dataset are shown as an example in Figure 10, together with the visualised cores of each anticipated mood that match to those tweets (opinion). Table 6 provides an example of the overall results that can be anticipated using Deep Twitter Opinion Mining, and Figure 11 provides a representation of the results that can be predicted using Deep Twitter Opinion Mining.

Table 6. Overall Tested Samples of the Mood Predict Visualization Results

Twitter Test No.	The Requested Twitter	The Predicted Mood	Actual Mood	Prediction Score %
1	virginamerica view of downtown los angeles the hollywood sign and beyond that rain in the mountains httptcodw5nf0ibtr	neutral	neutral	0.517001
10	virginamerica is it me or is your website down btw your new website isnt a great user experience time for another redesign	negative	negative	0.737704
40	united customer service is atrocious you have disrupted my travel plans you have lost my luggage and it is impossible to talk to a human	negative	negative	0.902131
80	virginamerica i spoke with a representative that offered no solution i am a loyal customer who flies on virginatlantic as well	negative	negative	0.713250
100	united premier gold desk changes flight waives fees gives me wrong flight now jana acosta in salt lake refuses the same service angry	negative	negative	0.713250
1300	americanair delaney and shawn at dfw showed exceptional customer service today will happily choose aa whenever possible now thank you	positive	positive	0.625260

6.Conclusion

Twitter has once again been the focal point of attention due to recent happenings, which demonstrates the growing significance of social media in today's world. Regardless of how you feel about social media in this day and age, there is no disputing the fact that it has become an essential component of our lives online. People freely express their thoughts and ideas with the wider public on Twitter, making it an excellent place to begin an investigation into social media. This is in stark contrast to social networking sites like Facebook, where conversations between users are often kept private. In this research, we offer a Deep Learning framework for Twitter opinion mining, with the goal of classifying user sentiment. In addition to this, we implemented the Twitter Opinion abstraction and visualisation method using a Deep Learning model. The most important addition that this study makes is the proposal of a novel visualisation model for mood prediction on Twitter that is based on the Deep Learning method. After the Deep Learning model is trained using the learned dataset that has been randomly picked from the original dataset, the proposal system visualises the top three predictions and their scores for each Twitter in the testing data. This is done after the testing data has been loaded. Then, the proposal system will sort the prediction scores, choose the top three values, and show it as the primary mood that has been projected for Twitter (opinion).

7. References

[1] Yao, F., Chang, K. C. C., & Campbell, R. H. (2015). Ushio: Analyzing news media and public trends in Twitter. Proceedings of the 2015 IEEE/ACM 8th Int. Conf. Util. Cloud Compute. UCC 2015 (pp. 424-429).

[2] Mohata, P. B. (2015). Web data mining techniques and implementation for handling big data. International Journal of Computer Science and Mobile Computing, 4(4), 330-334.

[3] Krebs, F., Lubascher, B., Moers, T., Schaap, P., & Spanakis, G. (2017). Social emotion mining techniques International Journal of Computer and Communication Engineering 14 Volume 8, Number 1, January 2019 for Facebook posts reaction prediction. Cornell University. arXiv: 1712.03249v1 [cs.AI].

[4] Molla, A., Biadgie, Y., & Sohn, K. A. (2014). Network-based visualization of opinion mining and sentiment analysis on twitter. Proceedings of the 2014 Int. Conf. IT Converge. Secure, ICITCS 2014 (pp. 2012-2015).

[5] Haddi, E., Liu, X., & Shi, Y. (2013). The role of text pre-processing in sentiment analysis. Procedia Compute. Sci., 17, 26-32.

[6] Kitaoka, S., & Hasuike, T. (2018). Where is safe: Analyzing the relationship between the area and emotion using Twitter data. Proceedings of the 2017 IEEE Symp. Ser. Compute. Intel. SSCI 2017 (p. 18).

[7] Bhardwaj, A., Vanraj, A., Kumar, Y., Narayan, & Kumar, P. (2016). Big data emerging technologies: A case study with analyzing twitter data using apache hive. Proceedings of the 2015 2nd Int. Conf. Recent Adv. Eng. Compute. Sci. RAECS 2015.

[8] Kumar, M., & Bala, A. (2016). Analyzing Twitter sentiments through big data. Proceedings of the 10th Indiacom - 2016 3rd Int. Conf. Compute. Sustain. Glob. Dev. (pp. 262-263).

[9] Buntain, C., & Golbeck, J. (2017). Automatically identifying fake news in popular twitter threads. Proceedings of the 2nd IEEE Int. Conf. Smart Cloud, Smart Cloud (pp. 208-215).

[10] Algur, S. P., & Patil, R. H. (2018). Sentiment analysis by identifying the speakers polarity in Twitter data . Proceedings of the Int. Conf. Electra. Electron. Commun. Compute. Technol. Optima. Tech. ICEECCOT 2017 (pp. 898-902).

[11] Bhardwaj, A., Vanraj, A., Kumar, Y., Narayan, & Kumar, P. (2016). Big data emerging technologies: A case study with analyzing twitter data using apache hive. Proceedings of the 2015 2nd Int. Conf. Recent Adv. Eng. Compute. Sci. RAECS 2015.

[12] Nestorov, S. (2000). Data mining techniques for structured and semi structured data. PHD thesis, Stanford University.

[13] Huiqing, L. (2004). DM: Effective use of data mining technologies on biological and clinical data. PHD thesis, University of Singapore.

[14] Choromanska, A. (2014). Selected machine learning reductions . PHD thesis, Colombia University.

[15] Smola, A., & Vishwanathan, S. V. N. (2010, October 1). Introduction to Machine Learning. Cambridge, United Kingdom: Press Syndicate of the University of Cambridge the Pitt Building.

[16] Harrington, P. (2012). Machine Learning in Action. United States of America.

[17] Sainath, T. N., Mohamed, A., Kingsbury, B., & Ramabhadran, B. (2013). Deep convolutional neural networks for lvcsr. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 8614-8618).

[18] Sainath, T. N., Kingsbury, B., Mohamed, A., Dahl, G. E., Saon, G., Soltau, H., Beran, T., Aravkin, A. Y., & Ramabhadran, B. (2013). Improvements to deep convolutional neural networks for lvcsr. Proceedings of the Workshop on Automatic Speech Recognition and Understanding (ASRU) (pp. 315-320).

[19] Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2014). Semantic image segmentation with deep convolutional nets and fully connected CRFs. Cornell University. arXiv:1412.7062 [cs.CV].

[20] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 3431-3440).

[21] Shalev, S., & Ben-David, S. (2014). Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press.

[22] Alpaydn, E. (2010). Introduction to Machine Learning, Second Edition. Massachusetts: MIT Press Cambridge.

[23] Swadia, J. (2016, May). A study of text mining framework for automated classification of software requirements in enterprise systems. Arizona State University. International Journal of Computer and Communication Engineering 15 Volume 8, Number 1, January 2019

[24] Aggarwal, C. C. (2015, March 8). Data Mining: The Textbook. New York: IBM T. J. Watson Research Center Yorktown Heights.

[25] Koushik, J. (2016). Understanding convolutional neural networks. ArXiv preprint arXiv: 1605.09081.

[26] Vedaldi, A., & Matconvnet, K. L. (2015).Convolutional neural networks for matlab. Proceedings of the ACM International Conference on Multimedia (pp. 689-692).

[27] Gibian, A. Sky Convolutional Neural Networks. Retrieved from http://andrew.gibiansky.Com/blog/machine-learning/convolutional neural-networks/

[28] Convolutional neural networks for visual recognition. Retrieved from http://cs231n.github.dio/convolutional-networks/

[29] Iftene, M., Liu, Q., & Wang, Y. (2016). Very high resolution images classification by fine tuning deep convolutional neural networks. Proceedings of the Eighth International Conference on Digital Image Processing (p. 100332).

[30] Aggarwal, C. C. (2012). A survey of text classification algorithm. Yorktown Heights: IBMT. J. Watson Research Center.

[31] Guduru, N. (2006). Text mining with support vector machine and non-negative matrix factorization algorithm. University of Rhode Island.

[32] Addis, A. (2016). Study and development of novel techniques for hierarchical text categorization. PHD thesis, Department of Electrical and Electronic Engineering, University of Cagliari.

[33] Haykin, S. (2008). Neural networks and learning machines third edition. McMaster Thesis, University Hamilton Canada.

[34] Han, J., & Kamber, M. (2006). Data Mining: Concepts and Techniques. University of Illinois at Urbana Champaign, Elsevier Inc.

[35] Mahindrakar, P., & Hanumanthappa, M. (2013). Data mining in healthcare: A survey of techniques and algorithms with its limitations and challenges. Int. Journal of Engineering Research and Applications, (3), 937-941