A project report on

MULTILINGUAL SEMANTIC ANALYSIS USING NEURAL NETWORK

Submitted in partial fulfillment of the requirements for the award of degree of

Master of Technology

In

Information System

Submitted by:

Sachin Kumar

(2K15/ISY/15)

Under the guidance of

Dr. O.P. Verma

(Professor, Department of Computer Science and Engineering, DTU)

2015-2017

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

Bawana Road, Delhi-110042

Certificate

This is to certify that **Sachin Kumar (2K15/ISY/15)** has carried out the major project entitled "**Multilingual Semantic Analysis using Neural Network** " in partial fulfillment of the requirement for the award of Master of Technology Degree in Information System during session 2015-2017 at Delhi Technical University.

The major project bonafide piece of work carried out and completed under my supervision and guidance. To the best of my knowledge, the matter embodied in the thesis has not been submitted to any other University/Institute for the award of any degree or diploma.

Dr. O.P. Verma

Professor

Department of Computer Science and Engineering

Delhi Technological University

Delhi – 110042

Acknowledgments

I express my sincere gratitude towards my project mentor **Dr. O. P. Verma**, Professor, Department of Computer Science and Engineering, Delhi Technological University, Delhi, for providing valuable guidance and constant encouragement throughout the project. It is my pleasure to record my sincere thanks to him for his constructive criticism and insight without which the project would not have shaped as it has.

I thank God for making all this possible, my parent and friends for their constant support and encouragement throughout the project work.

Sachin Kumar

Roll No. 2K15/ISY/15

M.Tech (Information System)

Department of Computer Science and Engineering

Delhi Technological University

Abstract

Emotions are the main influence of human behavior. Emotions can be defined as an opinion or feeling, a neutral feeling or an attitude towards something. The sense test is a method of computationally recognizing and classifying the opinions outlined in a piece of text, especially to determine whether the authors are positively, negatively or neutral towards a careful topic, product etc.. Inspiration of the senses uses data mining techniques, extract and capture natural language resources and text classification data and analyze it over here. Our focus here is to make sense analysis on twitter data. Millions of Tweets are posted daily on microblogging, which basically reflects the thoughts and feelings of users of the world. These tweets can be present in different languages. Our goal is to make sense analysis in many languages. Therefore, we propose a system that uses the Google translation API so that tweets can be converted into English language in many languages and then be able to analyze the feelings about it. Before translation, we must prioritize the Tweets to filter useful information from raw data. Multilingual cement analysis has been done using different classification algorithms. These algorithms are basically performing well, but when we apply deep learning, we have seen all of these algorithms so far. Our proposed methodology is used to describe multi-lingual sense analysis, using multi-lingual sense analysis more efficiently and accurately, with the accuracy of 63.5% and 73.5% respectively, using the naive Bayes and Max Entropy algorithms, the multilingual signal analysis has been described. And our deep neural network provides the accuracy of 93.5%.

Content

List of Figures

Chapter 1
Introduction

Social Website Twitter is an online news and social networking service where users post and interact with communication, called “tweets”, have 140 characters..Twitter is a social platform where share their views/idea toward the any social, religion context , share their opinions, and find out what is happening in the world right now. A single opinion from one person may not seem important. [1]But among millions of person, opinion through tweet can form a comprehensive picture. Sentiments are the main influence of human behavior, it can be a neutral feeling or an opinion expressed. Whenever we want to take decision regarding a product or a thing we first seek users opinion about that particular thing. Sentiment expressed towards a product or a thing can be useful both to the business organizations and the users too. Business organizations are keen to find out the users opinion regarding their product through users review and feedback. In the similar way users look for the reviews and feedbacks given by exising users before buying that product. [2]Twitter is a microblogging site, where millions of people share their views, opinions or feedback regarding a product or thing. Our aim is to do Multilinguil Sentiment analysis over twitter data.

1.1 BACKGROUND

Microblogging has become the most common and powerful medium among internet users today. Some of the commonly used microblogging sites are Twitter, Facebook, Tumbreck In these sites, users write about their lives, share their opinion with something like a film or product and also discuss current issues. [1]Due to the ease of use of an independent format and the use of these microblogging websites, Internet users have shifted their interest from traditional communications devices such as traditional blogs or mailing lists to these microblogging services. Since users who use microblogging sites have posts about their products or services, or they express their political and religious views, these sites are therefore a source to know about people's feelings or viewer topic. Are formed. Since the number of microblogging presentations and services increases every day, the statistics of these sites can be used in assessment mining and emotional testing. In order to get the answers to the following questions, companies can analyze the price of their product-What do people believe about their product ?

1> How positive or negative people are regarding our product ?

2> What do people favor our product to be like ?

1.2 MOTIVATON FOR THESIS

The sense study is also called as opinion mining, a form of information extraction from the increasing research and lessons of commercial interests.[3] The analysis of the emotion of the text is important for many natural language resource works. Especially with the development of social media, through the emotional analysis, there is a great need to lose useful information on the internet with large data. Big data is a term that considers data too large and complex that the traditional data processing work is inadequate . [4]It will be of 3 v i.e. Volume, velocity and diversity are also defined by. From Volume we mean that the data collected from many sources is very heavy, we mean the unprecedented speed of data. It is necessary to handle within the appropriate time frame. Built-in data comes in a variety of formats, which are unstructured, multilingual, email etc. Millions of Tweets are posted daily in which the feelings and thoughts of users around the world are included, for example, movie reviews or product reviews. There are more than one language in tweets, so multi-lingual is facing the biggest challenge.[1] Our goal is to write a program that automatically reads the reviews written in different languages and finally gives results that the product is good The reason for implementing the machine learning is bad or because there are millions of reviews about the customer or film and we cannot do it manually, It will take millions of years to do. So we have to go for different classification algorithms. There are so many proposed algorithms and they are doing well. But when you apply a deeper education algorithm, they actually overcome all the algorithms that we have seen so far.

1.2 THESIS OUTLINE

The main goal in the first part of the thesis is the purpose of which the proposed work is done according to purpose, which is done earlier in the field of multi-level sense analysis, thus the weakness is described which is present in the work presented earlier. . The next part includes the proposed method in the thesis, where we define the purpose of our system. We have a brief introduction about the neural network, while explaining the main concepts of neural networks and artificial neural networks. The next part gives a theoretical description of the phases of emotional analysis and then after the details of practical implementation. In the next part the results are obtained. It consists of three parts, the exactness received and the subsequent accuracy was obtained. The next part is compared with other algorithms already implemented. After all, we have concluded the thesis with the final conclusion and future part, which can be done in this.

Chapter 2
Literature Survey

In 2011, (Shi and Lee, 2011) developed a supervised machine learning technique to test the spirit of online reviews in English using Unigram feature technology. To be positive and negative, the facilities used to classify polarization of documents and TF-IDF were the difference in the form of training and test set with different data examples. Examples of training data sets were added to target values of data instances. SVM is a model that predicts the target values of data examples. It is chosen as a better performance than any other classifier. The review includes 4000 positive and negative reviews, reviews have been labeled as preprocessing and positive and negative. Sentiment classification model was then used to classify the sequential flow as negative and positive. TF-IDF was considered a better feature than the event.

In 2009, (Boi E, Mons M-F, 2009) another study was done which used the supervision classification technique for the recognition of emotions in a document .. applied to law sentences, which are reviewed, blogs And found in films. The figures obtained in these blogs are in full form and before the classification algorithm can be implemented, they should be processed first. After preprocessing, unrelated features are selected which include Unigram, Stems, Prohibition and Speech Features. For classification multimodal kinase biis, vector machines and maximum entropy were supported. For different data linearly, the SVM classification algorithm gives a minimum error rate. The English language corporation was created from blogs, reviews and meeting sites such as www.livejournal.com and www.skyrock.com. The maximum entropy showed 83% accuracy, in contrast to the extra classifier used in this learning.

In 2010 (Barbosa et al, 2010) has given another direction, in which direction is shown in the Twitter data, they collected 1000 tweets from three different websites and labeled them positive, negative and neutral . They made another 1000 tweets for the test. The characteristics used to identify the sentiments of tweets include the way in which it was written and the words which included the tweets. In addition to the polarization of words and POS of words, they use syntactic features such as actuators, hashtag punctuation marks and exclamation marks. The accuracy gained due to the abstract representation of words rather than the raw representation of words was very high. And the better quality labels provided by the mass data sources

In 2011 (Ghorell H and Jacot T, 2011) planned an approach for French film reviews. In this method, three types of characteristics were used, i.e. literally, meaning and morphosyntactic an unigram was chosen as a feature. The goal of this system to find the division of words was also used in addition to the speech tag. In the sequence, it will be used to select the overall division of the evaluation. Scientistnet is an English language that is preserved. To use SentiWordNet,

French reviews were first converted into English before the return of polarization. The movie review contains 2000 tweets in the dataset, 1000 of which positive and 1000 out of which 10 movies have negative. SMM classifier was used for this classification, before it can be used in SentiWordNet, the word needs to be translated into inconvenient English language. The quality of change has a negative impact on the classmeter's performance. As the conversion of words does not protect the direction of meaning due to the differences between languages.

In 2012, (Balakrishnan Gokulakrishnan, 2012) proposed a model for emotional classification where he used a publication on the Elephant Dataset, which contained 53000 tweets, containing about 41,000 positive tweets and 8500 negative tweets. Tweets were made for testing and then pre-processed to get critical figures of raw data. Then the different classification algorithms were functional in addition to the sense analysis. They were Kannai Breeze, Random Forest, Sequential Mining Adaptation (SMO) and Support Vector Machine (SVM) and were set accurately. It was seen that compared to SMO, SVM and Random One, according to the DMNB show, Kannai twenty two classifier was better.

In 2013 (Guan and Rang, 2013) had planned a method for Chinese examination on Chinese microblog. The reason was to classify opinions in positive or negative in Microblogs. The required preprocessing steps for the method such as word segmentation and noise removal classification require the removal of specifications for every communication and self-training, TTOcActualAnulebell data was used. A method for semi-supervision is self training, where jointly labeled and unlocked data is used as training corpus. First of all, this training label starts with data; When the iterations are introduced, it is able to organize the unenilabeled data with the labeled data, compared to the supervision education algorithm, the self-training for classification of classification is not the best performance.

In 2015, (Rui Jiayet., 2015) proposed a novel data expansion approach, which is called Dual Sense Test (DSA), which talks about the problem of division for emotion classification. The main reason behind this step was to create back-up reviews which are contrary to the original reviews of emotions and then using the pair of reversed and unique to train the spirit classifier (dual training); In order to predict further emotions (double prediction), he further clarified that in the case of polarization classification compared to other classifiers, the DSA algorithm supports the preparation of effective DSA and a selective data expansion technique, With high degree of degree required to expand the data along with a much more efficient performance. He has also worked on the expansion of the DSA to the DSA, which can deal with the 3-class (positive, negative, neutral) emotion classification. Finally, to overcome the dependence of the DSA on the External Antonyms dictionary, they suggested a corpus-based technique for the creation of a pseudo-Antonim dictionary.

Chapter 3
The Proposed System

With the development of the World Wide Web, the increasing quantity of data has increased rapidly, people express their feelings about different topics. For any company related to a brand on the Internet, its product Very important for quality so it makes the company important for understanding the feelings of the people. Many approaches to products or services have been proposed to classify emotions expressed in various channels such as Twitter, blogs and user comments. There are huge data generated by millions of users on Twitter as opinions, opinion and comments. Processing this raw data can be a very challenging task to extract useful information. Sense analysis is done in two stages in tweets-

1> Identify expressions

2> Determine the polarization of sentiments expressed in tweets.

In many ways, many microblogging sites such as Twitter, Facebook etc have been classified as emotions, the importance of current emotion analysis systems is usually the same language, English is connected around the world.[5] To write comments in different languages in detail, sense analysis in one language is written in other languages. The purpose of the proposed method is to analyze tweets in various languages, to analyze emotion and classify them in the form of tweets, positive, negative or neutral blocks of the proposed methodology used by the neural network. Shown below .

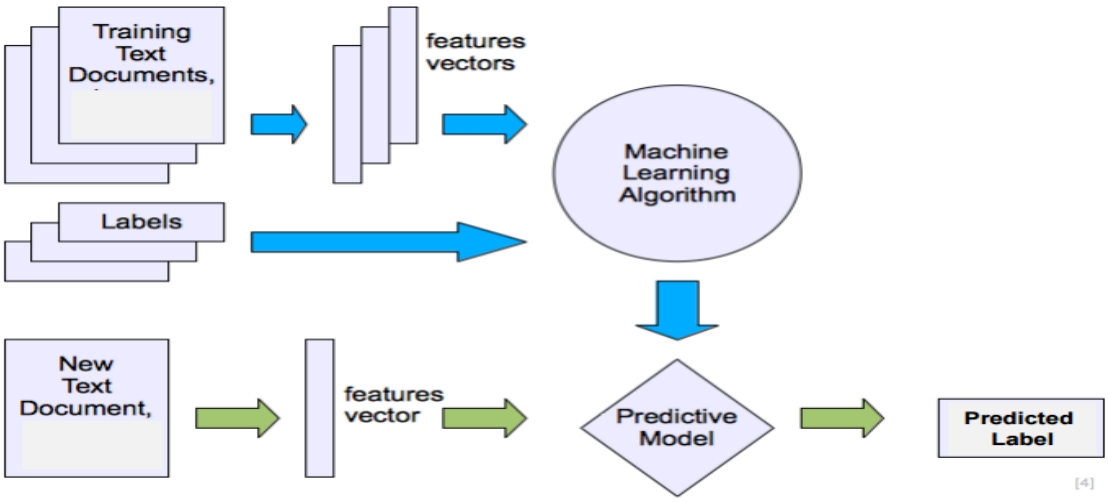


Fig1 : BLOCK DIAGRAM OF PROPOSED METHODOLOGY

The main purpose of this method is to propose an effective way to filter out valuable information from twitter tweets and to identify the emotions made by these tweets. For this purpose, we have proposed an algorithm that allows the reader to respond positively, Negative or neutral is called neural network to be classified. For this purpose, we need to have two data Training and testing with a 70:30 ratio of training, we are bringing data using Twitter API, this figure is usually present in raw form, which requires it to be processed so that it contains The next thing that we are doing is to pre-process the data, in which stop words, URLs, hashtags, slang words, spaces, repeated repetitions. Its etc. involves the removal. The processed data is then converted into numerical vectors and we analyze the feelings using neural network algorithm by these vectors are referred to as features.

We use neural network approach to train the system and some of the features of the neural network approach are

Imagine a computer with thousands or even millions of nodes. Some areminputs connected to others in the network and finallynare the output ‘neurons’ or variables. For a ‘signal’ to pass from one neuron to the next the inputmto the neuron must match what the neuron is looking for if it is the right input it will pass through the neuron to the next. In this way patterns are formed between inputs and outputs. Let's take an example of speech recognition or ‘speech to text’.

In this example the inputs to the computer variables are frequencies of a spoken word. The different neurons will respond only to certain frequencies. Thus the output will be unique for each word. We then train the net by identifying the output related to the input word. So wjen that wors is spoken the compute has learned it. Exams a certain word. Thus with the neural net and training the computer can now recognize human words.

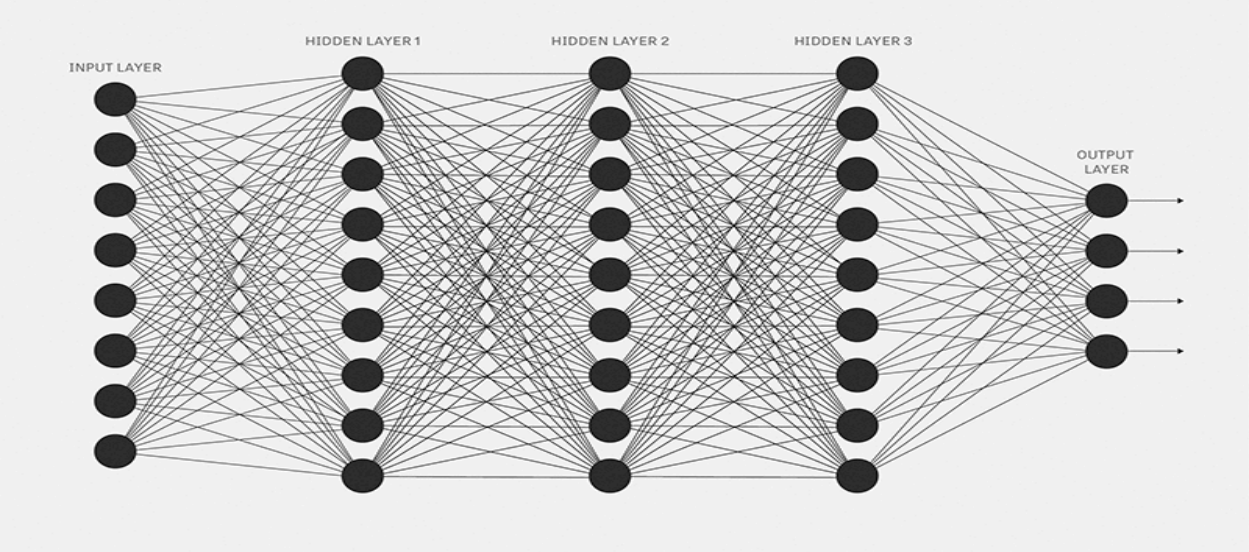


Fig 2 : Deep Neural Network Model

each node decides what to send on to the next tier based on its own inputs from the previous tier -- neural networks use several principles. These include gradient-based training, fuzzy logic, genetic algorithms and Bayesian methods. They may be given some basic rules about object relationships in the space being modeled. For example, a facial recognition system might be instructed, "Eyebrows are found above eyes," or "moustaches are below a nose. Moustaches are above and/or beside a mouth." Preloading rules can make training faster and make the model more powerful sooner. But it also builds in assumptions about the nature of the problem space, which may prove to be either irrelevant and unhelpful or incorrect and counterproductive, making the decision about what, if any, rules to build in very important.

Here I use google API for the translation of the tweets . It was basically its own version of a Neural Machine Translation system. The Google Neural Machine Translation (GNMT) system tremendously improved the efficiency of apps like Google Translate. They had kicked off their service with German, French, Spanish, Portuguese, Chinese, Japanese, Turkish and Korean in 2016. This year, they launched the tool for Russian, Vietnamese and nine Indian languages. Let us take a look at what this new approach of translation entails.

Neural Machine Translation (NMT) systems existed before Google's version. These are machine learning models that use neural networks to predict the outcome of a particular problem. These predictions arise from similar patterns that the system had 'memorized' earlier.

But these NMTs were expensive for training the network as well as for applying it to actual problems. Google's NMT made the process of learning easier by deploying parallelism. Low-level arithmetic computations increased the output translation times quite impressively. Also, special conditions for the interpretation of rare words helped improve the accuracy of the approach.

Earlier, Google used Statistical Machine Translation to translate languages. But it had its limitations as it used a "phrase-based translation." Therefore the translation wouldn't work for language pairs with significantly different word order. Also, the linguistic difference between the source and target languages made the translation a complex task. For example, English and Hindi, although both belonging to the Indo-European language family, possess enough distinctive characteristics in word order and morphology.

But with the new state-of-the-art neural machine translation, translation has become easier, more accurate and comprehensive as it translates whole sentences at once, rather than doing it by phrases.

Chapter 4
Introduction to Deep Neural Networks

Deep learning (also known as deep structured learning or hierarchical learning) is to present the knowledge of artificial neural network (ANN) to the counterparties, which keeps more than one hidden layer. [5]Unlike the deep learning task data algorithms, knowledge Based on data representations, the machine is part of a broader understanding of learning methods. Natural language processing is a field which provides us the opportunity to tokenize documents and extract patterns to better understand the structure, sentiment, polarity, style of writing, contextual information, and much more. It's a relatively early field but patterns can be extracted from text to:

- predict the way constituents will vote in an election
- to recommend results in a search engine
- provide summaries of articles, etc.

When you throw deep learning into the mix you are strengthening the initial observations from the pre-processing of your natural language document term matrix/corpus and allowing the opportunity for better understanding of new bodies of text based on classification, similarities, distance measurements and more.

Machine learning methods have classically relied on features selected or created by humans. However, for some tasks this hasn't worked too well.

Deep learning is all about making your model create the features it needs to fulfill its task. In other words, you want your model to represent concepts as a hierarchy: lower level concepts like lines and corners build up to squares and circles, and those build up to even more abstract concepts like cars and billboards.

In practice this is often implemented using multiple layer models like deep neural networks. The layers are simply representations of lower to higher level concepts.

The difference between classical models and deep learning, is that deep learning refers to a way of building hierarchical representations of data, whilst classical models concentrate on a more straightforward task: use features to classify.

Deep Education is a class of machine learning algorithms that:

• The properties or representations of different levels of data (uneducated) are based on learning. High level characteristics are derived from low level characteristics so that there is a hierarchical representation.

• There are some parts of the field of comprehensive machine learning for presenting data learning.

• Representing many levels according to different levels of separation; Level build hierarchies of concepts.[12]

In these definitions, there are several layers of general (1) non-liner processing units and (2) the information of the requested or unincorporated learning in each layer, with the layers, creates a hierarchy of low-level to high-level facilities.) [15]The composition depends on the problem of solving a layer of non-liner processing units used in a deep learning algorithm. The layers used in deep learning are Ritrim set of sources of hidden layers and complex presentation of neural network ... they deeply generating models such as endangered variable levels can also include such deep faith networks and nodes in Deep Boltzmann machines [17].

Deep Learning was first designed and implemented by World School Council London, which uses algorithms to change its input through layers more than the shallow teaching algorithm. On each layer, the signal is replaced by a processing unit Is, such as artificial neurons, whose parameters are adjusted through training.

• Credit Assignment Path (CAP) - A series of changes in output from inputs describe the possible connection between the CAP input and the output connection.

• Depth of the cap - For a feed forward neural network, the depth of the CAP (this type of network) is the number of hidden layers and one (also parameterized as the output layer), but for the recurrent neural network, in which one Indications can be promoted through one layer more than once, the CAP depth is probably unlimited[19].

• Deep / shallow-deep learning is not universally agreed on in-depth learning, but most of the field researchers agree that there are many nonlinear layers in deep education (CAP> 2). Schmidburg understands CAP> 10 to learn "too deep

4.1 Fundamental Concept

Distributed representations have the underlying assumption that observed data is generated by interaction between layered factors, deep education is the assumption that these layers of factors combine consistent levels of abstract or composition, variations of the layers and layer shapes differ in abstract It is this idea of a deep-education-oriented level, where lower level people should get higher levels and more deep academic architecture is often a greedy With a layer-by-layer method, you can take advantage of this, deep learning helps in solving these absorption and facilitating improvements in improving performance, whether giving birth or learning. Useful, deep learning methods feature Use engineering, [6] the data is compact like a compact intermediate main component. The level of structures that translate and receive while reducing the redundancy representation learning how to learn algorithms can be applied to learning unchanged learning tasks. This is an important advantage. Ecosis label data is more abundant than the examples of deep structures that can be trained in an unexpected way. Unlock from, these are the nerve Ihas compressors and the trust network.

4.2 Artificial neural network

Artificial neural networks (ANNs) or connection systems are systems induced by the biological nervous system that constitute the animal brain. Such systems (in order to improve successfully), with examples of working, usually learn without work-specific programming. Image recognition, they can learn to identify pictures in which the pictures are analyzed by the bill, allegedly, "cat" or "no cat" is said to have been used and analytical results are used to identify cats in other images. The use of traditional computer algorithm has been used more and more. ANN, the collection of connected units Between half neurons, each connection (synapse) can transmit a signal to the neuron, which is called artificial neurons (in the biological brain, the axis The receptor (postsynaptic) can process the neuron signal (PRO) and then point to the downstream neurons associated with it. The state may be in neurons, usually indicating the actual number, usually 0 Between 1 and 1 Neurons and surgery may also have a weight that differs in learning form, which can increase or decrease the signal strength, which sends it down. In addition, they may have a limit, only when the total signals are below (or above), the level downward signal is sent. In general, neurons are organized in layers, different layers can make different types of changes on their input. To take the signals from the first (input) to the last (output) layer, possibly after taking the layers several times, the basic goal of the approach of neural networks was to solve problems in the same way that a human brain is focused on time, specific Corresponds to mental abilities, thereby distracts from biology in biology, or goes backwards in the information direction and adjusts that information Ratibimbit to the network [5]

Deep Learning = Artificial Neural Networks + more of everything

4.3 Deep Neural Networks

deep learning is a fancy new term for multi layer neural networks, with the difference now being we have found efficient ways to train these deep (5+ hidden layers) neural nets that wasn't possible before.

Deep learning is possible due to three areas of advances, hardware, techniques and data. Hardware being gpu computing (not really popular before Jeff Hinton in 2006), and techniques being better weight initialization from unsupervised techniques such as denoising autodecoders. And of course, we now have large public datasets (image net, etc) and even more private data (think Google, uber, quora) which are what deep neural nets need as they have some many parameters to tune.

Same "deep" architecture can be used for other variants of neural networks such as ones with recurrent connections, such as the popular LSTM. This similar stacking fashion of deep LSTM networks is leading breakthrough research in modelling highly complex sequential data, such as speech recognition and machine translation.

4.4 Natural Language Processing And Neural Networks

At its core, deep learning (and neural networks) are all about giving the computer some data, and letting it figure out how it can use this data to come up with features and models to accurately represent complex tasks - such as analyzing a movie review for its sentiment. With more common machine learning algorithms, human-designed features are generally used to model the problem and prediction becomes a task of optimizing weights to minimize a cost function. However, hand crafting features is time consuming, and these human made features tend to either over-represent the general problem and become to specific or are incomplete over the entire problem space.[14]

Chapter 5
Phases of Sentiment Analysis

The Sentiment analysis is basically a step-by-step process which consist of the following phases comprising-

1>Data collection i.e. basically collection of tweets

2>Preprocessing of tweets

3>Dataset generation i.e. basically testing data set & training data set.

4>Sentiment classification using ANN into positive, negative and neutral.

Following is a block diagram of the entire process-

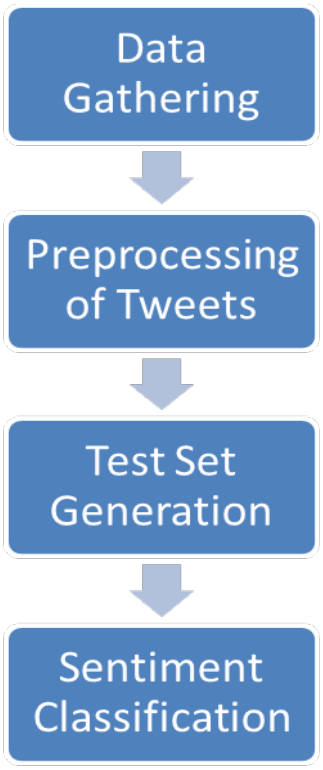


Fig 3- Block diagram showing Sentiment Analysis Phases

5.1 Data collection in twitter

"We collect tweets from twitter to analyze emotion on twitter data. We can collect tweets using the tweeter api i.e Tweepy Master. To install Tweepy, you must first install" pip ". The following is a snippet The use of PP, which reflects the establishment of the Tweep. The TPE, which are collected, are in raw form, which have proceeded for further steps.

5.2 Data Processing

The numbers collected so much have a large amount of generally ambiguous information, which is necessary to filter on the first priority.

1> First of all, do not express any feelings like unnecessary tweets such as tweets, replies and tweets, stop words and hash tags, URLs, disturbances, spaces, misspellings.

2> Apart from this, we have seen that the same word and characters are repeated several times in tweeting so that it can be more trendy. Words like Küduüüllü, Guadüuüüd come under this category ohhohhh

3>. This continuous repetition of letters should be removed only in this phase.

4 Apart from this, there are some words in the English language that do not actually contribute to any kind of emotion which is called the stop word which is also removed at this stage.

5> Apart from this, here we are working with multilingual facility so that all tweets are translated into English using a Google translator. This study will use the N-Gram model[11]

The language we are working on is dot-net. It has open source natural language toolkit (NLTK). This library is a rich set of natural language resource tools and datasets, whose purpose is for educational purposes. For example, the combination of NLTK and dot net makes the right combination for teaching NLP techniques[20]

5.3 Algorithm for Data Preprocessing Module-

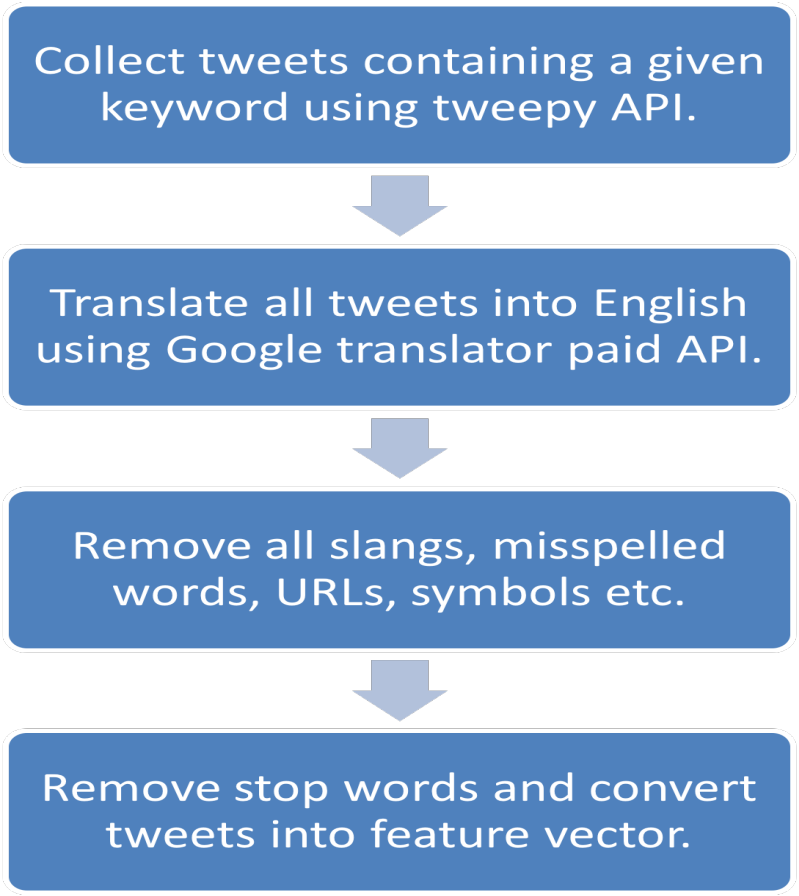


Fig 4 : Block diagram for preprocessing of tweets

5.4 TRANSLATION OF MULTILINGUIL TWEETS

After collection of Tweets using Twitter API i.e Tweepy, we collect all the tweets in a CSV file. There may now be tweets in different languages like Hindi, Urdu. To analyze the emotion, we need to convert it to English language. For this purpose we will use the Google Translator API for this purpose. We will use Google Cloud Translate which is a Google translation API. This is the payment API. Since we are using dot net for our project, dot net has a library as Google Cloud Translation. We must first establish the library[10]

5.5 Feature Extraction

After the preprocessing phase, we will remove the features from it and we will use various combinations such as UniGram, POS tagging, emotional words (happy, tragic, good, bad etc), emotional mark, punctuation, word of protest, excitement and try will do. Latitization etc. Characteristics are nothing but only recommended words, which are to be converted into further numerical values, which are called feature vectors, which are then fed as input of machine to learn. [7]

In the case of neural networks, all values are displayed as 0 and 1. For neural networks, you will use Word 2 VK for numerical values for the conversion of these features. The WORD2 AVC is a group of related models used to make the word embedded. These models are shallow, two-level neural networks that are trained to reproduce the linguistic references of the words

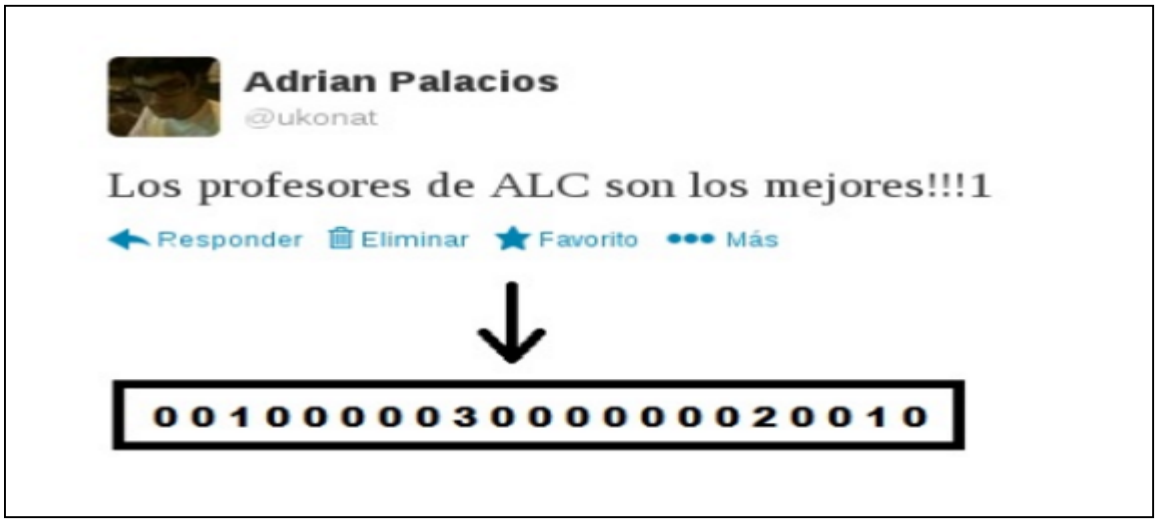


Fig 5 : Conversion of words to numerical vectors

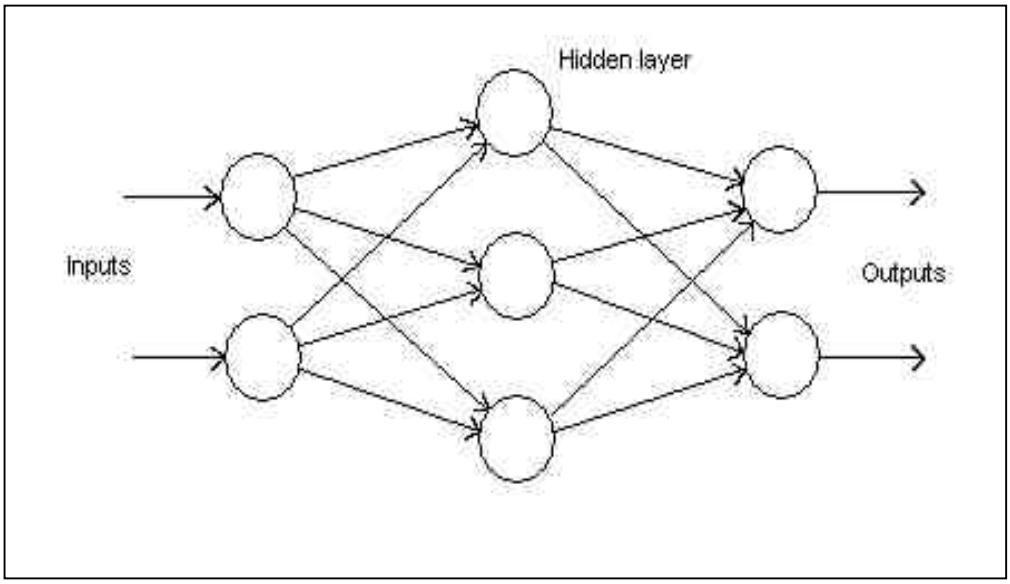
Word2vec takes its input as a large part of the text and usually produces a vector space of several hundred dimensions, where each specific word is assigned to the same vector in the space in the corpus. Words are vector vectors positioned in space such as words that share common references in the corpus, are located in close

5.6 Apply Classification Algorithm

The Backpropagation algorithm was originally introduced in the 1970s, but its importance was not fully appreciated by David Rummelhart, Geoffrey Hinton, and Ronald Williams in a famous 1986 paper. Today, Backpopagration Algorithm in Neural Network The scope of education is the area of neural networks, artificial intelligence, machine learning, parallel processing, relationships with data and other areas. Of being can be thought. The attraction of neural

networks is that they are most suitable to solve the problems that are the most difficult to solve in a traditional computational way ".[14]

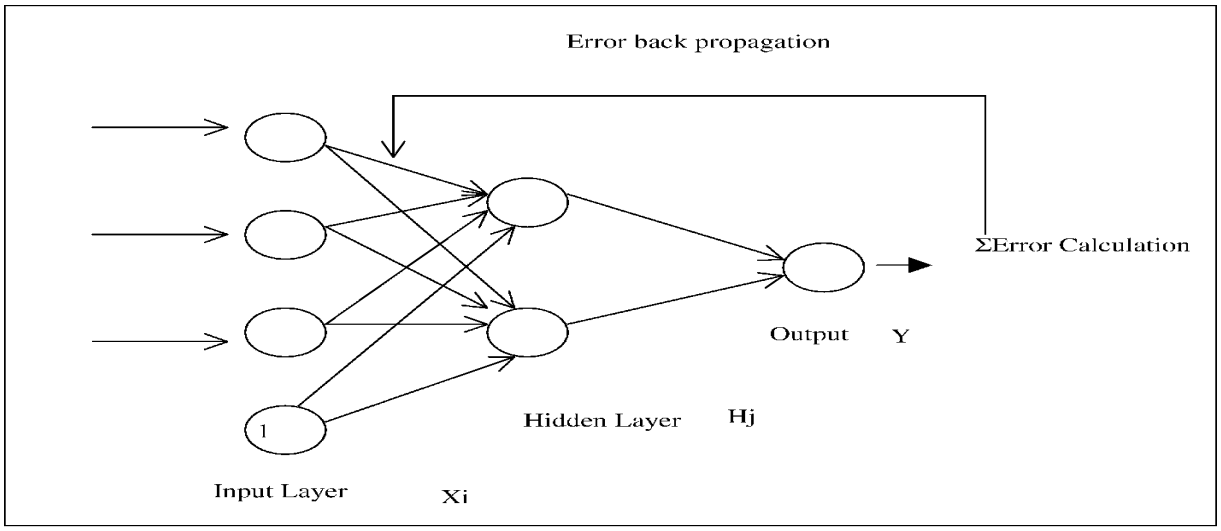
If we believe that the human brain is the 'ultimate' neural network, then ideally we want to make tools that mimic the brain's functions. However, due to the boundaries in our technology, we have a very simple design. Transfer functions similar to a biological neuron designed to design an electronic device, and for each neuron connector to neurons. Using an RLC network to duplicate dendrakes, exon and syncs, such an electronic model is not too complicated to apply, and we have difficulty in teaching 'network' so that the design is useful in making the design more manageable and obstacles First of all, change the connectivity between the neurons first, so that they can separate the layers, in this way each next step In addition, we have defined that the signal flows in the same direction across the entire network, and we compare the comparison of other neuronstruments, analog resistors, to simplify the neuron and synapse design, powered by To do now we have a feed-forward neural network model that can be really practical to "build and use



The above figure displays a normalized network, stimulation is applied to the inputs of the first layer, and the signal spreads through the (hidden) layer (the output layer). Each link between neurons has a unique load value[2].

The structure of the neuron is different from one or more pre-neurons, then it is expressed. The result is that there is no similarity between 0 and +1, and output values are given to neurons in the next layer.

Since the actual specificity of the network or reality exists in the value of the weight between neurons, we need a method of adjusting the weight to solve a particular problem. For this type of network, the most common learning algorithm is called Back Promotion (BP). A BP network learns by example, that is, we have to provide a teaching set, which includes some input examples and known output of each input. Therefore, we use these input-output examples to show the network what kind of behavior is expected, and allows the BP algorithm to optimize the network.



Notes: The weight connecting node *i* in the input layer to node *j* in the hidden layer is denoted by *Wji*, and the weight connecting node *j* to the output node is represented by *Vj*

Fig 6 : Working of Backpropagation Algorithm

BP learning process works in small running stages: For example, a network is applied, and the network produces some output based on its current state, which is synchronous load (initially, the output will be random) . This output is compared to known-good outputs, and the average class error signal is calculated. Then the error value is spread backwards through the network, and small changes are made in each layer. Changes in weight are calculated to reduce the error signal for the case in question. The whole process is repeated in the case of each instance, then again the first case again, and so on. The cycle is repeated from some predetermined threshold until the total error value is reduced. At this time we say that the network has detected a "good" problem - the network never learns the ideal function, rather it will reach an idealist approach to an incompatible form.[8]

5.6.1 Backpropagation algorithm training process

There are two ways to learn to choose: In the stochastic and batch stochastic learning, the weight of each reproduction is updated promptly. Before accumulating errors in the batch above the samples in batch above the samples, many promotions are obtained before improving the weight;

Stochastic learning, introduction of "noise" in slope descent, using local shield created from the data point; This reduces the possibility of a network trapped in the local network. Still learning a batch produces a fast, more stable line at the local level, because each updated batch is directed towards the average error of samples. A common compromise option is to use "mini batches" in modern applications, which means learning batch, but with small size batch and with stochastic selected samples.

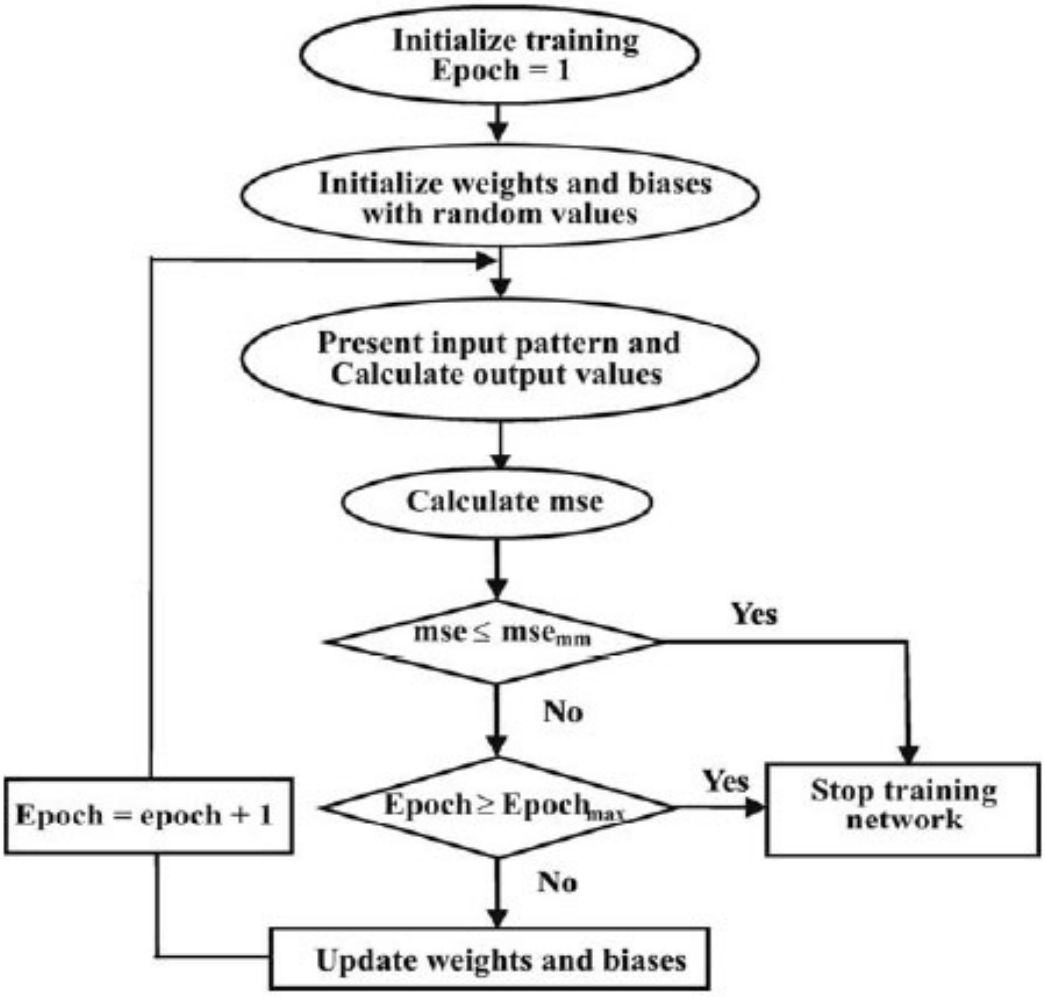


Fig 7 : Block diagram of training process of backpropagation algorithm

5.6.2 Error Assessment Function of Back propagation Algorithm

Consider the feed-forward network with N input output units. It can contain any number of hidden units and can display any desired feed-forward connection pattern. We will also get a training set. ({1 (x1, t1)) ..., (XP, TP)}, together with the commands of P, N and M-dimensional vectors, which are called input and output patterns, on each node of the network. Let the constant work be constant and separate. The actual number of edges is chosen at random when the input pattern is presented in this network from the training set, then this target typically produces an output OI from TI. We want to make I = 1 equal for Oi and TI ... using P., P, using a learning algorithm. More precisely, we want to reduce network error function.

E = 1/2 Σ_ (i = 1) ^ paa_ (i) -t_i ‖ ^ 2

After reducing this function for the training set, the new unknown input pattern is presented in the network and we expect it to be interpolated. The network should understand that a new input vector is similar to the learned pattern and generates a similar output .

Back propagation algorithm is used to find at least the error function locally. The network is started with randomly chosen weight, to calculate the gradient of the error function and to fix the initial load. Our job is to calculate this gradient .

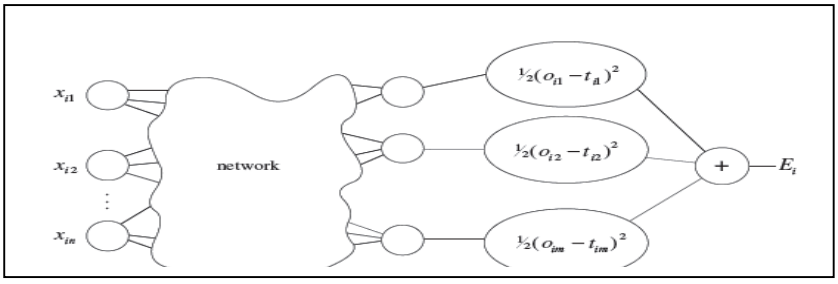


Fig 8 : Extended network for the computation of the error function

The first step of the reduction process involves expanding the network, so that it automatically calculates the error function, the above figure shows how this is done. Each of the network's J Output units is connected to the node which is function 1 / 2 (oz-tease) 2, where oz and the remaining output vectors represent the J-th component of the OE and the target T. No output of the extra meter nodes Is collected on D, which adds them and gives the amount EI as its output. The same network extension should be made in the form of each pattern ti. A computing unit collects errors of all classes and their Zodiac E1 + • • • Ep + The output of this extended network is error function .

Chapter 6
Experimental Setup

In our case, we will use the neural network ie back propogation to train the system. We are basically using dot net for our project, so in the dot net we have to implement the neural network in dot net Nimranet Library. It basically works in the concept of hidden layers which act as intermediate layers between input and output The more hidden layers we use, the more accurate it is. When we only take a hidden layer, it is said to be learning as a machine, but when we take two more hidden layers, In our case

There are 3 steps in our entire experiment-

6.1Preparing the Training Data Set

Our training data set basically all the type of English language modeling and library. This library is provided by the standardford University. For rich dataset and better accuracy this is the best option which one can try for sentimental analysis project .

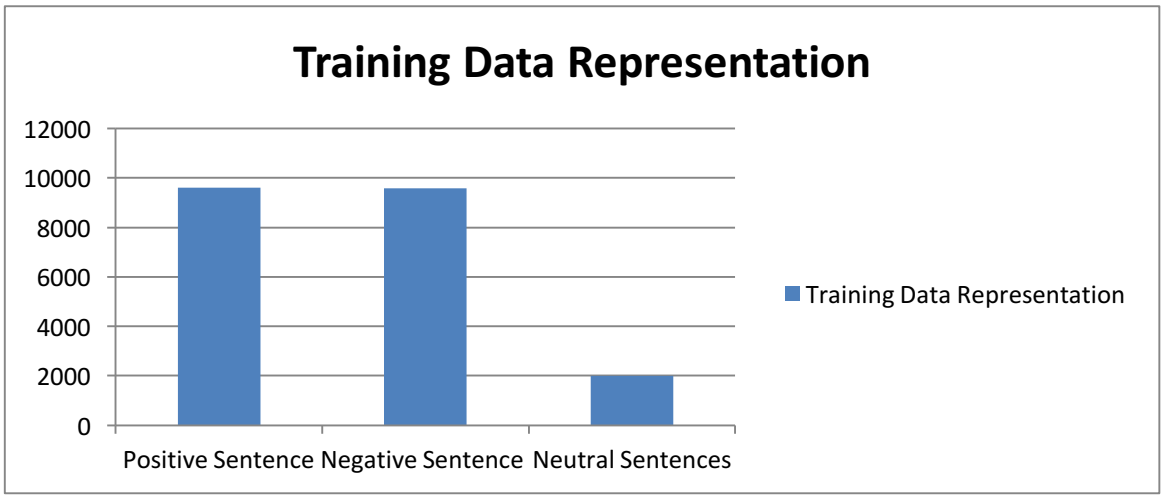


Fig 9: Training Data Representation

6.2 Accuracy Estimation

When we use the Nimbernet Library and import it into the dot net to create our neural network, and then use the back propogation algorithm to train our network using the above training data set of 21604 tweets, So we receive our accuracy as 93%. Basically divided into our training data set in the ratio of 70:30. 70% of total tweets are used for training and validation and 30% tweets are used for testing.

Backpropage Neural Network Algorithm is applied to the training dataset after the above, when the following diagram shows the accuracy.

Hence on training the machine using the Backpropogation Neural Network Algorithm we get the accuracy which is approximately equal to 93%.

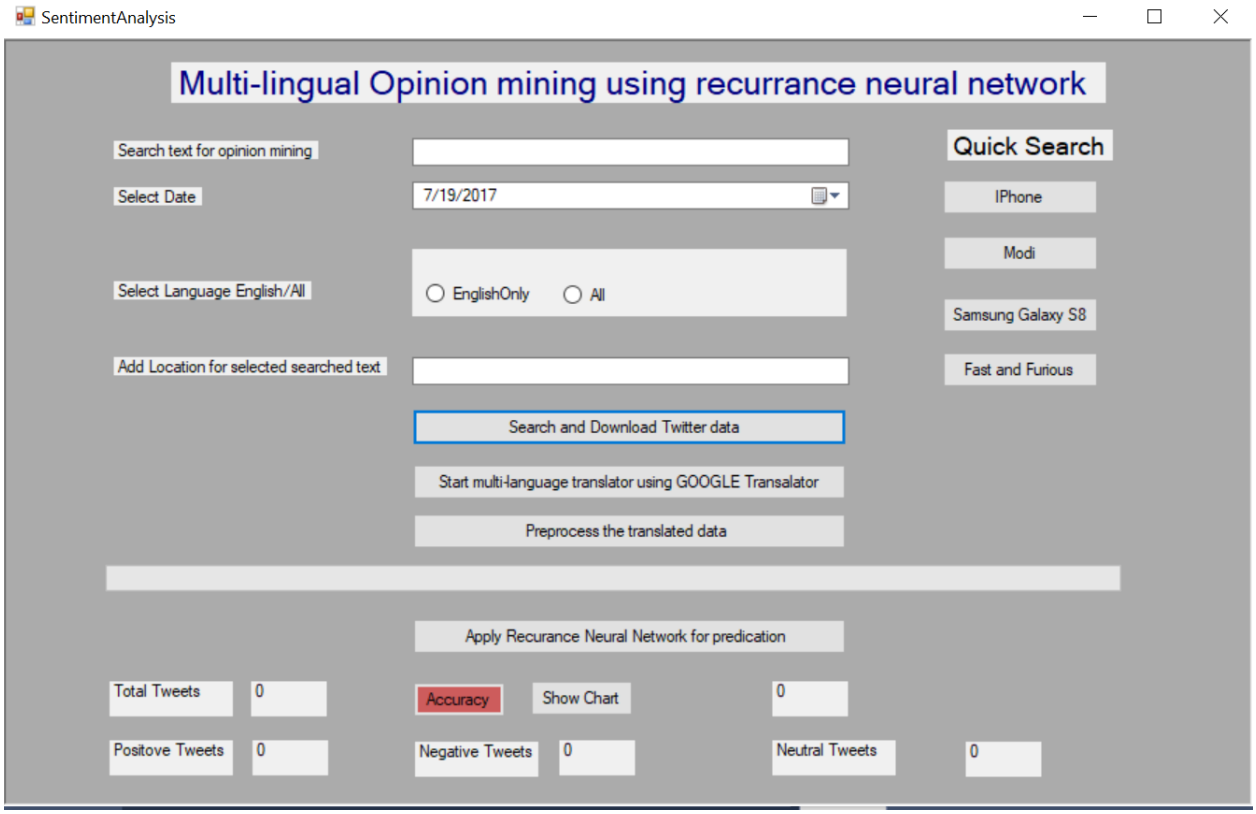


Fig 10 : Initial Screenshot of the output

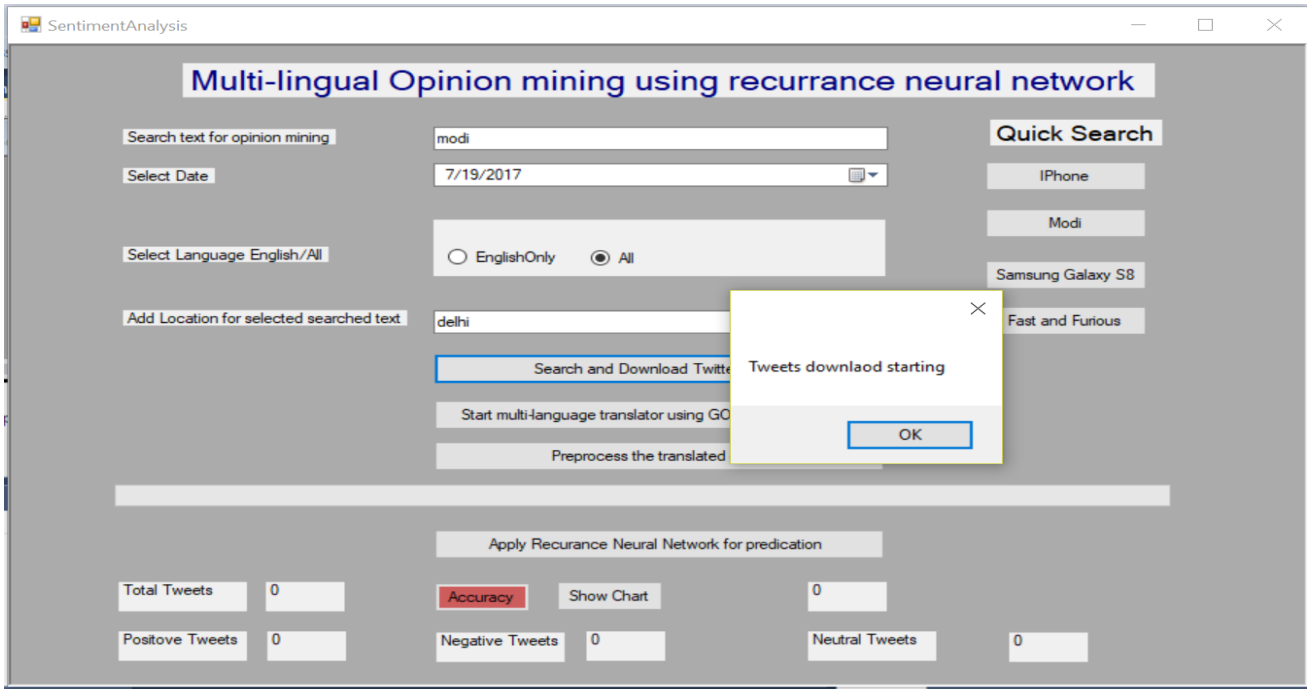


Fig 11 : When tweets start downloading

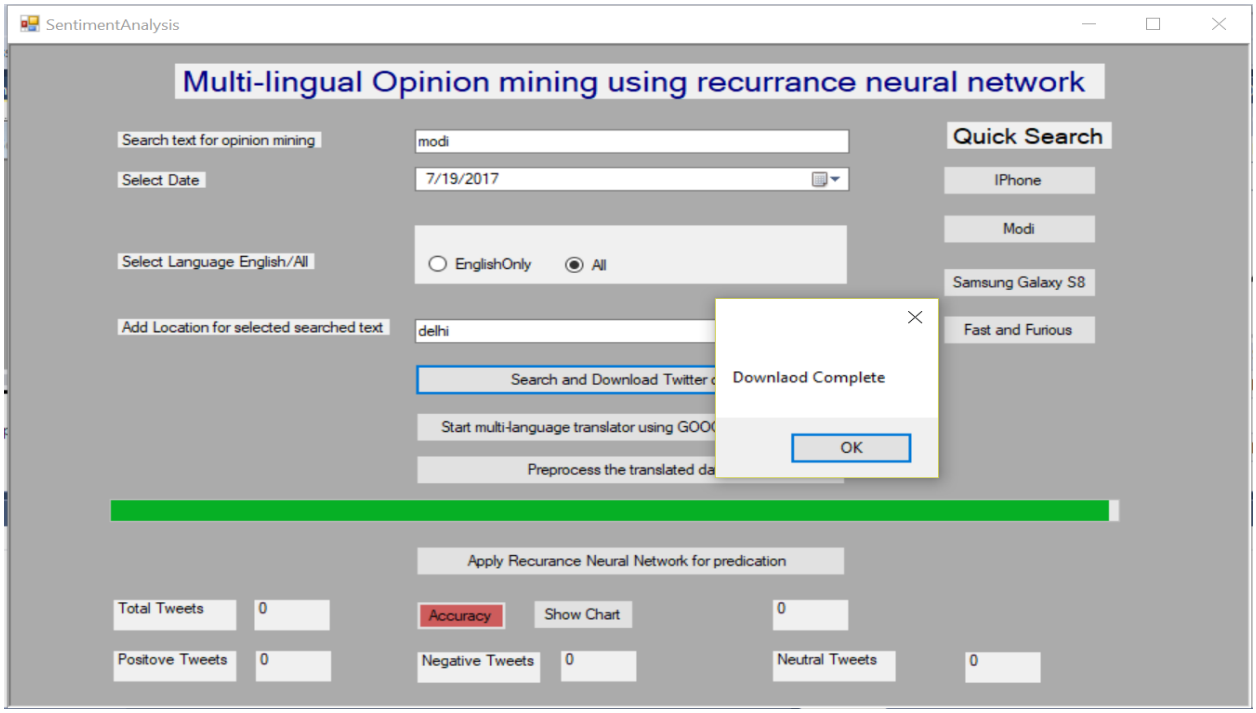


Fig : 12 When tweets finish download

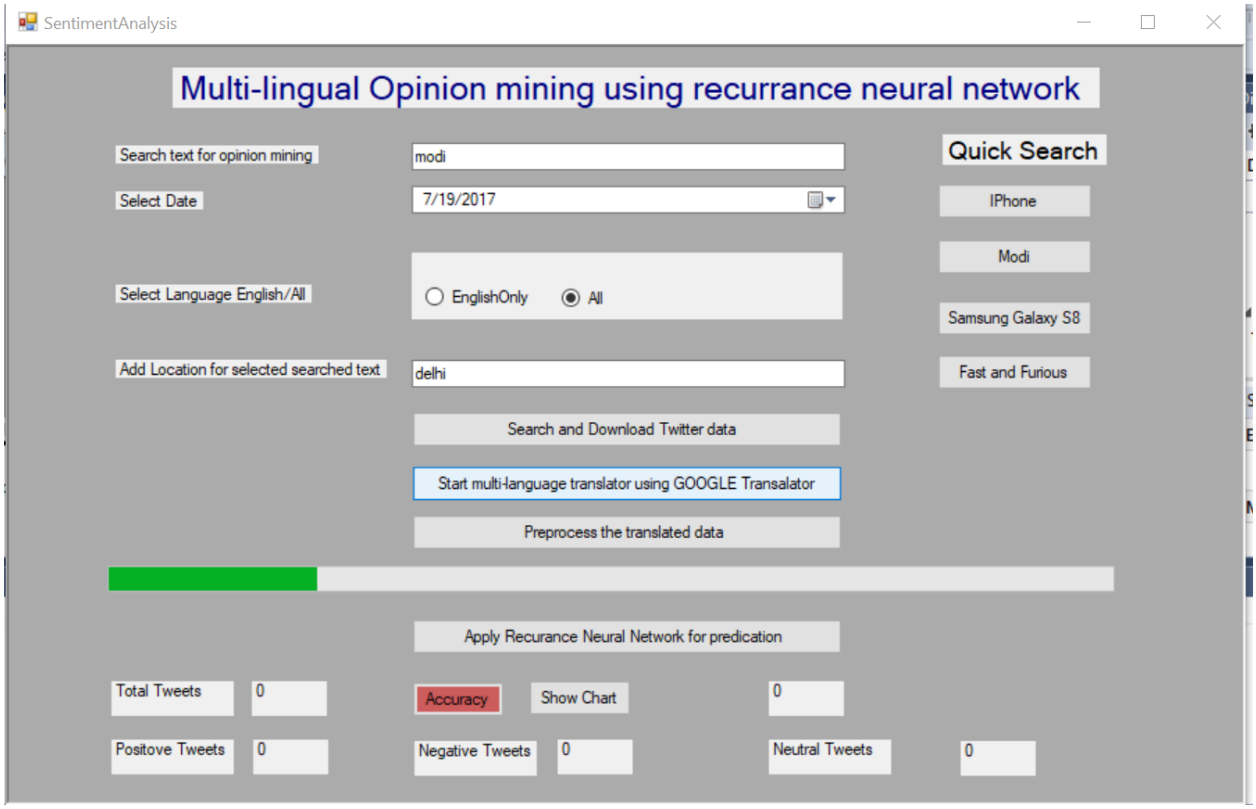


Fig13 : Other language tweets start converting

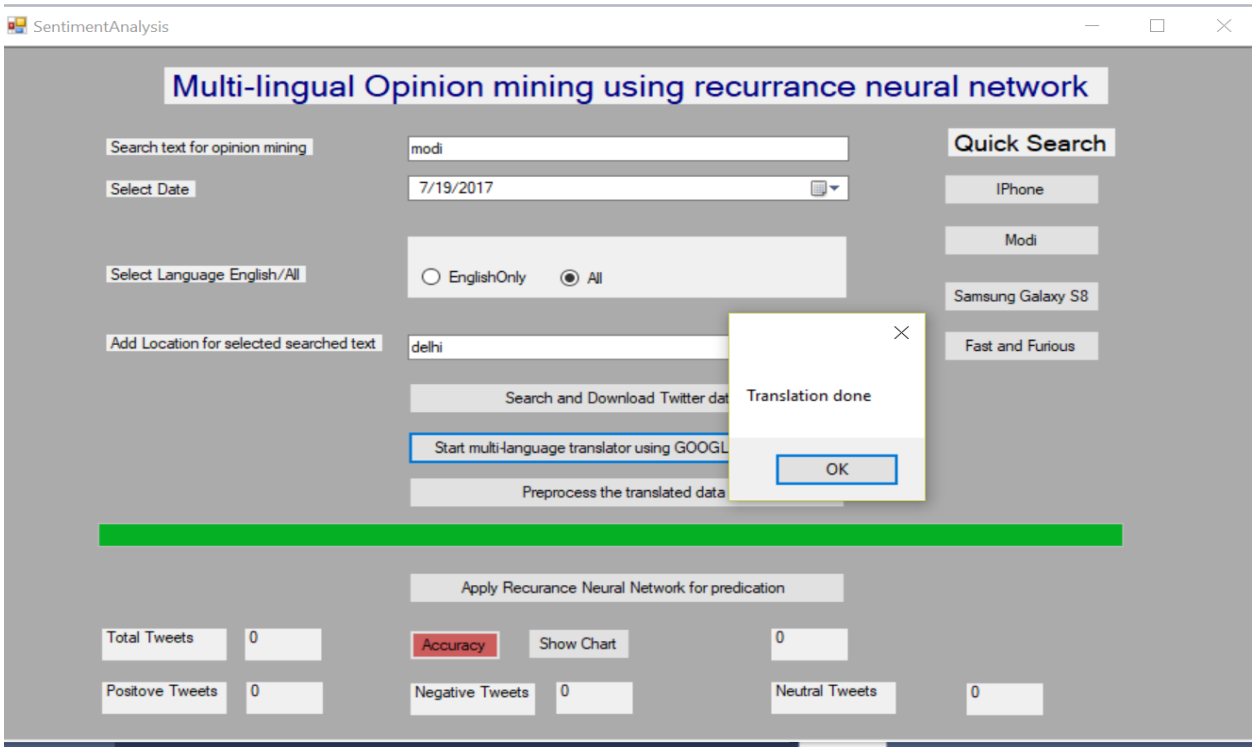


Fig 14 : Translation Done

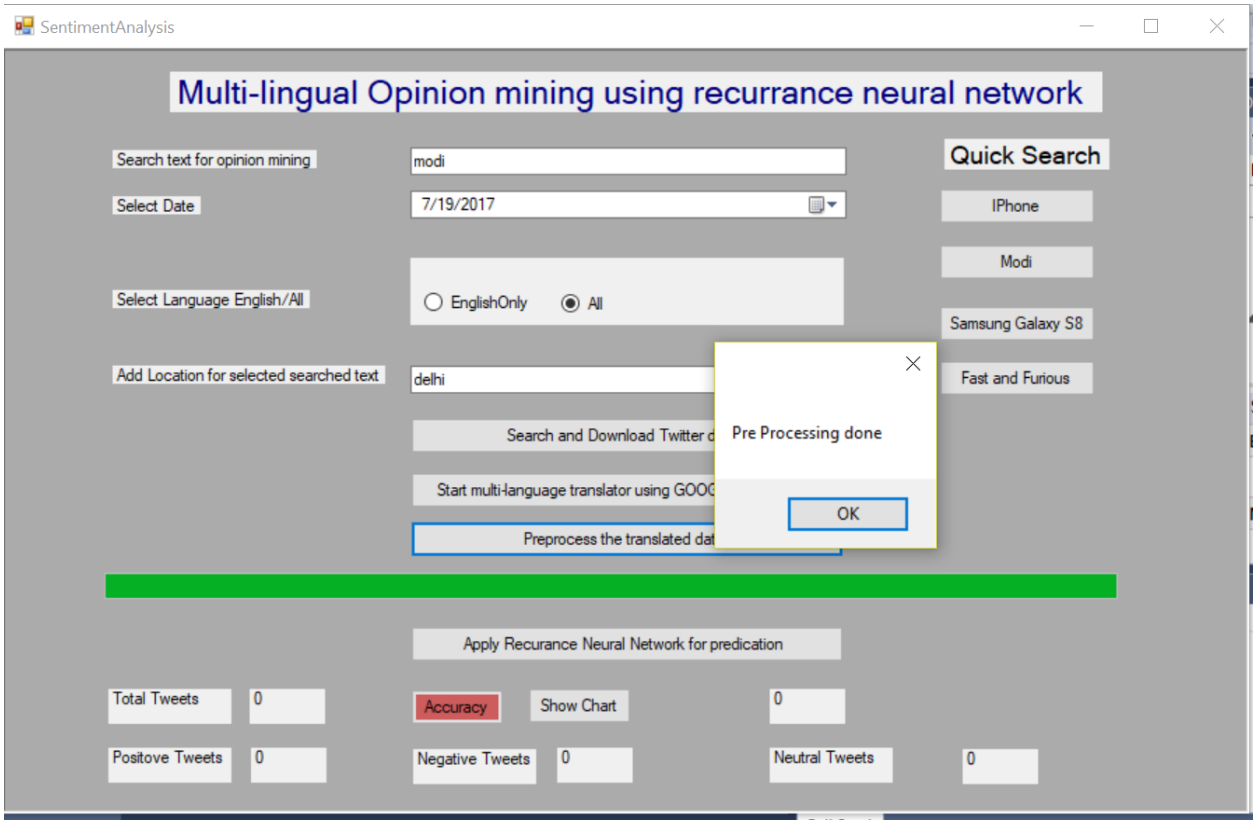


Fig 15 : After Preprocessing

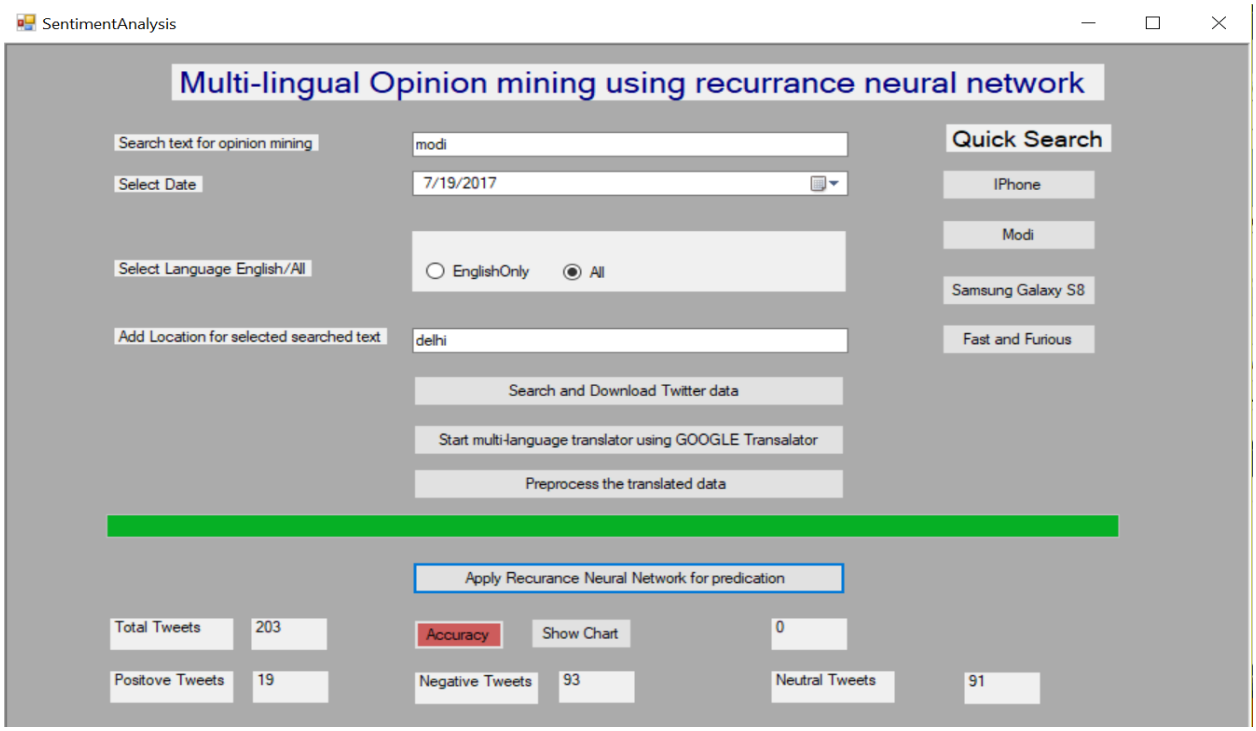


Fig 16 : After Apply Recurrence Neural Network

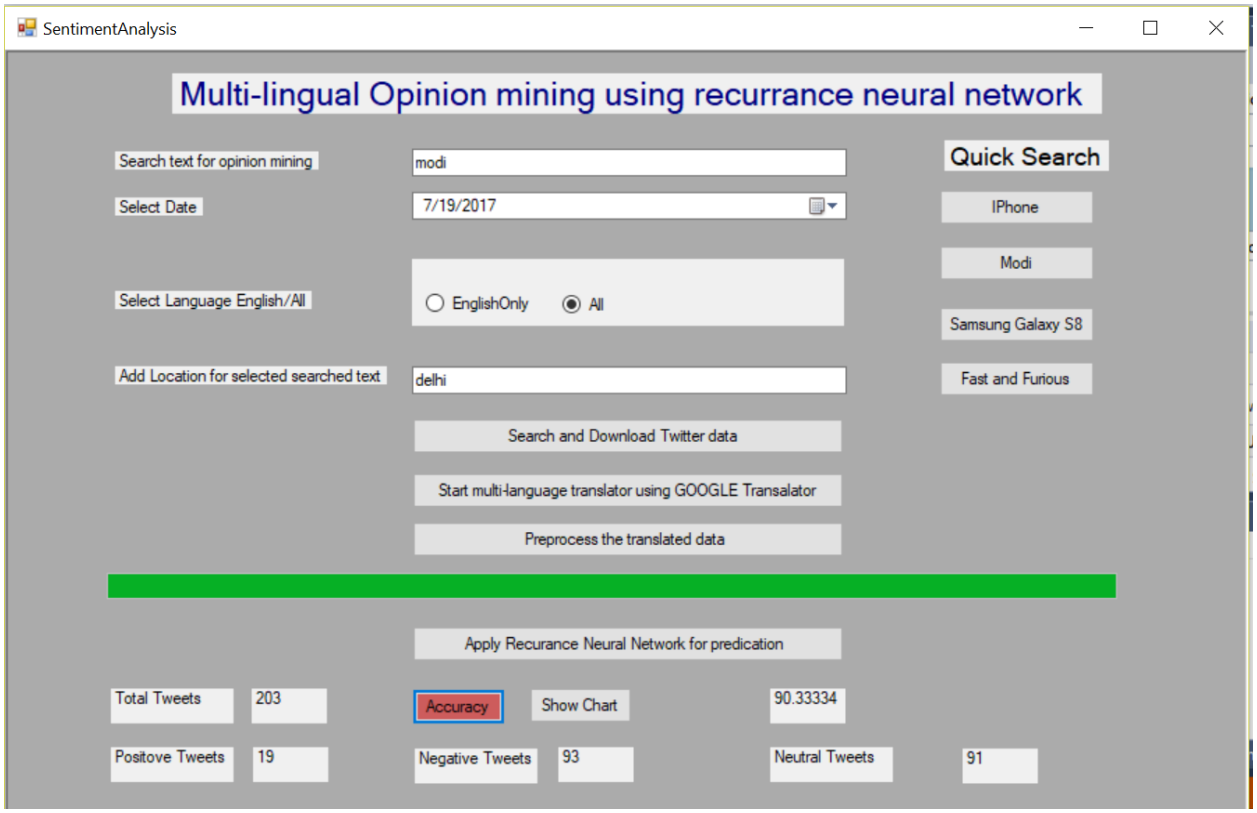


Fig 17 : After Calculation of Accuracy

6.3 Prediction over Testing Data Set

After training the machine on training datasets using back propogate neural network, we will now predict tweets. Following is a screenshot of the following results, which was estimated by the Back propgation Neural Network Algorithm Testing Data Set C, using a following test data set, is a graphical representation of the reproduction of the prophecy, using 3,000 tweets using Neural Network algorithms using 3,000 tilts Is used. Neural network algorithm

Forecasting is at the center of almost every scientific discipline, and the study of generalization (which, prediction) from data is a central topic of machine learning and statistics, and more generally, data mining. Machine use and statistical methods are used to combat "overload surcharges" that reflect our current digital age. The data of the Artificial Intelligence Community has evolved mainly in the last 30 years, as well as the availability of modern computing, the statistics have made great progress, although some parts of these two areas should be targeted at the same goal, which Data is anticipated. The following is a graphical representation of the predicted data.

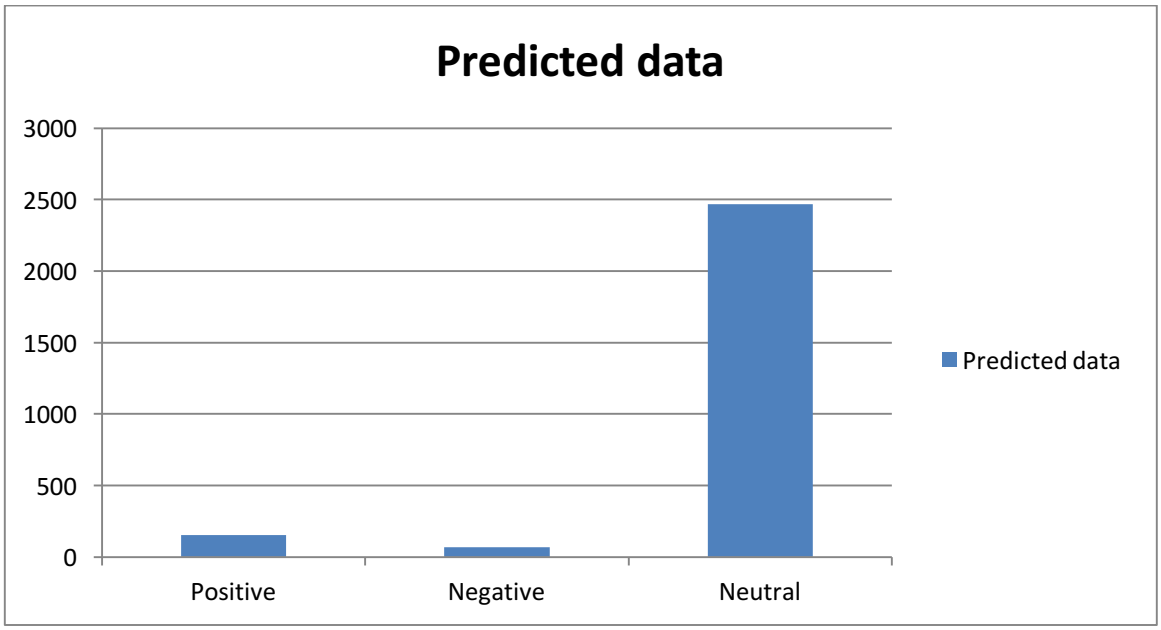


Fig 18 Graphical representation of predicted data

While performing Prediction over the data set during the testing phase an error was encountered which basically displayed the number of tweets which were misclassified during prediction The following graph represents the error which was encountered and hence shows the tweets which were misclassified.

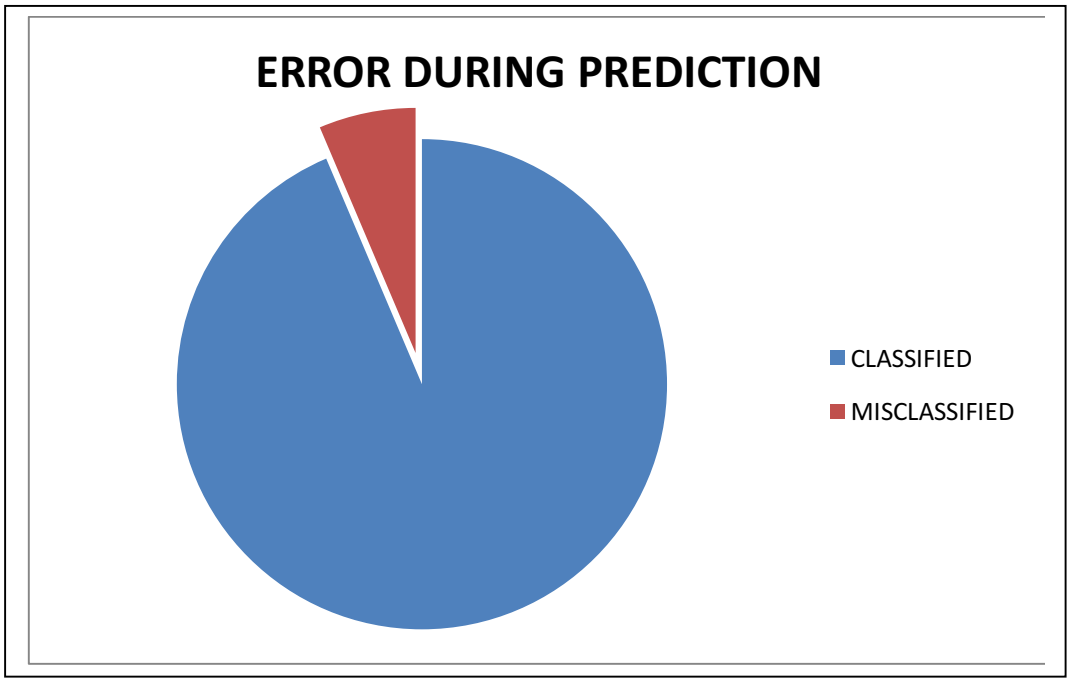


Fig 19 Graphical representation of misclassified tweets during Prediction

6.4 Comparison With Other Algorithms Used Before

Several algorithms were proposed for multi-lingual emotion analysis, these algorithms used to perform good for emotion classification, but when the nervous system was applied to it, it would overcome all other algorithms to give more accuracy. The previously used algorithm was discussed below-

6.4.1 Inactive bayous algorithm

A potential classifier that uses the properties of the Bayes theorem further assumes a strong freedom among the features. One of the important benefits of this classifier is that it uses a small amount of training data for the next calculation of the required parameters for the prediction phase, it is also an estimate that the predictions are independent of each other. We classify tweets using training facilities like we use training datasets. The probability of class for polarization is calculated using meaningful bayus logarithmic possibilities. Classifier then specifies the class label on given tweets.

In general terms, according to the niive biis classifier, the presence of any particular feature of a particular class is not related to the presence of any other feature. Ask for a given document 'X' and it is calculated as a conditional probability P (C / D) for each class 'C' (for both positive and negative). According to the Bayes theorem, the possibility can be calculated as:

$$\boldsymbol{P(c|x) = \frac{P(x|c)P(x)}{P(x)}}$$

It can be further extended in the following manner-

$$\boldsymbol{P(c|x) = P(x1|c) \times P(x2|c) \times P(x3|c) \ldots\ldots \times P(xn|c) \times P(c)}$$

6.4.2 Maximum Entropy

Max Antoopy Classifier is also known as one of the potential classifier, which falls under the category of exponential models. It does not believe that the facilities need to be independent of each other because it was inevitably assumed in the case of twenty-two. MaxEnt mainly focuses on the principle of maximum entropy and selects maximum entropy among models that choose the training. This classifier is primarily used to classify the variety of texts

Where N is given by the size of the training dataset and the xi has relevant information related to the document (sparse array) and its class of y. Using the above empirical probability distribution to create a statistical model of the random process, which considers their relevant information

allocates to a particular class. Building blocks in our model are a group of our statistics that come as a result of our training dataset, i.e. empirical probability distribution.

We have also compared our results to using neural network algorithm for mechanism analysis with those results, which came on the implementation of the Knee Bayes and Max Entropy algorithms, which were previously published and published. There is a graphical representation of the differences received for the above.

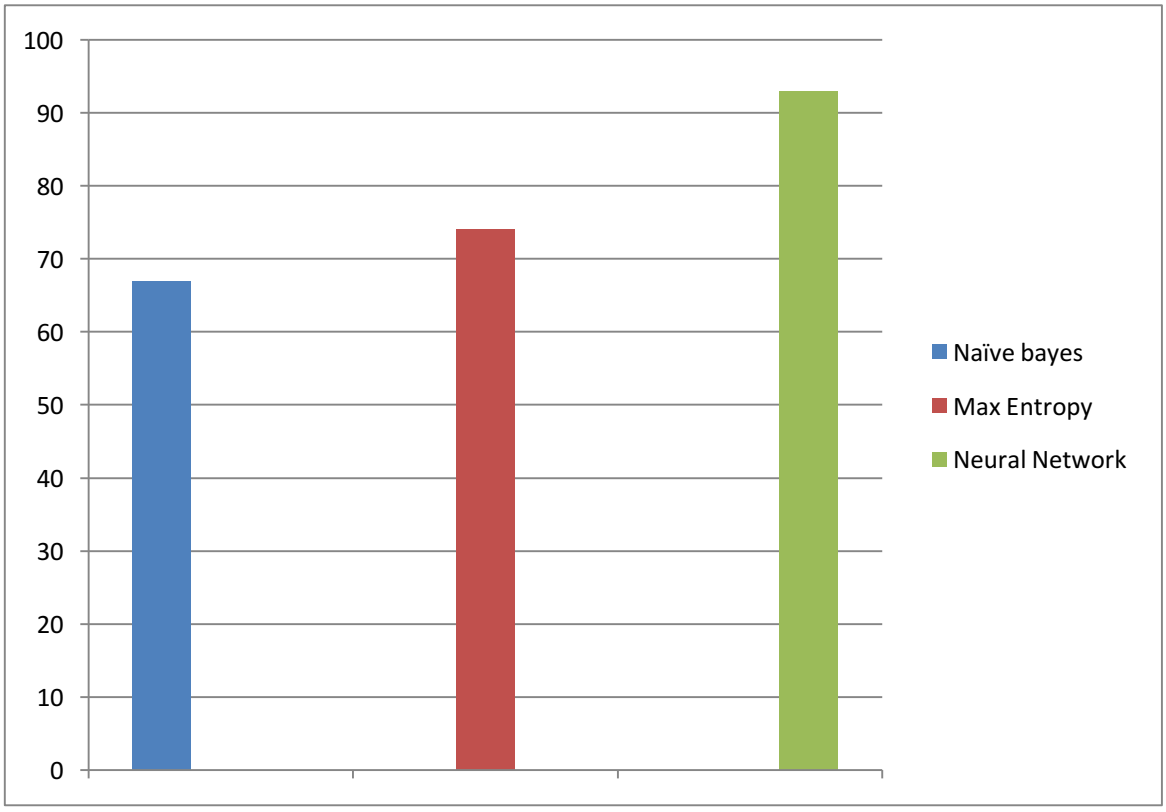


Fig 20- Accuracy Comparison of different algorithms

After the graphical presentation shown in the next picture, the comparison of the capabilities of different classification algorithm and neural networks for the same language is shown. I.e Non-multilingual and multilingual tweets i.e. Multiple Lnguages The following graph is shown below and the algorithm used for comparison is the maximum interpp and the Naver Bayes classification.

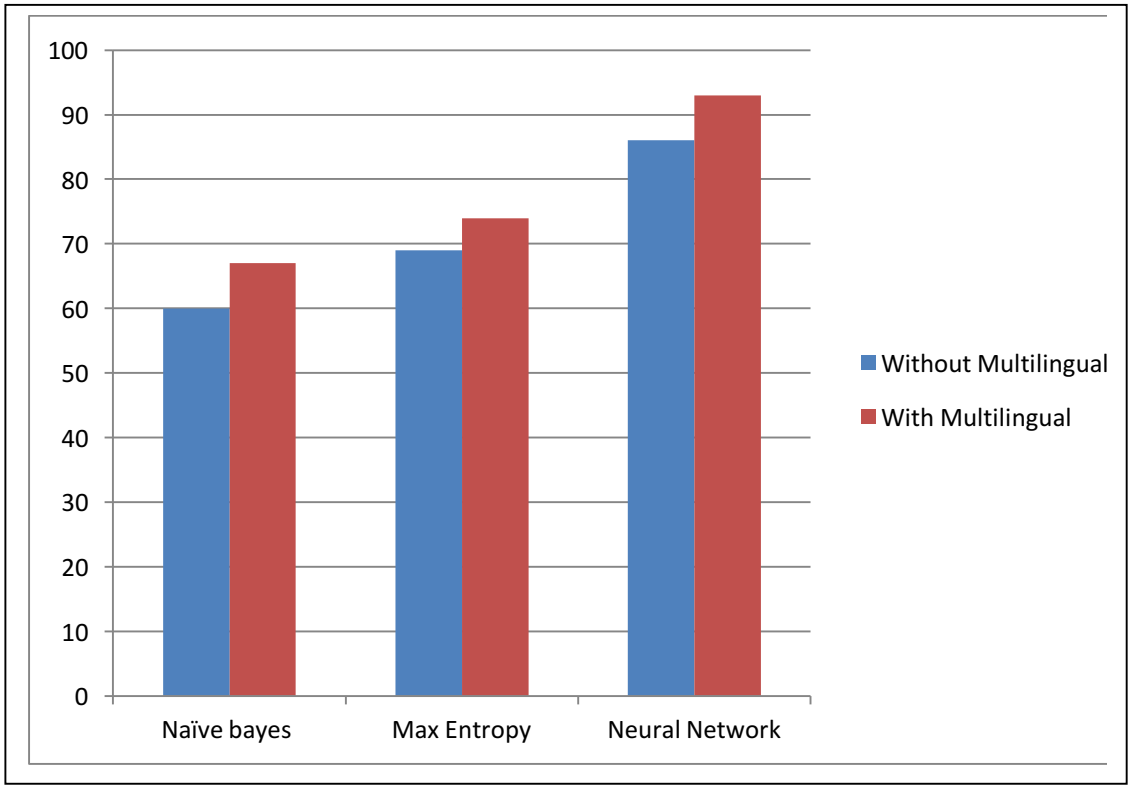


Fig 21- Comparison of Neural Network with different algorithms for both Multilingual and Without Multilingual

Chapter 7

Conclusion

In our research, we have made an attempt to classify emotion analysis that tweets that we have brought from twitter using the neural network algorithm i.eBackpropogation. We came to be the accuracy of our algorithm which came out to be approximately 93% .Multilinguil sense analysis has been performed using different classification algorithms. These algorithms are basically performing well, but when we apply deep learning, then we have seen this algorithm very far away. Our proposed methodology is to use the artificial neural network to interpret multi-lingual sense analysis with the accuracy of 63.5% and 73.5% respectively, using pre-operations inexperienced Bayes and Max Entropy algorithms, respectively, we also have the previous implementation of two algorithms With our algorithm comparative comparison reflecting multilinguil sense analysis ie naïve Bayes (nb) and max entropy .. comments it is very neutral Make sure Etvrk is 93.145% beats other classification algorithms work to predict the emotions with nearly accuracy. So we were successful in executing Mulilinguil sense analysis in a simplified manner with a high accuracy.

Chapter 7

Future scope

In our research, we have successfully implemented neural network algorithms to classify large amounts of datasets to address the topic. There are various challenges for emotional analysis first. In an attempt to combat this, we have used neural network algorithm to classify the emotions made by slangs, misspellings, emoticons and multilingual words and achieved accuracy of 93.145%.

The more difficult challenges of emotional analysis will also include that it can be used for the next advancement of this study in the coming future. Apart from this, we can use a hybrid of fuzzy logic and neural networks i.eNeurofuzzy to improve and improve the system's accuracy.

References

[1] ekeCai, Scott Spangler, Ying Chen, Li Zhang, "Leveraging Sentiment Analysis for Topic Detection", ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2008.

[2] Balakrishnan Gokulakrishnan, Pavalanathan Priyanthan ThiruchittampalamRagavan, NadarajahPrasath, AShehanPerera, Opinion Mining and Sentiment Analysis on a Twitter Data Stream, The International Conference on Advances in ICT for Emerging Regions - ICTer 2012 : 182-188.

[3] Federico Neri, Carlo Aliprandi, Federico Capeci, Montserrat Cuadros, Tomas, "Sentiment Analysis on Social Media", ACM International Conference on Advances in Social Networks Analysis and Mining, 2012

[4] Y. Singh, P. K. Bhatia, and O. Sangwan, "A review of studies on machine learning techniques," International Journal of Computer Science and Security, vol. 1, no. 1, pp. 70–84, 2007

[5] KekeCai, Scott Spangler, Ying Chen, Li Zhang, "Leveraging Sentiment Analysis for Topic Detection", ACM International Conference on Web Intelligence and Intelligent Agent Technology, 2008.

[6]BalakrishnanGokulakrishnan, PavalanathanPriyanthan, ThiruchittampalamRagavan, NadarajahPrasath, AShehanPerera, "Opinion Mining and Sentiment Analysis on a Twitter Data Stream", The International Conference on Advances in ICT for Emerging Regions - ICTer 2012 : 182-188.

[7] Federico Neri, Carlo Aliprandi, Federico Capeci, Montserrat Cuadros, Tomas, "Sentiment Analysis on Social Media", ACM International Conference on Advances in Social Networks Analysis and Mining, 2012.

[8] S. Argamon, K. Bloom, A. Esuli, and F. Sebastiani, "Automatically determining attitude type and force for sentiment analysis," in Human Language Technology.Challenges of the Information Society Springer, 2009, pp. 218–231.

[9] Y. Singh, P. K. Bhatia, and O. Sangwan, “A review of studies on machine learning techniques,” International Journal of Computer Science and Security, vol. 1, no. 1, pp. 70–84, 2007.

[10] Praveen Kumar, Tanupriya Choudhury, Seema Rawat, Shobhna Jayaraman Analysis of Various Machine Learning Algorithms for Enhanced Opinion Mining using Twitter Data Streams , 2016 International Conference on Micro-Electronics and Telecommunication Engineering

[11] Santhosh Kumar K L, Jayanti Desai , Jharna Majumdar , Opinion Mining and Sentiment Analysis on Online Customer Review , 2016 IEEE International Conference on Computational Intelligence and Computing Research

[12] G.Angulakshmi, Dr.R.ManickaChezian, “An Analysis on Opinion Mining: Techniques and Tools.” International Journal of Advanced Research in Computer and Communication Engineering, Vol. 3, Issue 7, July 2014.

[13] Vidisha M. Pradhan, Jay Vala, Prem Balani, “A Survey on Sentiment Analysis Algorithms for Opinion Mining”, International Journal of Computer Applications, Volume 133, No.9, January 2016

[14] G. Sneka, CT. Vidhya, “Algorithms for Opinion Mining and Sentiment Analysis: An Overview”, International Journal of Advanced Research in Computer Scinece and Software Engineering, Volume 6, Issue 2, February 2016

[15] D. Davidov,A. Rappoport, “Enhanced sentiment learning using twitter hash tags and smiles” in Proc. 23rd International Conference on Computational Linguistics: Posters, 2010, pp. 241-249.

[16] B. Pang, L. Lee, and S. Vaithyanathan, “Thumps up? Sentiment Classification Using Machine Learning Techniques” in Proc. Conference on Empirical Methods in Natural Language Processing (EMNLP), 2002.

[17] B. Pang and L. Lee, “Opinion Mining and Sentiment Analysis” in Foundations and Trends in Information Retrieval, vol. 2, pp. 1-135, 2008.

[18] R. Ghani, K. Probst, Y. Liu,M. Krema, and A. Fano,"Text mining for product attribute extraction," SIGKDD.voI.8, pp.41-48.2006.

[19] H.Y.Liu. Social Computing: Users Online Behavior Analysis Mining. Beijing, CATsinghua University, pp.l05-114. 2014.

[20] R Nitya, D Maheshwari, “Sentimental Analysis on Unstructured Reviews”, in Proceedings of IEEE International conference on Intelligent Computing Application, pp. 367-371, Coimbatore , 2014.