Sentiment Analysis of Twitter Data Using Machine Learning and Deep Learning

Submitted By:

Name: Swapan Pal

ID: 2018-2-50-016

Name: Mahamudul Hasan

ID: 2018-2-50-029

Name: Amit Paul

ID: 2018-2-50-030

This Thesis Paper is Submitted in Partial Fulfillment of the Requirements of the Degree of Bachelor of Science in “Information & Communication Engineering”.

Department of Electronics & Communication Engineering

East West University

Approval

The thesis paper titled "Sentiment Analysis of Twitter Data Using Machine Learning and Deep Learning" submitted by Swapan Pal(ID:2018-2-50-016), Mahamudul Hasan (ID:2018-2-50-029) and Amit Paul (ID:2018-2-50-030) to the Department of Electronics and Communications Engineering, East West University, Dhaka, Bangladesh has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Information and Communications Engineering and approved as to its style and contents.

Approved By

Supervisor

Dr. Mohammad Arifuzzaman

Associate Professor

Department of Electronics & Communication Engineering

East West University

Dhaka, Bangladesh

Declaration

We declare that our work has not been previously submitted and approved for the award of a degree by this or any other University. As per of my knowledge and belief, this paper contains no material previously published or written by another person except where due reference is made in the paper itself. We hereby, declare that the work presented in this thesis paper is the outcome of the investigation performed by us under the supervision of Dr. Mohammad Arifuzzaman, Associate Professor, Department of Electronics Communications Engineering, East West University, Dhaka, Bangladesh

Countersigned

Supervisor name

Dr. Mohammad Arifuzzaman

Signature

Swapan Pal

ID:2018-2-50-016

Mahamudul Hasan

ID:2018-2-50-029

Amit Paul

ID:2018-2-50-030

Acknowledgement

We would like to expose our deepest gratitude to our supervisor, Dr. Mohammad Arifuzzaman, for his extremely valuable guidance and advice during the topic investigation and our detection technique trial. His motivation to volunteer his time and effort is truly appreciated. We would also want to express our gratitude for his companionship and understanding. We gained numerous valuable information and thoughts about machine learning approaches from him during our research. His courage encouragement inspired in us the confidence to carry out our mission.

Abstract

In recent years, sentiment analysis has become a popular topic of discussion because of the rapid growth of various social media sites and e-commerce sites. At its core sentiment analysis analyze people's opinion and try to determine the polarity of that opinion. It is very handy to determine customer's review and also to determine various social trends that is going on. Its aim is to identify if a text contains a positive or negative meaning. Nowadays people can post anything on social sites like facebook, tweeter, Instagram. In our thesis we used a dataset that contains more than 29530 tweets of various users. The aim of this thesis is to determine if these tweets contain any hatred content or they are not hatred. Different machine learning technique such as Naïve Bayes, Support Vector Machine, Logistic Regression, Random Forest are implemented in this thesis work. Using these algorithms, we have performed the classification task and the performance evaluation using various parameters such as precision, recall, f1-score, and accuracy. Furthermore, we analyzed their results and compared their outcome with one another. All our model gave almost the similar result varying within 1-2%. Among the machine learning algorithms, we got high accuracy (96.24%) using the Random Forest algorithm. Our work does not finish here as we also tried to implement deep learning to our project. Our aim was to compare the result between machine learning and deep learning and see which model provides the best outcome. From deep learning we used Bidirectional-Long Short Term Memory for this thesis work. However, the result we got from it was a tiny bit less accurate that our machine learning model. Therefore, our final decision is machine learning performs better for our small dataset than deep learning. Finally, in this thesis we also discussed the challenges and limitations of our research.

Table of Contents

Subject	Page
Approval	i
Declaration	ii
Acknowledgement	iii
Abstract	iv
List of Tables	viii
List of Figures	viii

Chapter One (Introduction)

1.1: Introduction	1
1.2: Problem Statement	2
1.3: Motivation	3
1.4: Thesis Organization	4

Chapter Two (Literature Review)

2.1: Literature Review	5

Chapter Three (Introduction To Sentiment Analysis)

3.1: Sentiment Analysis or Opinion Mining	8
3.2: How Does Sentiment Analysis work?	9
3.3 Sentiment Analysis (SA) & Natural Language processing (NLP)	10
3.4 Different level of sentiment analysis	14
3.4.1: Document level analysis	14
3.4.2: Sentence level sentiment Analysis	14
3.4.3: Aspect level sentiment analysis	14

3.5: Sentiment Analysis Models 14
3.5.1: Bag of Words 15
3.5.2: Lexicon 15
3.5.3: Part-of- Speech Tags 16
3.6: Sentiment Analysis Approaches 17
3.6.1: Machine learning approach 17
3.6.2: Deep learning approach 18
3.7 Sentiment Analysis Challenges 19

Chapter four (Machine Learning)

4.1: Machine Learning 22
4.1.1: Supervised Learning 23
4.1.2: Unsupervised Learning 27
4.1.3: Reinforcement Learning 28
4.2: Evaluation Parameters 29
Precision 30
Recall 30
F1-measure 30
Accuracy 31

Chapter five (Deep Learning)

5.1: Deep Learning 32
5.2: Deep Neural Network 34
5.3: How Deep Neural Network Works 34
5.3.1: RNN (Recurrent Neural Network) 37

Chapter Six (Research Methodology)

6.1: Methodology 40
(i) Data Pre-Processing 41
(ii) Text Processing 46
(iii) Analyzing the Data 48
(iv) Model Building 48

Chapter Seven (Result and Analysis)

7.1: Result and Analysis 50
7.2: Applying Machine Learning Algorithm 50
7.3: Applying Deep Learning Algorithm 52
7.4: Applying Confusion Matrix 53

Chapter Eight (Comparison Between ML and DL)

8.1: Advantages of Machine Learning 54
8.2: Challenges of Machine Learning 54
8.3: Advantages of Deep Learning 55
8.4: Challenges of Deep Learning 55
8.5: Comparison Between Machine Learning and Deep Learning 56

Chapter Nine (Conclusion)

9.1: Research Challenges 58
9.2: Future Scope 58
9.3: Conclusion 59
Reference 61

List of Tables

Table No	Title	Page
Table 4.1: Confusion Matrix		29
Table 7.1: Experimental Result of Machine Learning Algorithms		50
Table 7.2: Experimental Result of Deep Learning Algorithm		52
Table 8.1: Comparison between the accuracy rates of ML & DL		56

List of Figures

Figure No	Title	Page
Figure 3.1: Sentiment analysis conceptual graph		8
Figure 3.2: Working Mechanism of Sentiment Analysis		10
Figure 3.3: Various approaches of sentiment analysis		17
Figure 3.4: Overview of machine learning		18
Figure 3.5: Architectural overview of Deep learning		19
Figure 3.6: Challenges of sentiment analysis		20
Figure 4.1: How Machine Learning Works		22
Figure 4.2: Supervised Learning		23
Figure 4.3: How Naïve Bayes Algorithm works		24
Figure 4.4: Support Vector Machine Algorithm		25
Figure 4.5: Logistic function models the conditional probability of the response		26
Figure 4.6: How Random Forest Algorithm works		27
Figure 4.7: Unsupervised Learning		28
Figure 4.8: Reinforcement Learning		29
Figure 5.1: What is Deep Learning		32
Figure 5.2: Deep Learning		33
Figure 5.3: Deep Neural Network		34
Figure 5.4: Basic Model of Deep Learning		35
Figure 5.5: Perceptron		36
Figure 5.6: RNN		37
Figure 5.7: RNN have Loops		38
Figure 5.8: The repeating module in a standard RNN contains a single layer		38
Figure 5.9: The repeating module in an LSTM contains four interacting layers		39
Figure 6.1: The small sample screenshot of the dataset of the Twitter Data		40
Figure 6.2: Work Process		41
Figure 6.3: Distribution of train and test dataset		42
Figure 6.4: Amount of hatred and non-hatred tweet in our dataset		42
Figure 6.5: Distribution of length of the hatred tweet		43
Figure 6.6: Distribution of length of the non-hatred tweet		43
Figure 6.7: World Cloud view for most used word in all tweets		44
Figure 6.8: World Cloud view for normal word in all tweets		44
Figure 6.9: World Cloud view for positive word in all tweets		45

Figure 6.10: World Cloud view for negative word in all tweets 45
Figure 6.11: Quantity of top 20 positive word 46
Figure 6.12: Quantity of top 20 negative word 46
Figure 7.1: Confusion Matrix of SVM 53
Figure 7.2: Confusion Matrix of Naïve Bayes 53
Figure 7.3: Confusion Matrix of Random Forest 53
Figure 7.4: Confusion Matrix of Logistic Regression 53

CHAPTER ONE

Introduction

1.1: Introduction:

At this moment in twenty first century, we are living in an era of information and technology where terabytes of data and information is accessible in the palm of our hand. This information or data can be used to make vital decisions in every aspect of our lives. Among this information a significant portion comes as opinion or point of view of millions of people which are openly floating around the internet. Anyone can share his or her opinions, thoughts or interests with millions of other in a matter of seconds because of the rapid growth of social media or blog sites in the internet. This growth of internet has changed how people express their perspective. In the past two decades a massive number of human populations have been attracted towards social media sites such as facebook, twitter, snapchat, instagram. Among these users most of them use these social media sites for expressing their beliefs, interests, and opinions, emotions about things or persons. It is done through facebook post, twitter tweet, commenting or writing review or feedback of things in particular website. When we try to analyze or take decision about something trending, we always come across such opinions or posts of other people to have necessary idea for taking our decision. For example, when we want to buy a product, we first look for reviews or opinions about the same product from different users. After analyzing others opinions, we make our decision about the product. This can be also true for other things like movie review, travel experience and so on. However, there is also a matter of concern which cannot be ignored at any cost. As we all know that the internet is open and anyone can share their opinions or feedback about anything, which also creates opportunity for fake post, fake review or can spread hatred speech. Therefore, it is more necessary than ever to analyze if the shared opinions or reviews are genuine or fake. And here is where sentiment analysis plays one of the most vital and biggest roles to detect if the opinions are beneficial or not.

Sentiment analysis can be regarded as the automated process to find out whether a text or sentiment gives positive, negative or neutral expression about something or a topic. It analyzes human language as text and try to find out the polarity of the expression or opinion. Nowadays it has become a vital tool for businesses and customer services. For example, customer can express their feedback or opinion about a service or product of a company or brand more openly and freely than ever before. When the company needs to analyze these opinions of the customer it is quite impossible to go through each individual feedback manually. This is where sentiment analysis can become a very powerful tool to analyze thousands of customer feedback and opinion. Analyzing millions of reviews and feedback at a time has become the biggest advantage for businesses around the world. Companies can analyze these customer expressions and adjust their services accordingly. Sentiment analysis helps companies monitor their brand reputation on social media, gain insights from customer feedback, and much more. Not just in business, sentiment analysis is also used for analyzing the behavior of users on social media like facebook and twitter.

In the rise of machine learning and deep learning it has become easier than ever before to analyze millions of data within few minutes. With the help of machine learning and deep learning we can perform sentiment analysis over any given dataset consist of thousands of sentiments or text feedback with very decent accuracy. In this paper we will describe how sentiment analysis can performed using various machine learning and deep learning algorithm.

The goal of the thesis is to find out the sentiment polarity of tweets using Machine Learning and Deep Learning Models. When we researched about sentiment analysis, we found that Random Forest algorithm works best for sentiment analysis. We have also used other algorithms of machine learning model like Naïve Bayes, Support Vector Machine and Logistic Regression. In the recent studies we have seen that from Deep Learning Model the Bidirectional-Long Short Term Memory algorithm is doing tremendous job when it comes to analysis of sentiment. That is why we have also used Recurrent Neural Network with Bi-LSTM in our experiment. By using all of the model we are trying to analyze sentiments of a twitter dataset. Additionally, this paper also tries to serve an idea of sentiment and tries to compare Machine Learning and Deep Learning model.

1.2: Problem Statement:

In this recent era, the world around us has become incredibly dependent on digital technology. We live in a moment where technology does amazing stuff that was once fantasy for human mind. One of the most amazing inventions of modern technology is online social media where millions of people are now able to share their thoughts, perspective and opinion with one another. For various purposes this opinions or statements needs to be analyzed. For example, business may want to analyze the feedback of people over their product to improve their services, or sometime it is necessary to detect if a statement made by a person is harmful or not, or may be if a statement is true or false. So, it is necessary for businesses and organizations to detect the polarity of the statement made by the users. Detecting this polarity of statements is known as sentiment analysis.

There are lots of researches and works have been done in this field of sentiment analysis using machine learning and deep learning. For this project we will try to make more accurate model using two approaches. We will try to analysis our dataset by applying machine learning and deep learning algorithms and compare the outcomes of the two methods. The overall goal of our project is given bellow:

- Selection of dataset which contains few thousands of tweets.
- Preprocessing and filtration of the dataset.
- Analyzing the dataset using supervised machine learning algorithms.
- Using RNN with Bi-LSTM and comparing the result with machine learning algorithms.
- Showing the results and finding the best model for our dataset.

1.3: Motivation:

As we have discussed earlier, that sharing opinions or feedback has become wide spread after the rise of the social media. Therefore, it has become necessary to analyze people's feedback to understand their opinion. For example, analyzing the reviews or customer feedback is crucial for modern business to meet the customer expectations. If business uses sentiment analysis it becomes very easy for them to track the customers expectations and adjust their services accordingly. This is one of the main reasons behind choosing sentiment analysis as our thesis topic as it is one of the most trending topics which is still very essential for businesses or organization to grow. Moreover, we know that social media is available for all of us at the palm of our hand. We share our thoughts and opinions every now and then in social media sites like facebook and twitter. However, for the easy access of social sites it is also possible for people to spread fake news and hatred messages. Therefore, it also very necessary for the organizations that control social sites to filter out such tweet or status and only allowing non harmful posts. This is one of the most vital factors that motivated us to perform research on this field. We therefore, selected a dataset that contains thousands of tweets where some of them actually contains hatred speech. Our goal is to detect those hatred tweets using both machine learning and deep learning models. If we can detect such tweets, it will be very beneficial as the model will be able to differentiate harmful tweets that may cause hatred. For approaching to our goal, we have used few machine learning and one deep learning model for this thesis work. From machine learning we have selected Naïve Bayes, Support Vector Machine and Logistic Regression and from deep learning we have also used RNN with Bi-LSTM. The reason for using both deep learning and machine learning is to find out which method works best for our dataset.

1.4: Thesis Organization:

Thesis organizations consists of nine chapters. These are organized as follows:

In chapter one, we introduced the introduction, problem statement, motivation, thesis organizations. Chapter two, we discussed the Literature Review. In chapter three, we covered an Introduction to Sentiment Analysis. In chapter four, we covered about Machine Learning, also discussed about the evaluation parameters. In chapter five, we talked about deep learning. In chapter six, we discuss about research methodology where we clarified the methodology briefly. In chapter seven, we showed the result and analysis. Also showed the best classifier model. In chapter eight we discussed advantages and disadvantages of ML & DL model in addition to the comparison between ML & DL model. Finally in chapter nine, we analyzed the research challenges, limitations of research, future scope and conclusion.

CHAPTER TWO

Literature Review

2.1: Literature Review:

Sentiment analysis is known as gathering people's opinions. Sentiment analysis is a field that automatically evaluates and analysis many variables such as people's opinions, attitudes, and emotions based on user-generated text or rating. Sentiment analysis is the use of natural language processing.

There are so many research papers, articles, and survey articles published, which are directly and most relevant papers. Some of those which the current study was inspired are discussed below:

In author [1] analyzed public view towards a product. To analyze the product, they selected one of the most famous social media platforms, which is Twitter. They developed a Natural Language Processing (NLP) which was based pre-processed data framework to filter tweets. Then they implement, Bag of Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDE) model concept to analyze sentiment.

In author [2] analyzed Bangla Text by using sentiment analysis and Deep Learning Algorithms. They implemented the rule-based method Bangla text Sentiment Score (BTSC) algorithm for extracting polarity from large texts. Then these feds went to the neural network. They used a famous deep learning algorithm which was Long Short-Term memory (LSTM).

In author [3] analyzed supervised Machine Learning algorithms for spam email detection. They used many many machine learning algorithms to detect spam mail. They collected data from email and then analyzed those data. They used Naïve Bayes, Support Vector Machine, Random Forest Classifier machine learning algorithms. They got a decent accuracy in each algorithm such as 98.8%, 97.6%, 91.5%, 97.8%, 98.5% accuracy in Multinomial Naïve Bayes, Bernoulli Naïve Bayes, Gaussian Naïve Bayes, Random Forest classifier, Support vector machine (SVM) respectively.

In author analyst [4] they evaluated a text valence classifier based on SVM in total of four languages. Their results showed that the classifier is not confined to a single language, with an accuracy rate of 95.31 % for English, which is 11 % higher than traditional approaches.

In author analyst [5] build a sentiment analyzing model by using Graphlab. They used machine learning algorithm. For showing the result, they used Support Vector Machine (SVM) classifier along with N-grams feature.

In author [6] analyzed on multi-tire sentiment analysis using supervised machine learning. Document management is a big thing in the modern world. They classified the text and organized different categories. They used four different machine learning algorithms which were Naïve Bayes, SVM, Random Forest, and SGD.

In author analyst [7] they created classification of sentiment review. They tested with 1000 positive and 1000 negative examples in an English data set. The accuracy of the NB the technique is 89.53 %, whereas the accuracy of the SVM approach is 94.06%.

In author [8] the analyzed sentiment with sentiment feature. They used a hybrid method and utilizes a Sentiment Lexicon to generate. Then a new set of features to train Support Vector Machine (SVM) classifier.

In author analyst [9] comparative study of Machine Learning. The effectiveness of three machine learning techniques Support Vector Machines, Naïve Bayes, and Maximum Entropy for classifying online reviews using supervised learning methods utilizing a web model.

In author analyst [10] they analyzed performance of supervised techniques for spam review detection. They compared and analyzed total five supervised learning techniques for review spam detection. The five methods, they compared those were Naïve Bayes, SVM, K-NN, Logistic Regression, and Decision Tree. They discovered that SVM has a higher accuracy 83.19 % than other strategies, but Decision Tree has a very low accuracy 51.00 %.

In author analyst [11] they used deep learning sentiment analysis on Amazon.com review and ratings. They created a model that learns low-dimensional review vector representation. The study was on a database of over 3.5 million people’s reviews. They used RNN, GRU and SVM algorithms.

In author analyst [12] they used sentiment analysis by using product review data. They collected total 5.1 million product reviews data. They used three different algorithms to analyze the reviews. Those algorithms were nave bayes, support vector machines, and random forest.

In author analyst [13] they used total four deep learning algorithms to deep sentiment representation. Those algorithms were LSTM, GRU, BI-LSTM and BI-GRU.

In author analyst [14] they analyzed data from many sources like Amazon data, particularly from the ecommerce site. They presented a new concept which is called toeken2vec. Here, all words, emoji came in a single vector region at the same time. This was a huge advantage. They also considered RNN and LSTM algorithms.

CHAPTER THREE

Introduction To Sentiment Analysis

3.1: Sentiment Analysis or Opinion Mining:

Sentiment analysis, often known as opinion mining, is a technique for assessing the emotional tone of a body of text using natural language processing (NLP). This is a standard corporate strategy for determining and categorizing consumer opinions about a product, service, or concept. It comprises employing data mining, machine learning (ML), and artificial intelligence to mine text for sentiment and subjective information (AI). According to Wikipedia Sentiments analysis is the invented science of psychology and sociology and both are the scientific study of people emotions, relationships, opinions, and behaviors. In simpler term, the method of determining whether a piece of text is good, negative, or neutral is known as sentiment analysis. Sentiment analysis uses a scoring method to track conversations and evaluate language and speech inflections in order to measure attitudes, views, and feelings about a product, service, or issue. Sentiment analysis aids data analysts in major organizations in gauging public sentiment, doing detailed market research, monitoring product and brand reputation, and comprehending customer experiences.

Figure 3.1: Sentiment analysis conceptual graph

[Source: https://www.analyticsvidhya.com/blog/tag/twitter-sentiment-analysis/]

Sentimental analysis mostly works with textual data and textual data is divided into two categories: facts and opinions. Opinions determine individual's feelings, ideas, or point of view about things, organizations, events, and their features, whereas facts are expressions of things, aspects, entities, occurrences, and their character traits.

3.2: How Does Sentiment Analysis work?

Constructing a vocabulary containing information about which words and phrases are good and which are bad is one way to apply sentiment analysis. SentiWordNet, for example, is a freely accessible lexical resource that assigns three number values to each WordNet synset expressing the characteristics of the words in the system [29]. This lexicon can be manually compiled or obtained automatically. Classifiers are trained with enormous sets of features to categorize a fresh batch of words or phrases, and lexical or corpora annotation is normally done by hand. Other ways to analyzing feelings, rather than relying on word parity, focus on the meaning of phrases or entire publications. This method is most commonly used with text document corpora. The main issue with polarity classification is that it must assess overall sentiment characteristics of a document, whereas the conveyed emotion might be contained in only one sentence or word. Other times, the feeling is expressed indirectly, making it much more difficult to recognize and categorize. The context surrounding these 'hidden' thoughts, on the other hand, might give extremely useful information for categorizing them. We commonly speak about word-level, sentence-level, and document-level sentiment categorization based on this split of the area of sentiment analysis. On the other side, we may locate another technique to emotion mining on the internet. The goal of Web opinion mining is to extract, digest, and identify different elements of independent online information [30]. Advertising agencies or trend watchers may find this useful. Sentiment analysis (sometimes known as opinion mining) is the process of identifying and extracting subjective information from source materials using natural language processing (NLP), text analysis (TA), and computational linguistics (CL). Sentiment analysis is extensively used in scope of goals, from advertising to customer service, in online reviews and social media.

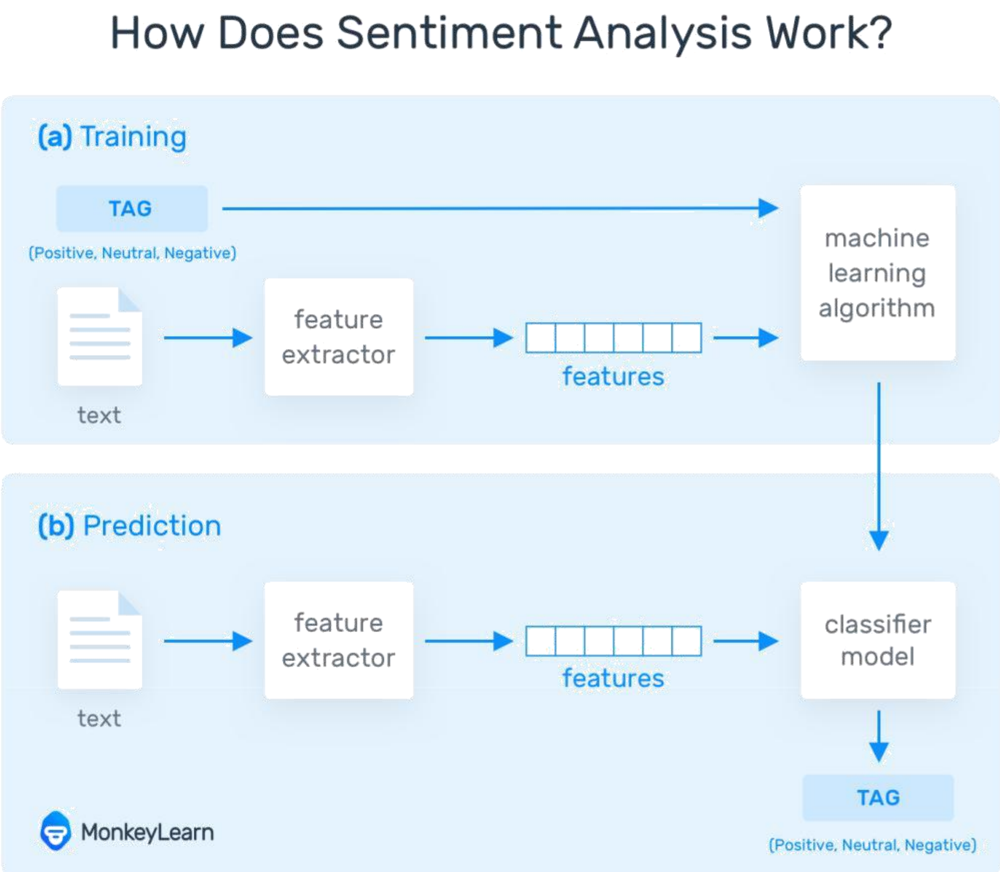

Figure 3.2: Working mechanism of sentiment analysis.

[Source: https://monkeylearn.com/sentiment-analysis/]

3.3: Sentiment Analysis (SA) & Natural Language Processing (NLP):

- **Natural Language Processing (NLP):**

A branch of computer science, computational linguistics, and AI concerned with computer-human interactions through natural language. Natural language processing (NLP) has fundamental relation with human–computer interaction (HCI). Natural language comprehension, or the ability for a device to understand human or natural language input, is one of the most difficult obstacles in NLP [29]. Natural Language Processing is a broad phrase that refers to a collection of techniques for automating the production, manage and exploration of natural language. Despite the fact that most NLP approaches are based on linguistics and AI, they are also influenced by disciplines such as Machine Learning, Computational Statistics, and Cognitive Science.

- **Token**: The incoming text must be differentiated as several components like words, phrase, separators, and other linguistic parts before any serious processing can proceed. Such component is known as tokens.
- **Sentence**: An ordered series of tokens.
- **Tokenization**: Method of separating the tokens that consist in a sentence.
- **Corpus**: Corpus refers to a full textual body containing multiple sentences.
- **Part-of-speech (POS) Tag**: Nouns, Verbs, Adjectives, and Articles are only a handful of the lexical or part-of-speech groups that a word can be classified into. A POS tag - NN (Noun), VB (Verb), JJ (Adjective), AT (Adjective) – is a symbol that represents a lexical category (Article). The Brown Corpus tag set is very popular and widely adopted.
- **Parse Tree**: It depicts a tree that is specified over a given sentence and understands the text's syntactic structure as described by a formal grammar.
- **Part-Of-Speech (POS) Tagging**: A mutual language processing problem is to automatically assign POS tags to each word in a sentence given a sentence and a collection of POS tags. For example, a POS tagger's output for the statement "The ball is red" would be the /AT ball/NN is/VB red/JJ. Modern POS taggers may achieve a level of accuracy as high as 96 percent. Part-of-speech tagging is quite useful to difficult NLP processes like parsing and machine translation.
- **Computational Morphology**: Natural languages include a vast number of words made up of fundamental building blocks called morphemes (or stems), which are the smallest linguistic units with meaning. Computational morphology is concerned with the use of computers to identify and analyze the inherent structure of words.
- **Parsing**: A parser constructs a parse tree given a sentence in the parsing task.

Some parsers require the existence of a set of grammar rules in order to parse, while more contemporary parsers are capable of deducing parse trees directly from the provided data using complicated statistical models. Most parsers work in a supervised environment and require that the sentence be POS-tagged before being processed. In the field of natural language processing, statistical parsing is a hot topic.

- **Subjective Sentence:** It's a statement whereby the writer communicates his or her thoughts or sentiments about entities, events, or their qualities. For example, "I enjoy swimming,".
- **Objective Sentence:** It is a factual sentence regarding objects, events, and their characteristics. An instance**,** “The to-do list includes cleaning, cooking and … ”

• **Opinion:** A judgmental point of view about a subject based on specialized knowledge. Opinions are occasionally expressed openly, such as "This phone's voice quality is fantastic." However, they are occasionally disguised in the tone of a statement, such as "The building collapsed in a year." Because the idea of opinion is so broad, sentiment categorization focuses on overall emotion exhibited by opinions (Positive / Negative). In truth, the Polarity of a viewpoint is determined by its positivity or negative. In other words, recognizing the positive or negative emotion for every subjective statement in the data set, is one of the key subtasks of sentiment analysis.

- **Opinion words:** Opinion words express the positive or negative expression. For example:
  Positive Sentiment {Wonderful, Decent, Caring}
  Negative Sentiment {Worst, Horrendous, dislike}

Sentiment Orientation-SO(Polarity): This express if any presented view is good, negative or neutral using opinion terms. For example: “The gaming performance of this computer is excellent” Positive.

• **Opinion Sentence:** A kind of sentence that comprises opinion words that express a point of view. For example, "The performance of the computer is amazing and the cooling system is also very good".

• **Object / Features:** Objects and their components or characteristics will be examined in opinionated papers, and attitudes about them will be stated on the basis of opinion words; such unit is referred to as object-feature [31]. For Example,

“The economic condition of the country is poor”.

Object Explicit object- feature: economic condition.

Opinion word: Poor.

The specified characteristic in the above case is economic condition, although object properties should occasionally be inferred from the text. This type of feature is known as "This method categorizes and labels various items as an output. Classifiers are used in sentiment analysis to assess emotion of a subjective statement on a subject. Two types of classification are used: Supervised classification, in which classifier is inferred from the building set. For every acceptable input item, the classifier determines the proper label (positive or negative). Unsupervised classification, on the other hand, infers the raw data's hidden structure. Both categorization types are commonly employed in sentiment analysis. Sentiment Analysis' major objective is to extract relevant features to build an engineering feature vector as an input for a classifier.

Text mining [32], sometimes known as text data mining and often synonymous with text analytics, is the technique of extracting high-quality data from text. High-quality data is often produced through the creation of patterns and trends using statistical methods. Text mining normally entails parsing the input text (along with the inclusion of certain derived linguistic characteristics and the removal of others, and subsequent insertion into a database), generating patterns from the structured, and, more recently, evaluating and interpreting the output. In text mining, 'high quality' generally refers to a mix of relevance, uniqueness, and interest.

Text mining is a recent branch of computer science that deals with natural language processing, data mining, machine learning, information retrieval, and knowledge management, among other things. The goal of text mining is to extract useful information from unstructured textual material by identifying and exploring intriguing patterns [33]. Information retrieval, lexical analysis, pattern recognition, tagging/annotation, information extraction, data mining techniques such as link and association analysis, visualization, and predictive analytics are all part of text analysis. The overall purpose is to convert text into data for examination through the use of natural language processing (NLP) and analytical methodologies [34, 35]. Communication experts employ textual analysis to characterize and understand the properties of a recorded or visual message. Textual analysis [35] is used to describe the content, structure, and functions of the messages found in texts. Selecting the sorts of texts to be researched, getting acceptable texts, and deciding which technique to use in evaluating them are all important concerns in textual analysis.

3.4: Different level of sentiment analysis:

We can classify sentiment analysis in 3 categories:

3.4.1: Document level analysis: In document level analysis, the whole document is considered as if it is a single statement. The analysis gives an output for the whole document whether the document gives negative or positive feedback, and a positive or negative polarity is provided depending on the feedback. As example, it determines if the review of a product or service gives negative or positive feedback. This process is known as document level sentiment analysis, where each document expresses a perspective on a certain topic. Therefore, such technique is not much useful where a single document contains information that compare or evaluate number of different elements [36].

3.4.2: Sentence level sentiment Analysis: In this method each and every single sentence or statement is being analyzed or examined to decide its polarity. An opinion where it is not decided if the point of view is positive or negative is considered as neutral and it is treated as having no opinion. This technique is comparable to feature extraction, which seeks to separate phrases based on precise data and express them as subjective viewpoints [37].

3.4.3: Aspect level sentiment analysis: Aspect level sentiment analysis (ABSA) is a text analysis approach that divides data into aspects and determines the sentiment associated with each. Customer feedback may be analyzed using aspect-based sentiment analysis, which associates specific attitudes with different characteristics of a product or service. Aspect level sentiment analysis is significant because it may assist businesses in automatically sorting and analyzing client data, automating procedures such as customer service chores, and gaining useful insights on the go.

3.5: Sentiment Analysis Models:

A dataset and a classifier are used in sentiment text analysis, or more precisely positive/negative categorization. The classifier is applied in two ways in documents: positive and negative. Higher numbers of records being displayed in an informative manner during the test period, results the better categorization. Finding the finest document representations who can better explain it (the document) is crucial in sentiment analysis.

3.5.1: Bag of Words:

Text modeling with a bag of words is a Natural Language Processing approach. In technical words, we may call it a technique for extracting features from text data. This method of extracting characteristics from documents is easy and adaptable. A bag of words is a text representation that describes the frequency with which words appear in a document. We only keep records of number of words and don't pay attention to grammatical subtleties or word arrangement. Because all information about the sequence or arrangement of words in the text is deleted, it is referred to as a "bag" of words. The model simply cares about whether or not recognized terms appear in the text, not where they appear [38]. According to this approach, a dictionary is generated based on training data during the training phase and then utilized to distinguish between positive and negative articles in the testing step.

Let’s suppose we have two sentences below:

1) Welcome to the python training class.
2) Python is easier than java

So the dictionary which is formed for these two sentence using BOW will look like this:

1.Welcome.2.to. 3:”the”,4:”python”,5:”training”,6:”class”,7:”is”,8:”easier”,9:”than”,10:”java

As a result of the created dictionary, each document's feature vector has ten dimensionalities. Word appearance is particularly informative, as demonstrated in the sentiment analysis debate. Because of the nature of natural language, a single phrase may clearly communicate the authors' attitude, but a series of words cannot.

3.5.2: Lexicon:

Instead of building a vocabulary from documents, we used the WordNet dictionary in this study [39]. With 58058 words and four parts of speech tags, WordNet is a huge lexical database in English. Cognitive synonyms (synsets) are a collection of nouns, verbs, adjectives, and adverbs that individually communicate a separate notion. Conceptual-semantic and lexical relationships bind synsets together. WordNet was chosen for three main reasons. To begin with, the dictionary created using the "Aggregated Dataset" (an aggregation of six tiny datasets) was far too large (about 117660 words), rendering the produced matrix useless by MATLAB and PRTools. Because verbs have distinct tenses in documents and nouns have different forms of singular and plural, the dictionary's size has grown ineffectively, whereas words in the WordNet dictionary are in their lemmatized form, allowing MATLAB and PRTools to use the created matrix. The findings suggest that the varied tenses of verbs, as well as the single and plural forms of nouns, are unimportant in sentiment analysis, but their appearance is. As a result, using lemmatized versions of words is a safe assumption.

To be compatible with the WordNet words, all of the words are transformed to lower case in terms of their letters. In addition, all numerals, punctuation marks, and other features were eliminated from the dataset, with the exception of words with two or more letters, because they were not as informative as other words. As a result, keywords like 19, a, @, and others are deleted from the databases. For example, "corruption will not be extreme for the next four years," or "corruption will not be extreme for the next year." Documents are ready to be displayed by the BoW model once the pruning technique has been applied to the dataset. Each dataset is defined using a two-dimensional document matrix and the WordNet dictionary. Each document is verified against the WordNet lexicon, and the presence of each word in each document is indicated by a 1 if it exists, otherwise by a 0. As a result, each row in this matrix will correspond to a feature vector for a document. The resultant matrix has multiple zero columns, indicating that the column's head word was not used in any of the dataset's documents. As a result, they aren't useful and may be eliminated from the matrix.

3.5.3: Part-of- Speech Tags:

Part-of-speech tagging (POS tagging or POST), also known as grammatical tagging or word-category disambiguation, is the process of identifying a word in a text (corpus) as corresponding to a specific part of speech based on both its definition and context—that is, its relationship with adjacent and related words in a phrase, sentence, or paragraph. Tagging each word's grammatical properties is another effective method for improving accuracy ratios and detecting valuable trends for classifications. Instead of the past participle, subjective texts frequently utilize the simple past tense. Because many writers convey their negative thoughts regarding their loss or disappointment, the negative set has more verbs in the past tense than the positive set.

3.6: Sentiment Analysis Approaches:

Several approaches can be followed to perform sentiment analysis. Some of the approaches are discussed below:

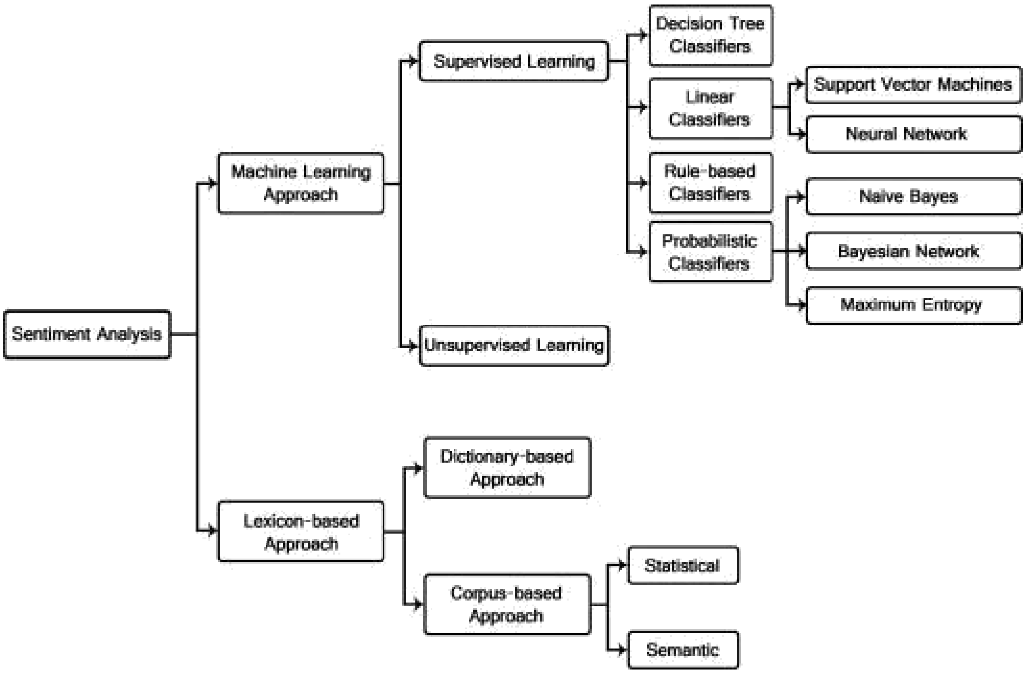

Figure 3.3: Various approaches of sentiment analysis

[Source: https://www.sciencedirect.com/science/article/pii/S2090447914000550]

3.6.1: Machine learning approach:

The fast development of machine learning has made sentiment analysis amongst the most popular topics in this domain. Sentiment analysis itself is a machine learning tool that looks for polarity in words, ranging from positive to negative. Machine learning tools learn how to identify sentiment without human involvement by training them with samples of emotions in text [3]. Simply said, machine learning enables computers to learn new tasks without having to be explicitly taught to do so. Sentiment analysis algorithms may be taught to go beyond definitions to recognize things like background, humor, and misconstrued words.

There are three types of machine learning approach such as supervised learning, un-supervised learning and reinforcement learning. Supervised learning is the first go to choice while performing sentiment analysis. In supervised learning, the dataset is labeled. Which means the data types are already identified prior to feeding the algorithm. The dataset is first divided into training and testing set, then the training dataset is fetched to the machine learning algorithm. Several algorithms such as Support Vector Machine, Naïve Base, Logistic regression, Decision Tree and many more. After the training set is fetched to the algorithm then the testing set is used to analysis the accuracy of the model. If the accuracy matches expectations, then the model is deployed in desired sector.

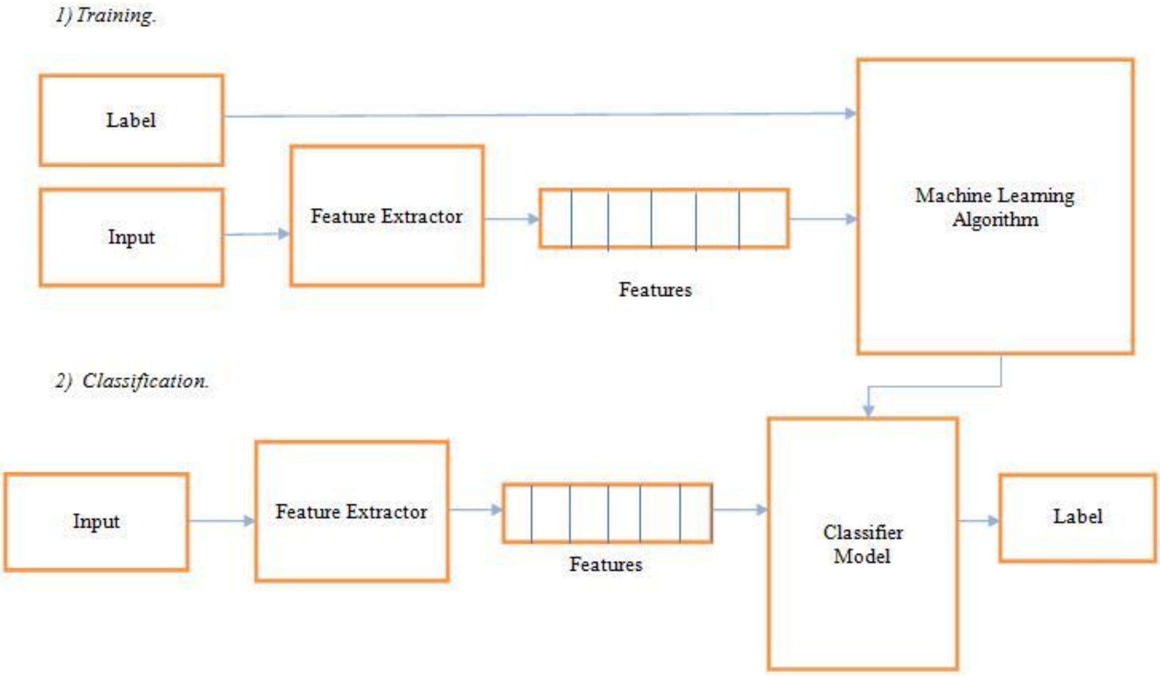

Figure 3.4: Overview of machine learning

[Source: https://laptrinhx.com/twitter-sentiment-analysis-of-movie-reviews-using-machine-learning-techniques-3813505517/]

3.6.2: Deep learning approach:

When dealing with enormous data sets, scientists encounter a variety of challenges, including data intricacy and technology restrictions, since data accumulates at an exponential pace day after day. Numerous deep learning models, such as auto encoders, decoders, Unidirectional and Bidirectional Long Short-Term Memory models,21 Recurrent neural networks, and convolutional neural networks, have been utilized to handle massive volumes of data with breakthroughs in deep learning during the previous decade [40]. These models efficiently try to intuitively identify features or relevant information in the input dataset, and they use heuristic learning to learn interpretations and correlations of complexity in the data to tackle a variety of NLP issues. As a consequence, to tackle the problem of emotion detection categorization, we also use a deep learning model. The input is preprocessed in order to rearrange it for the embedding matrix, followed by the LSTM, and finally the fully - connected layers for text classification.

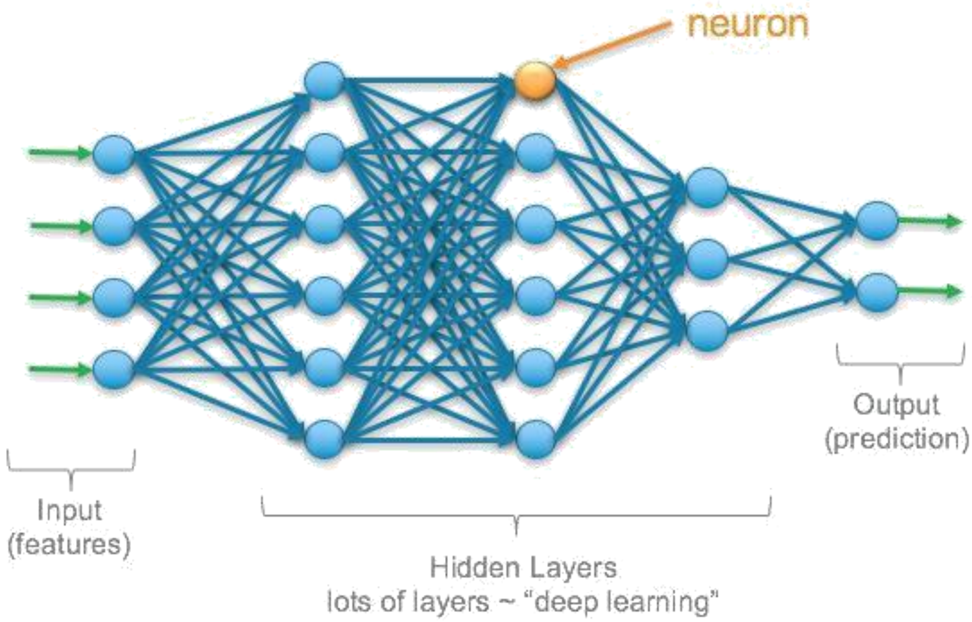

Figure 3.5: Architectural overview of Deep learning

[Source: https://srnghn.medium.com/deep-learning-common-architectures-6071d47cb383]

3.7: Sentiment Analysis Challenges:

Recognizing "sentiment analysis issues" include identifying sentiment challenges in assessment and detecting polarity for reviews, as well as identifying the most effected solutions for the best text accuracy. There is a lot of research in this field. We classify evaluation issues into Theoretical and Technical challenges [41, 42]. Challenges, such as our survey research for forty-seven publications, as described below:

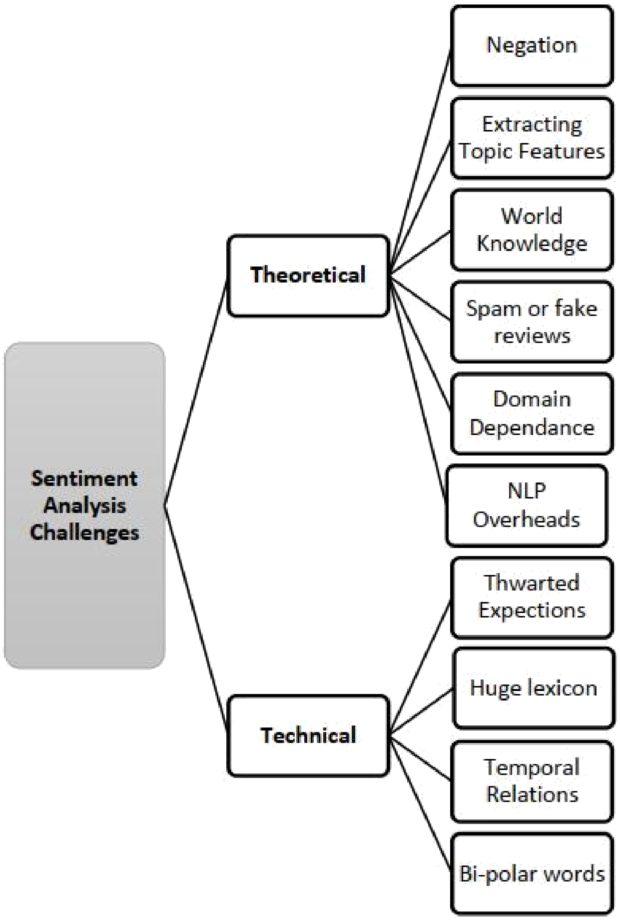

Figure 3.6: Challenges of sentiment analysis

The figure 3.6 describes the main difficulties of sentiment analysis. These difficulties obstruct the sentiment evaluation process and the comprehension of the sentiment reviews' significance.

Four technical Challenges:

1. **Temporal Relations:** For sentiment analysis, the time of the reviews may be important.

 For example: In 2008, the observer may have thought Windows Vista was wonderful,

but in 2009, he or she might have a opposite opinion due to the new Windows 7.

 Observation: As a result, examining certain types of feelings that fluctuate over time may improve sentiment analysis performance. This allows us to see if a product has improved over time or if people's opinions about a product have changed.

2. Thwarted Expectations: Often the writer purposefully creates a circumstance in order to discredit it afterwards. To demonstrate the notion of disappointed expectations, use an

English text. We are unable to distinguish or determine the emotion polarity of some hazy words. Although we can identify the topic domain, we cannot determine their polarity.

3. Bi-polar sentiments: Depending on the topic and qualities, or the implicit meaning of domain terms, a few idioms might have bi-polar interpretations.

- **Observation:** In virtually all situations, the term [Old] has a negative emotion score, however review 1 and 2 logically have a positive polarity, while review 3 logically has a negative polarity. To understand how to detect the polarity, we must first recognize traits, keywords, or the topic area.

4. Generate huge lexicons: The creation of large lexicons to encompass data evaluation is a challenge. The previous forms of sentiment problems have a significant impact on sentence comprehension. Some thoughts and phrases have more than one meaning, while others are based on real knowledge. Asymmetry in the availability of opinion mining software and Term Position are some of the problems.

CHAPTER FOUR

MACHINE LEARNING

4.1: Machine Learning:

Humans learn from past experiences. Humans store their past experiences and when it comes to the action, humans use their experiences and solve those problems in real time. On the other hand, the machine follows a set of instructions that are given by humans. The machine is as straight as a fellow instructor. But machine learning is quite different. What if humans can train the machines. Human set the previous data on machines and machines can act in a faster way. Basically, that is called machine learning.

Machine Learning is in the application of artificial intelligence (AI) that provides systems the ability to enables computers to grow, modify and learn by themselves when they are exposed to a new data. This means that the machine will learn from its experience and then build its own logic based on the problem. One of the important good aspects of machine learning is that algorithms use many computational methods to learn information directly from data without relying on predefined equations as models that easily analyze even large datasets [20].

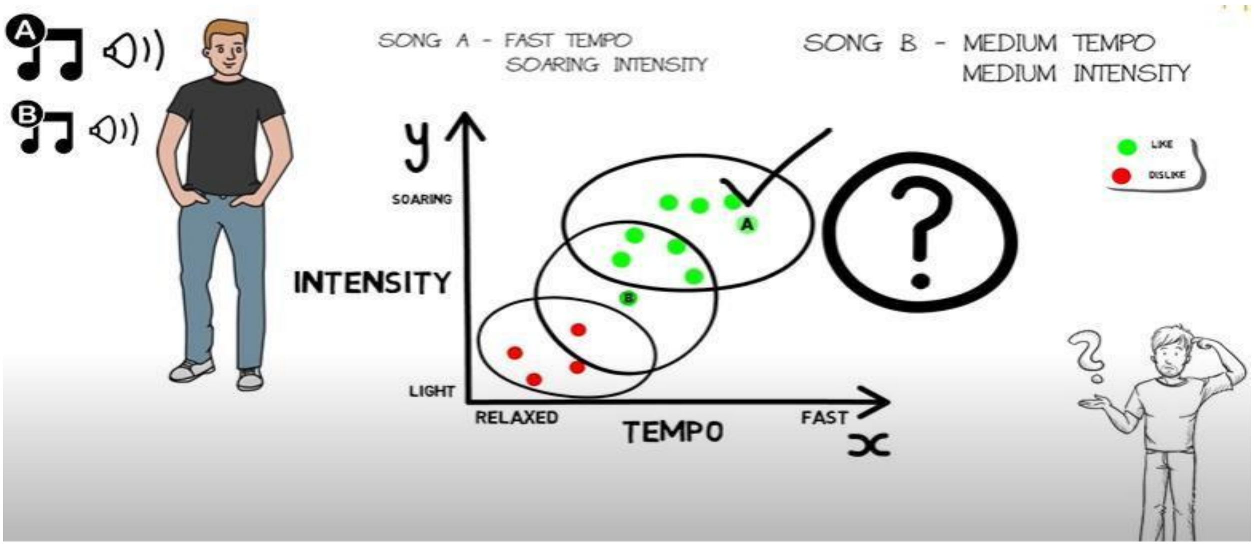

Figure 4.1: How Machine Learning Works

Types of Machine Learning:

There are basically many types of machine learning. But there are mostly common three types of machine learning and those are,

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning

4.1.1: Supervised Learning:

Supervised learning, also known as supervised machine learning, is a subcategory of machine learning and artificial intelligence. It is defined by its use of labeled datasets to train algorithms that to classify data or predict outcomes accurately. As input data is fed into the model, it adjusts its weights until the model has been fitted appropriately, which occurs as part of the cross-validation process. Supervised learning helps organizations solve for a variety of real-world problems at scale, such as classifying spam in a separate folder from your inbox [21].

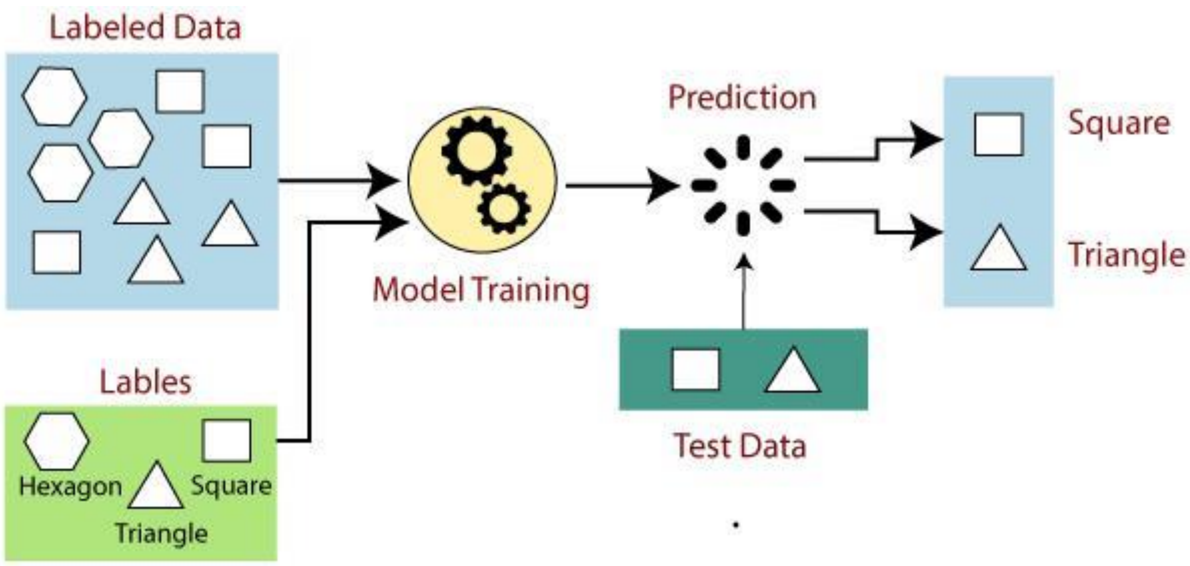

Figure 4.2: Supervised Learning

[Source: https://www.javatpoint.com/supervised-machine-learning]

In this thesis work, we will apply Four supervised learning algorithms

The following are some of the supervised machine learning approaches that were compared in this thesis**:**

Naïve Bayes Classifier:

It's a classification method based on Bayes' Theorem and the assumption of predictor independence. A Naive Bayes classifier states, in basic terms, that the existence of one feature in a class has no bearing on the presence of any other feature in the class. A Bayes classifier is a comparable classification algorithm to a classification technique (Nave Bayes) The Bayes theorem is used to create the Nave Bayes classifier, which is based on the premise of strong independence [16]. The concept of the Naïve Bayes classifier is explained with the help of Eq.

$$P(A|B) = \frac{P(A)\times P(B\vee A)}{P(B)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots[1]$$

When predictor B is already available, we calculate the likelihood of class A.

P(B) = B's prior probability

P(A) = denotes the class A prior probability.

P(B|A) = probability of predictor B occurring given class A.

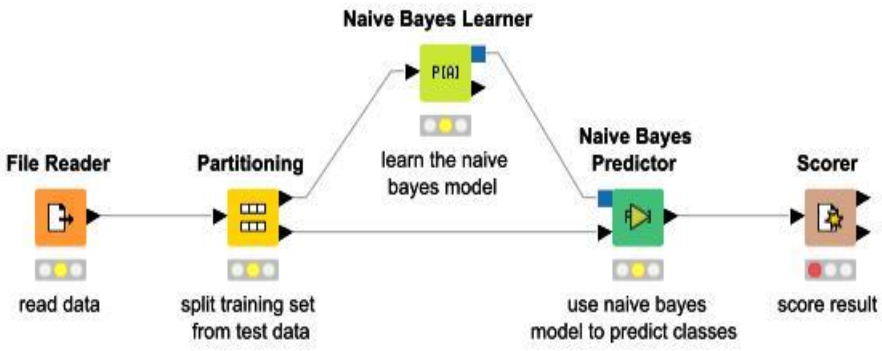

Figure 4.3: How Naïve bayes Algorithm works

[Source Link: https://i.stack.imgur.com/0QOII.png]

Support Vector Machine:

Support-vector machines are supervised learning models in machine learning that examine data for classification and regression analysis using learning methods. Support vectors are data points that are closer to the hyperplane and have an influence on the location and orientation of the hyperplane. These support vectors are used to increase the margin of the classifier. The hyperplane's location will vary if the support vectors are removed. These are the factors to examine while we construct our SVM [17].

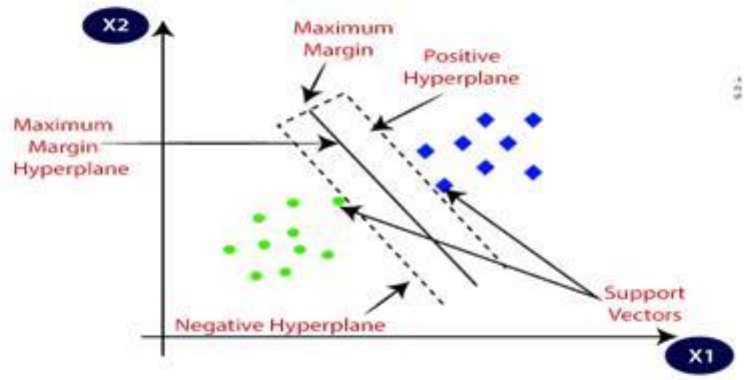

Figure 4.4: Support Vector Machine Algorithm.

[Source Link: https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm]

The base theorem for support vector machine is given below:

$R(\alpha)=Rtrain(\alpha)+\sqrt{\frac{f(h)}{N}}$ ………………………………[2]

$F(h)=h+h\log(2N)-h\log(h)-c$

$Margin=\rho$

Relative Margin$=\frac{\rho}{D}$

$h\leq \min(\{d,[\frac{D^2}{\rho^2}]\})+1$

Logistic Regression:

Logistic regression is one of the most often used Machine Learning algorithms in the Supervised Learning technique. It's a method for predicting a categorical dependent variable from a collection of independent factors. The output of a categorical dependent variable is predicted using logistic regression. As a result, the output must be discrete or categorical in nature. It can be Yes or No, 0 or 1, true or false, and so on, but instead of precise numbers like 0 and 1, it returns probabilistic values in the middle [18].

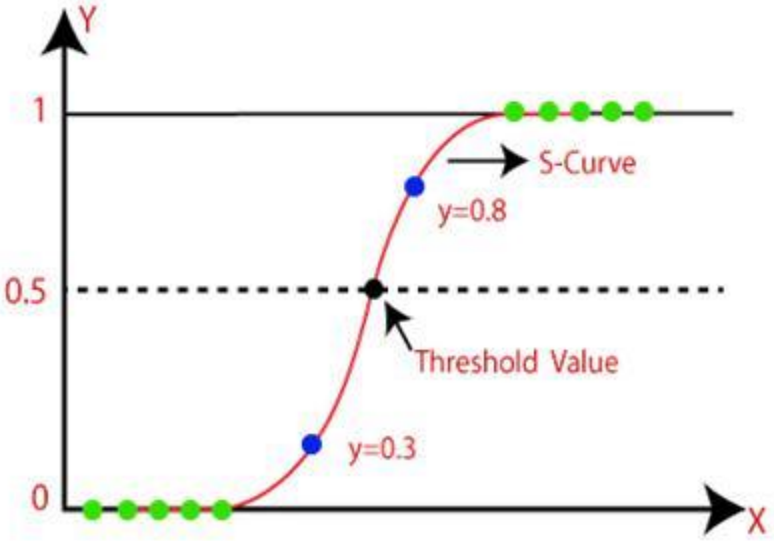

Figure 4.5: Logistic function models the conditional probability of the response.

[Source Link: https://www.javatpoint.com/logistic-regression-in-machine-learning]

The Logistic Regression can be obtain form the Linear regression equation. The mathematical step to get Logistic Regression equation are given below:

- The equation of the straight line cane be written as:

y = bo+b1+x1+b2+x2+b3+x3……+bn+xn……………………… [3]

- Logistic Regression y can be between 0 and 1 only ,so for this lets divide the above equation by(1-y):

$\frac{Y}{1-Y}$; 0 for y=0, and infinity for y=1

- But we need range between –infinity to +infinity ,then take logarithm of the equation it will become;

Log [$\frac{Y}{1-Y}$] bo+b1+x1+b2+x2+b3+x3……+bn+xn

The above equation is the final equation for the Logistic Regression.

4. Random Forest:

"Random Forest is a classifier that consists of a number of decision trees on various subsets of a given dataset and takes the average to improve the dataset's projected accuracy." Rather of relying on a single decision tree, the random forest collects forecasts from each tree and predicts the ultimate output based on the majority of votes[19].

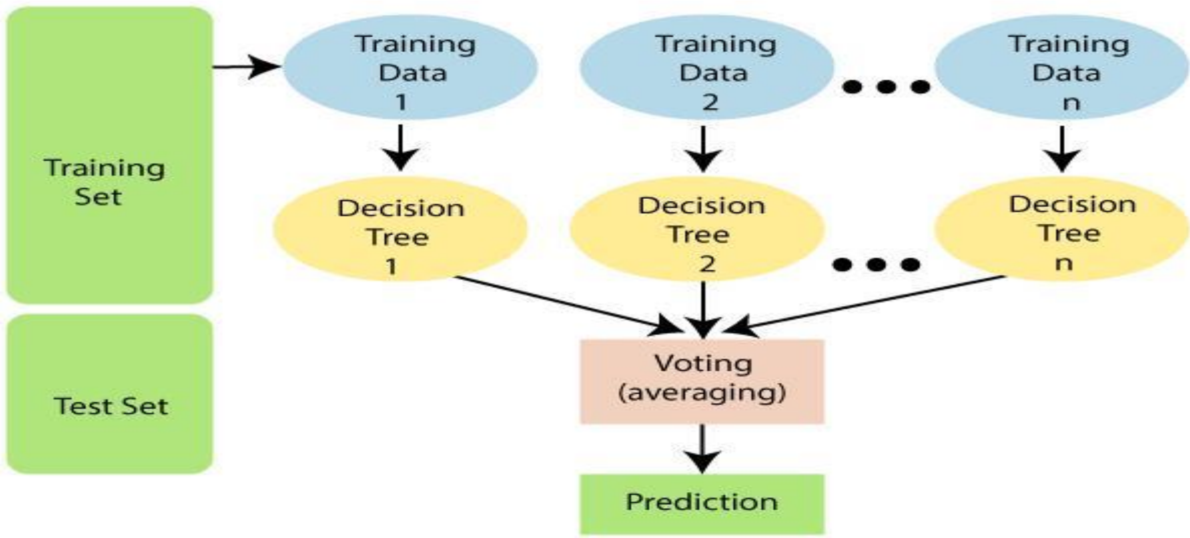

Figure 4.6: How Random Forest Algorithm works.

[Source Link: https://www.javatpoint.com/machine-learning-random-forest-algorithm]

4.1.2: Unsupervised Learning:

Supervised learning is basically labeled data learning but in unsupervised leaning, there are no labeled data. Unsupervised Learning is a machine learning technique in which the users do not need to supervise the model. Instead, it allows the model to work on its own to discover patterns and information that was previously undetected. It mainly deals with the unlabeled data [22].

There are no pre-classified training samples or training datasets in unsupervised learning, as demonstrated in Figure 03. Unsupervised learning makes use of data that hasn't been labeled. Unlabeled data is information that can be easily obtained from natural sources, such as photographs, text reviews, and videos. For each piece of unlabeled data, there is no explanation. In the classification of text, clustering is the most often used in unsupervised learning technique. Clustering can be done in a variety of ways, with K-means clustering being one of the most prominent.

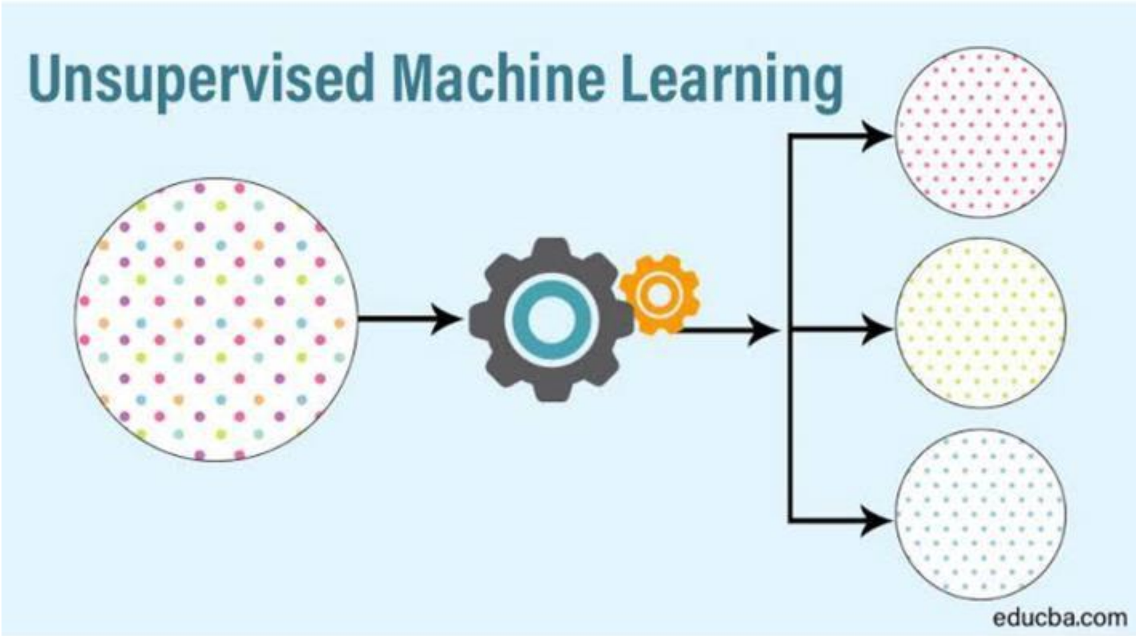

Figure 4.7: Unsupervised Learning

4.1.3: Reinforcement Learning:

Reinforcement Learning is a feedback-based Machine learning technique in which an agent learns to behave in an environment by performing the actions and seeing the results of actions. For each good action, the agent gets positive feedback, and for each bad action, the agent gets negative feedback or penalty [23].

Let us consider that, we have a model. Now we want to train that model by using reinforcement learning. Firstly, we show that model a dog picture. That model replies that the picture is the cat’s picture, which is false. Then we send feedback to the model that this picture is a dog picture. After that, whenever we show any kind of dog picture, the model will predict that as a dog picture by using previous experience.

This is how reinforcement learning work.

Figure 4.8: Reinforcement Learning

4.2: Evaluation Parameters:

Measurement of performance of a classifier is highly depends on evaluation matrix. For that reason accuracy is the most important things for a classifier. A classifier accuracy is the most sincere things when we perform a dataset through a model and train a dataset through a model after that test that dataset through that model. Then we find the exact accuracy by using a classifier.

Predict Positive(0)	True Positive(0)	False Positive(0)
Predict Negative(1)	False Negative(1)	True Negative(1)
	Actually Positive(0)	Actually Negative(1)

Table 4.1: Confusion Matrix

This is the table is confusion matrix where we can see that it shows the actual predicted scenario by using the machine learning algorithms.

• **True Positive (TP):** These are cases in which the tweet which is non hatred, also classified as positive by the classifier.

• **False Positive (FP):** False Positive are non-hatredtweet which the classifier classifies them as negative.

• **False Negative (FN):** False Negative are hatred tweet, but classifier classifies them as positive.

• **True Negatives (TN):** True Negative represents the tweet which are hatred and also classified as negative by the classifier.

Precision, recall, F-measure, and accuracy are the sentiment measures used to evaluate the performance of the classifier based on the data from the confusion matrix.

Precision: Precision is a metric that measure how accurate the classifier is, and how many of data returned as correct. A greater precision suggests that there are fewer false positives, whereas a lower precision means that there are more false positives. The ratio of properly categorized examples to total instances is known as precision (P).

$$\text{Precision} = \frac{Truth\ Positive}{Truth\ Positive + False\ Positive} \quad \text{.................................} \quad [4]$$

Recall: The sensitivity of a classifier is determined by recall, or how much positive data it returns. There are fewer false negatives when the recall is higher. Recall is defined as the proportion of successfully categorized examples to the total number of predicted occurrences.

$$\text{Recall} = \frac{Truth\ Positive}{Truth\ Positive + False\ Negative} \quad \text{................................} \quad [5]$$

F-measure: When precision and recall are combined, the outcome is the weighted harmonic mean of precision and recall, which is known as F-measure.

$$\text{F-measure} = 2*\frac{Precision * Recall}{Precision + Recall} \quad \text{....................................} \quad [6]$$

Accuracy: This is the percentage of times the classifier was right. We may achieve this by multiplying the total observations by the sum of the TP and the TN.

$$\text{Accuracy} = \frac{Truth\ Positive + Truth\ Negative}{Truth\ positive + Truth\ Negative + False\ Negative + False\ Positive} \ldots\ldots\ldots\ldots [7]$$

CHAPTER FIVE

Deep Learning

5.1: Deep Learning:

A computer model learns to perform classification tasks directly from images, text, or sound. Deep learning models can achieve state-of-the-art accuracy, sometimes exceeding human-level performance. Models are trained by using a large set of labeled data and neural network architectures that contain many layers [24].

The word "deep" in the phrase "deep learning" refers to the number of deep layers of a neural network. A neural network with more than two layers, some of which are hidden layers, is known as a deep neural network. Deep neural networks process input in a variety of ways using complex mathematical models. An adjustable model of outputs as functions of inputs, a neural network consists of several layers: an input layer, which includes input data; hidden layers, which include processing nodes known as neurons; and an output layer, which includes one or more neurons, whose outputs are the network outputs.

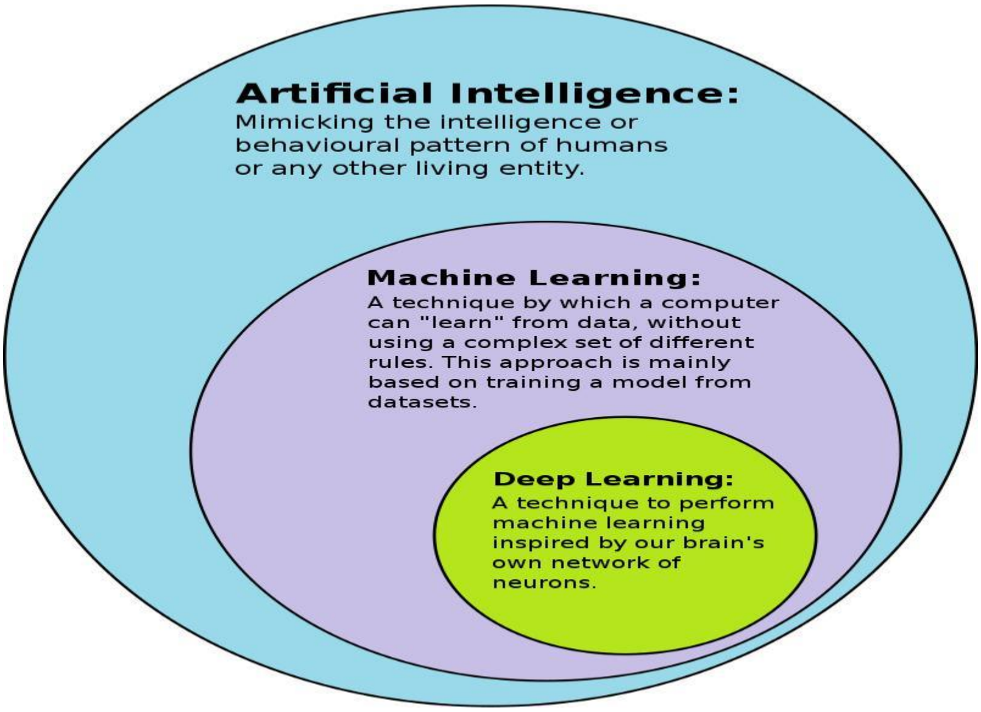

Figure 5.1: What is Deep Learning

[Source: https://en.wikipedia.org/wiki/Deep_learning#/media/File:AI-ML-DL.svg]

Deep learning always tries to work similarly with a human being. As human beings, we always try to think logically and rationally. Usually, other programming algorithms always work straightforward. Those algorithms always follow their instruction which is set in the program. On the other hand, Deep learning's algorithm is such an evolutionary invention in the computer science world. By deep learning, computers can work like a human brain. When we think to do something, we always think of the conscience. We obverse our around environment. Then we decided whether we should do that work or not.

That is why, deep learning is an evolutionary invention in computer science.

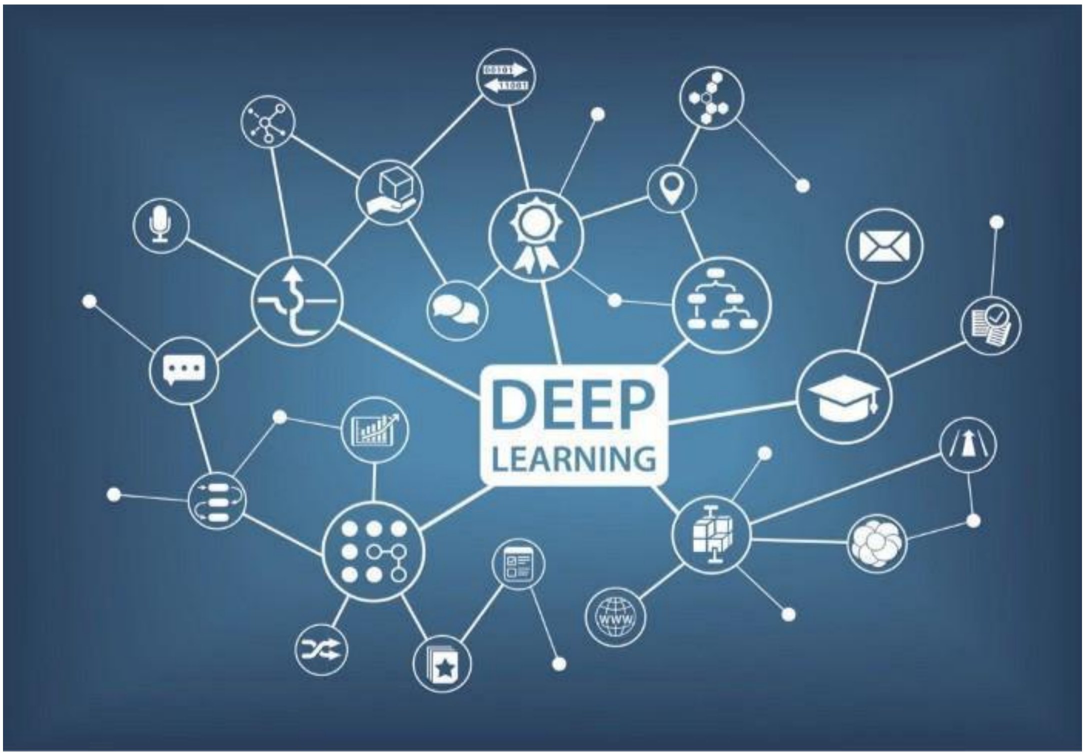

Figure 5.2: Deep Learning

[Source: https://towardsdatascience.com/does-deep-learning-really-require-big-data-no-13890b014ded]

5.2: Deep Neural Network:

Neural networks, also known as artificial neural networks (ANNs) or simulated neural networks (SNNs), are a subset of machine learning and are at the heart of deep learning algorithms. Their name and structure are inspired by the human brain, mimicking the way that biological neurons signal to one another[25].

A node layer store an input layer, one or more hidden layers, and an output layer in artificial neural networks (ANNs). Each node has a weight, and each node has some threshold value. If a node crosses the threshold value, the node will be on or active. And then data send to the next node of the network. And if the node do not cross the threshold, data will not sent to the next node.

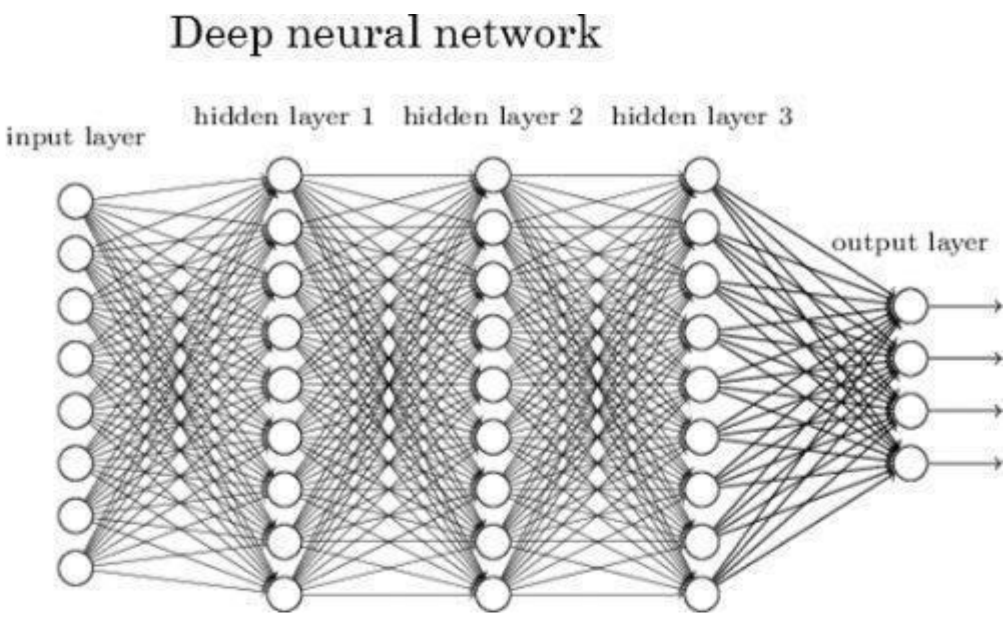

Figure 5.3: Deep Neural Network

5.3: How Deep Neural Network Works:

Let us consider that, there each node have a separate liner regression model, who have some input data, weights, a bias or threshold and an output. Formula:

∑wixi + bias = w1x1 + w2x2 + w3x3 + bias

output = f(x) = 1 if ∑w1x1 + b>= 0; 0 if ∑w1x1 + b < 0

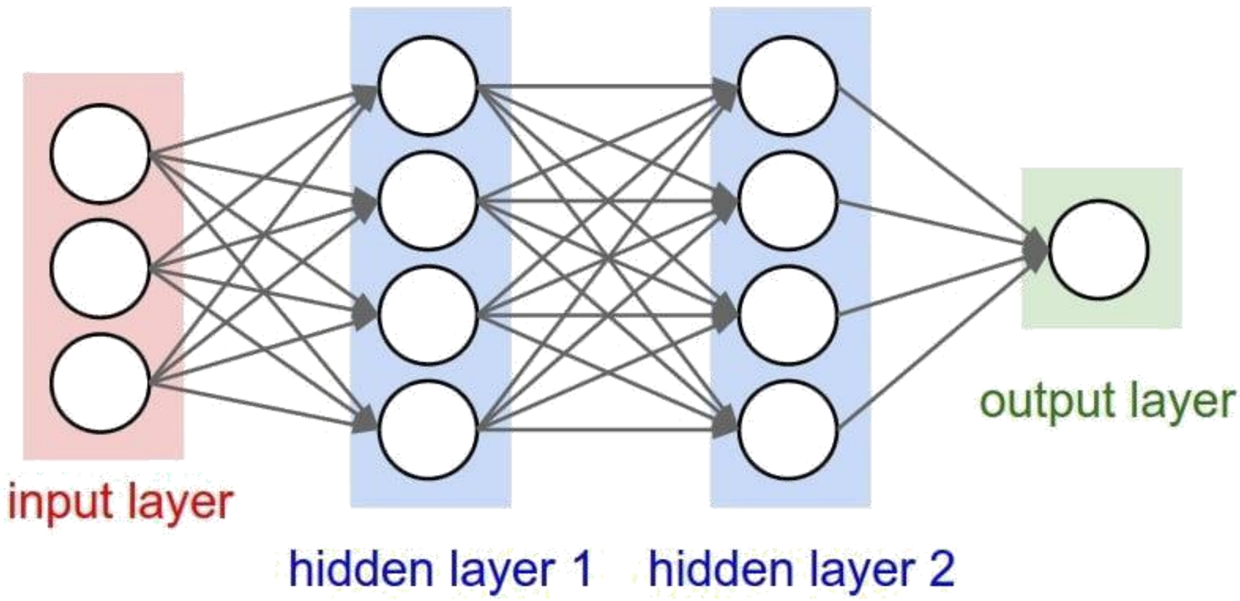

Figure 5.4: Basic Model of Deep Learning

[Source: https://www.bmc.com/blogs/deep-neural-network/]

According to the formula, if the input value and weight value together cross the bias or threshold value, the output will be one or positive, if not then output will be zero or negative.

Types of Neural Network:

There are different kinds of neural networks exist. Interestingly each neural network have different purpose. Here is some popular types of neural network:

Frank Rosenblatt invented the perceptron in 1958, and it is the earliest neural network. It is the simplest type of a neural network, with only one neuron:

The earliest neural network was invented by Frank Rosenblatt in 1958. This neural is the simplest type of neural network which has only one neuron.

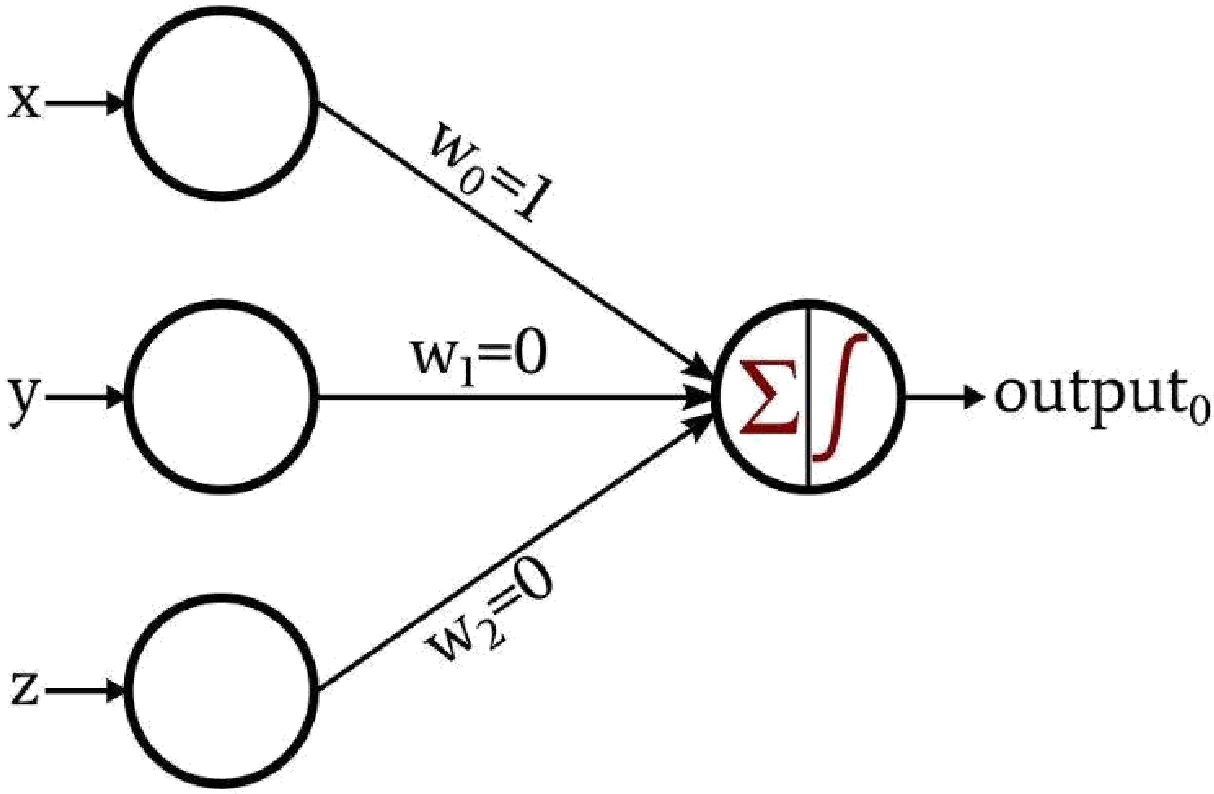

Figure 5.5: Perceptron

[Source: https://www.allaboutcircuits.com/uploads/articles/how-to-train-a-basic-perceptron-neural-network_rk_aac_image1.jpg]

After perceptron, we focused on feedforward neural network, who have multi layer perceptron's. This is known as Multi Layer Perceptron (MLP). MLP have more than one layer between input and output layer.

There are basically three different kinds of neural networks, those are,

- ANN - Artificial Neural Networks
- CNN - Convolution Neural Networks
- RNN - Recurrent Neural Networks

An artificial neural network is an attempt to simulate the network of neurons that make up a human brain so that the computer will be able to learn things and make decisions in a humanlike manner. ANNs are created by programming regular computers to behave as though they are interconnected brain cells [26].

Feedforward networks, convolutional neural networks (CNNs) basically used for to predict any image, pattern and computer vision.

The feedback loops distinguish recurrent neural networks (RNNs) basically used for predict future events.

5.3.1: RNN (Recurrent Neural Network):

Recurrent Neural Network(RNN) are a type of Neural Network where the output from previous step are fed as input to the current step. In traditional neural networks, all the inputs and outputs are independent of each other, but in cases like when it is required to predict the next word of a sentence, the previous words are required and hence there is a need to remember the previous words. Thus RNN came into existence, which solved this issue with the help of a Hidden Layer. The main and most important feature of RNN is Hidden state, which remembers some information about a sequence [27].

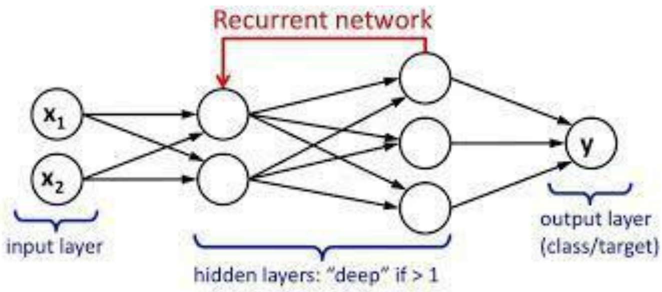

Figure 5.6: RNN

[Source: https://deepai.org/machine-learning-glossary-and-terms/recurrent-neural-network

The foundation for recurrent neural networks started by David Rumelhart. RNN discovered by John Hopfield in 1982.

RNN works interestingly. Let us see, how RNN works? Humans do not start thinking all over again every second. We compare word to our previous work. We do not read all the text all over again. We store our thoughts in our mind.

This is the traditional neural network or it is the fundamental flaw. Let us Consider, the following scenario: we divided types into many categorizes. It's unclear how a typical neural network might utilize prior events in the movie to guide subsequent ones.

This problem is addressed by recurrent neural networks. They're networks with loops in them that allow data to endure.

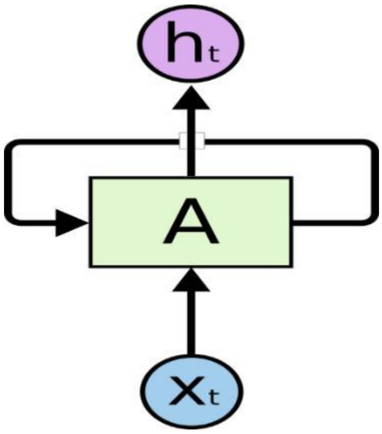

Figure 5.7: RNN have Loops.

[Source: https://colah.github.io/posts/2015-08-Understanding-LSTMs/]

LSTM (Long Short-Term Memory):

LSTM developed to prevent the problem of long-term reliance. They don't have to work hard to remember knowledge for lengthy periods of time; it's like second nature to them!

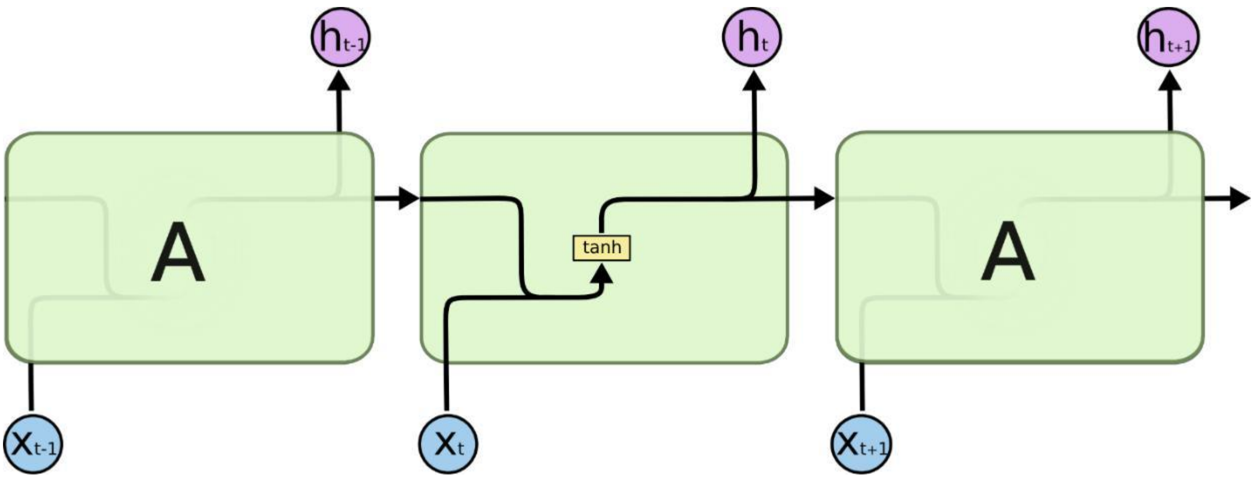

Figure 5.8: The repeating module in a standard RNN contains a single layer.

[Source: https://colah.github.io/posts/2015-08-Understanding-LSTMs/]

In LSTM, they work like a chain structure, although the repeating module is different. In RNN there are only one layer but in LSTM there are four layer and they connect to each other in a unique way.

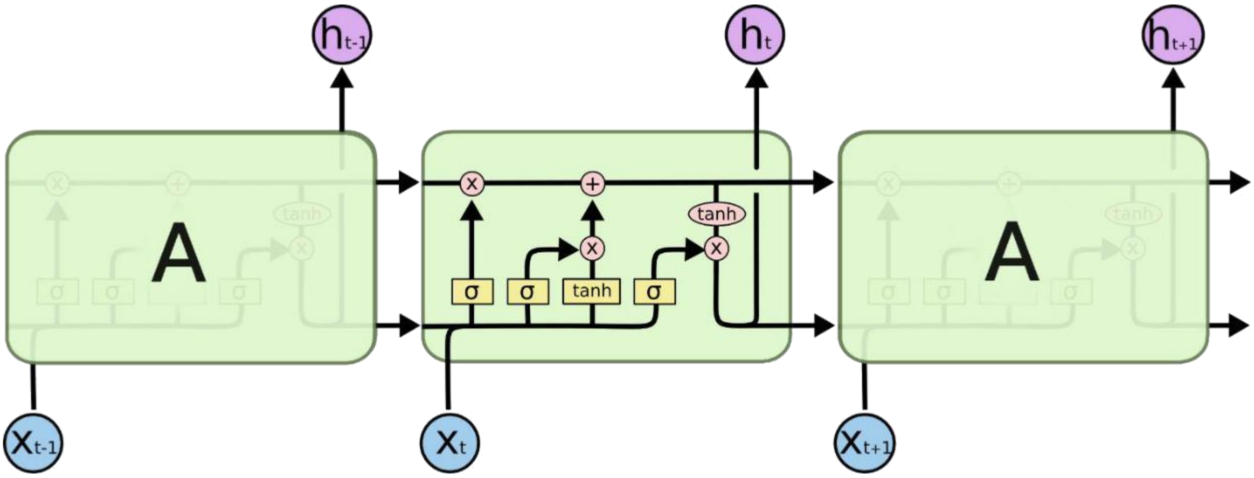

Figure 5.9: The repeating module in an LSTM contains four interacting layers.

[Source: https://colah.github.io/posts/2015-08-Understanding-LSTMs/]

LSTMs were a huge step forward in terms of what we could do with RNNs. It's reasonable to question if there will be another major move forward. "Yes!" said a majority of researchers. There's a next level, and it's all about paying attention!" The aim is to allow each stage of an RNN choose information to examine from a bigger pool of data. If you're using an RNN to build a caption for a picture, it can choose a different region of the image to look at for each word it produces. In fact, Xu, et al. (2015) do just that - it may be a good place to start if you want to learn more about attention! There have been a lot of really interesting breakthroughs employing attention, and it appears to be a promising field.

CHAPTER SIX

Research Methodology

6.1: Methodology:

Twitter is one of the famous social media sites in this world. We use twitter data for sentiment analysis that the tweet is hatred or not. This section of the thesis details how the selected methods have been implemented. We used data based on Twitter data and the dataset has been obtained by the following kaggle link [42]:
The data of this dataset is label and use it in a supervised learning model.

In our selecting dataset the data structure as follows:

“id”: ID of the Tweet account

“label” is containing the sentiment, if label is 0 it means this tweet is non hatred, if label is 1 it means this tweet is hatred.

“tweet” it defines the tweet.

	id	label	tweet
0	1	0	@user when a father is dysfunctional and is s...
1	2	0	@user @user thanks for #lyft credit i can't us...
2	3	0	bihday your majesty
3	4	0	#model i love u take with u all the time in ...
4	5	0	factsguide: society now #motivation
5	6	0	[2/2] huge fan fare and big talking before the...
6	7	0	@user camping tomorrow @user @user @user @use...
7	8	0	the next school year is the year for exams.ðŸ˜...
8	9	0	we won!!! love the land!!! #allin #cavs #champ...
9	10	0	@user @user welcome here ! i'm it's so #gr...
10	11	0	â† #ireland consumer price index (mom) climb...
11	12	0	we are so selfish. #orlando #standwithorlando ...
12	13	0	i get to see my daddy today!! #80days #getti...
13	14	1	@user #cnn calls #michigan middle school 'buil...
14	15	1	no comment! in #australia #opkillingbay #se...
15	16	0	ouch...junior is angryðŸ˜#got7 #junior #yugyo...

Figure 6.1: The small sample screenshot of the dataset of the Twitter Data.

In this thesis, we will try to analyze the sentiment a dataset of hatred or non-hatred tweet with both traditional machine learning algorithms including Naive Bayes, Logistic regression method, Linear Support Vector Machines (SVM), Random Forest and deep neural networks using such as Bidirectional Long Short-Term Memory (LSTM) is used.

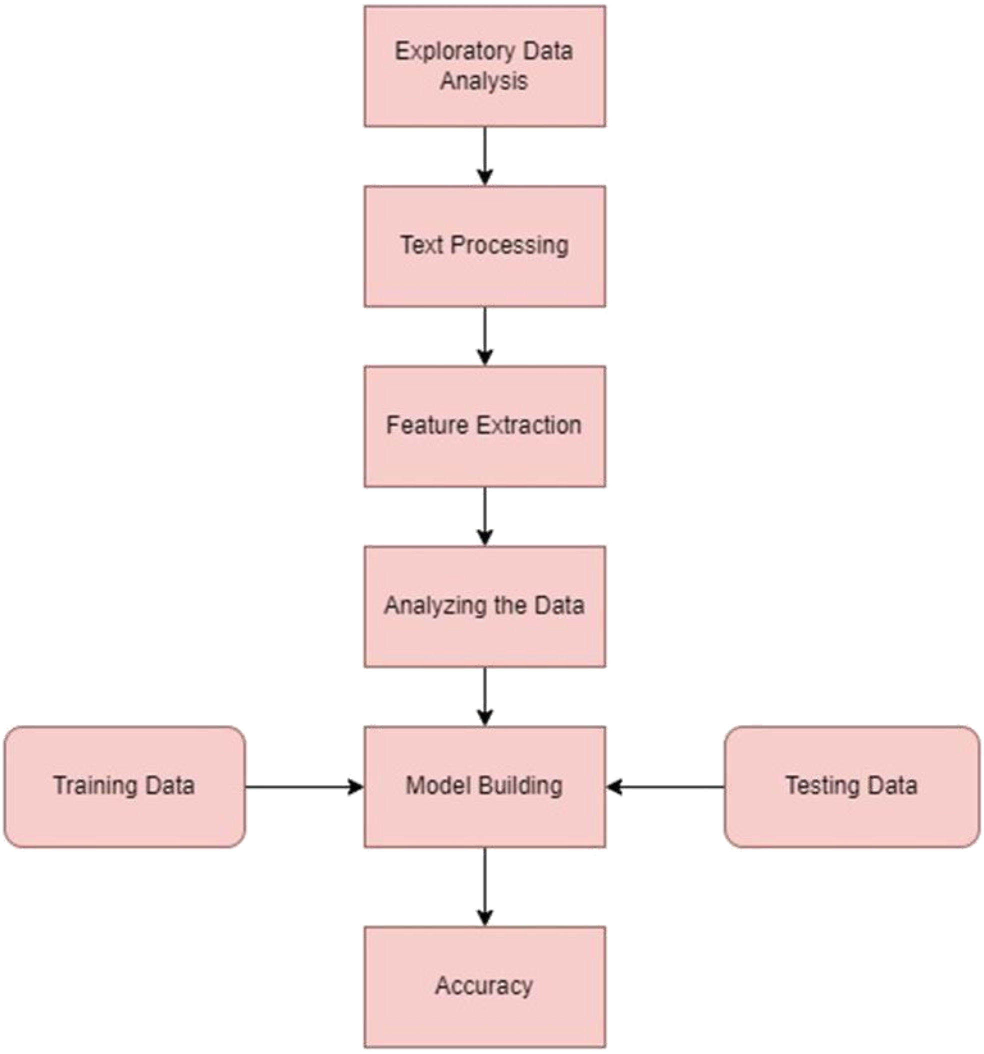

Figure 6.2: Work Process

(i) Data Pre-Processing

Our dataset is all about the sentiment analysis, for that we use two dataset one is train dataset and another one is test dataset. After processing the data, we combine both the dataset and passes through to the model for known the sentiment of that dataset.

We have tried to implement a sentiment analysis based on the tweet.

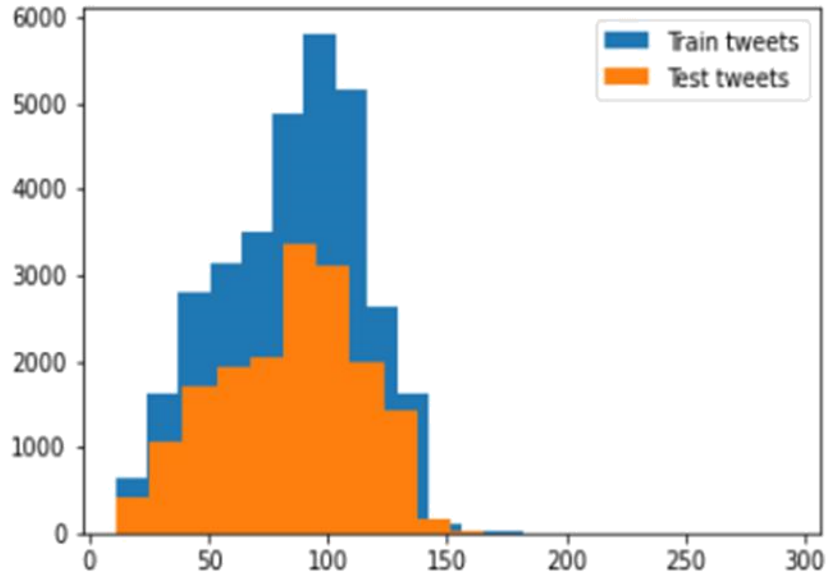

Figure 6.3: Distribution of train and test dataset.

Here we show two color graphs. The blue one is for train tweets and the brown one is for test tweets. Also, we can show that train tweet is almost 5800 tweets, and test tweet is 3200.These two datasets helps us to find the sentiment of the dataset.

After that we find the amount of the hatred tweet and non-hatred tweet. Also, we determine a graph for the representation. This graph is given below:

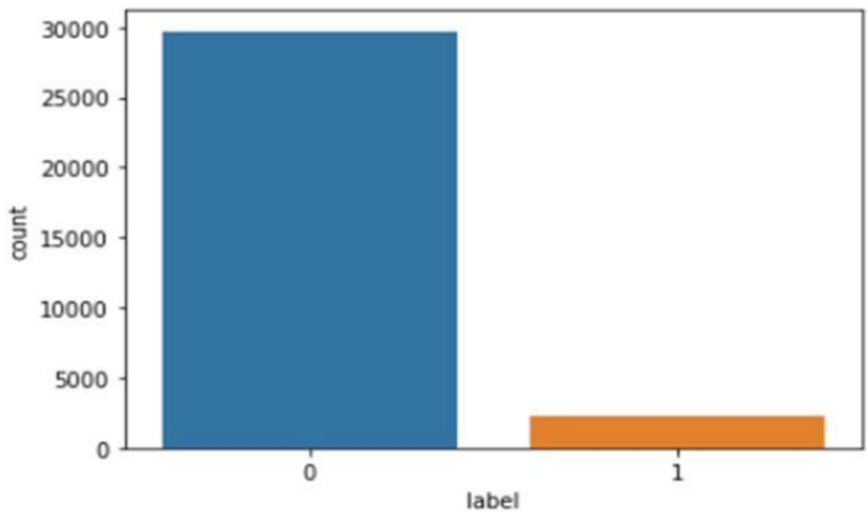

Figure 6.4: Amount of hatred and non-hatred tweet in our dataset.

In the x axis we define label of the sentiment and y axis we define the count of hatred and non-hatred tweet Here we show that a blue diagram that label is 0, it defines the non-hatred tweet. The amount of non-hatred tweet is almost 30000.The brown diagram that label is 1, it defines the hatred tweet. The amount of hatred tweet is 3000.

After that we have tried to implement the distribution of length of the hatred tweet and non-hatred tweet. We also find two graphs for the class 0 that defines non-hatred tweet and class 1 that defines hatred tweet.

Both the graph is given below:

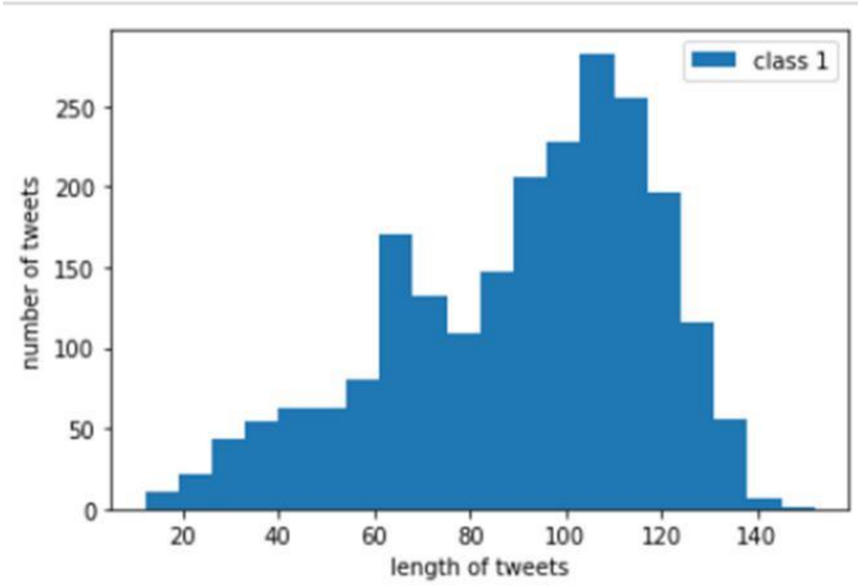

Figure 6.5: Distribution of length of the hatred tweet.

This is the distribution graph of length of the Hatred tweet. In the x axis we define the length of tweet and the y axis define the number of tweets for the class 1. Length of the tweet means that how many words on that tweet. Number of tweets defines that how many tweet for that word that define in the x axis.

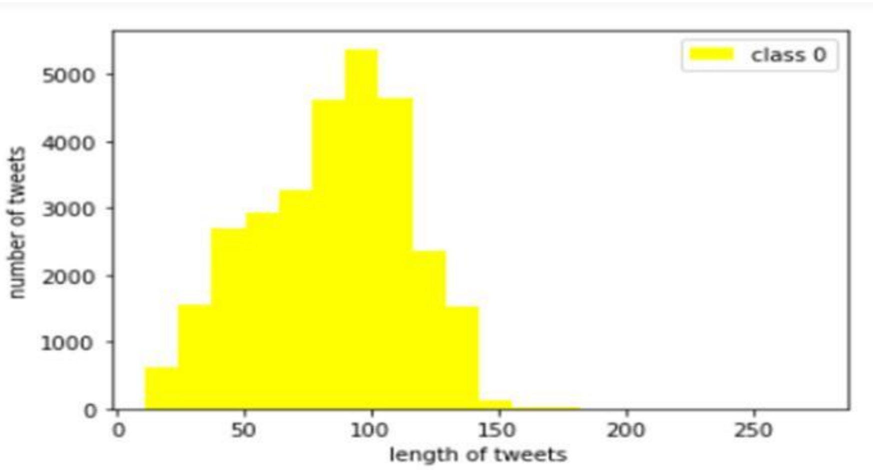

Figure 6.6: Distribution of length of the non-hatred tweet.

This is the distribution graph of length of the non-hatred tweet. In the x axis we define the length of tweet and the y axis define number of tweets for the class 0. Length of the tweet means that how many words on that tweet. Number of tweets defines that how many tweet for that word that define in the x axis.

We then start to verify the most common words used in text reviews using a word cloud to facilitate analysis. Also, we verify the normal words and find the positive words and negative words for determine the hatred words and non-hatred words. However, since we haven’t done preprocessing before we can see some words that make little sense to the text.

Most used words:

Figure 6.7: World Cloud view for most used word in all tweets.

This is most common words that used in all the tweets. Those all words nothing to do in our sentiment analysis. Now we differ and find the positive word and negative word for our sentiment analysis. Now we separately compute the positive and negative words.

Figure 6.8: World Cloud view for normal word in all tweets.

This is the view for all normal words that has no relation with the sentiment analysis or those word can not define the positive or negative word. That is why we find those word and after that we find the positive and negative word that actually need for sentiment analysis.

Figure 6.9: World Cloud view for positive word in all tweets.

This is the view for all the positive words that uses in the tweets. Those word define that the specific tweet is non-hatred tweet. We all know that positive word is only used for the positive thing. That is why all the positive word define the non-hatred tweet.

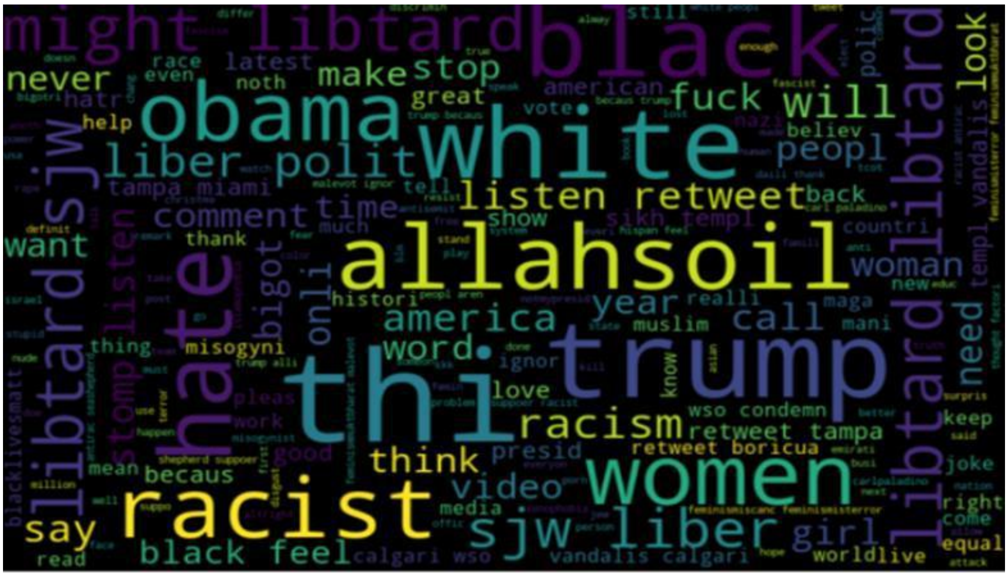

Figure 6.10: World Cloud view for negative word in all tweets.

This is the view for all the negative word that uses in the hatred tweets. Those word can define the specific tweet is hatred tweet. We all know that negative word is only used for the negative things. That is why all the negative word define the hatred tweet.

We now tried to know the actual word quantity for the positive word and negative word. That is why we extract top 20 word for the positive and negative word. Both the figure is given below:

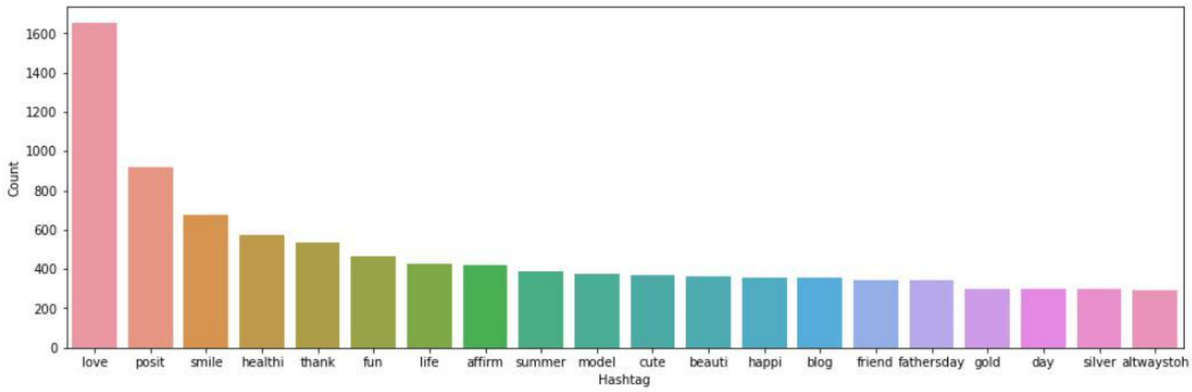

Figure 6.11: Quantity of top 20 positive word.

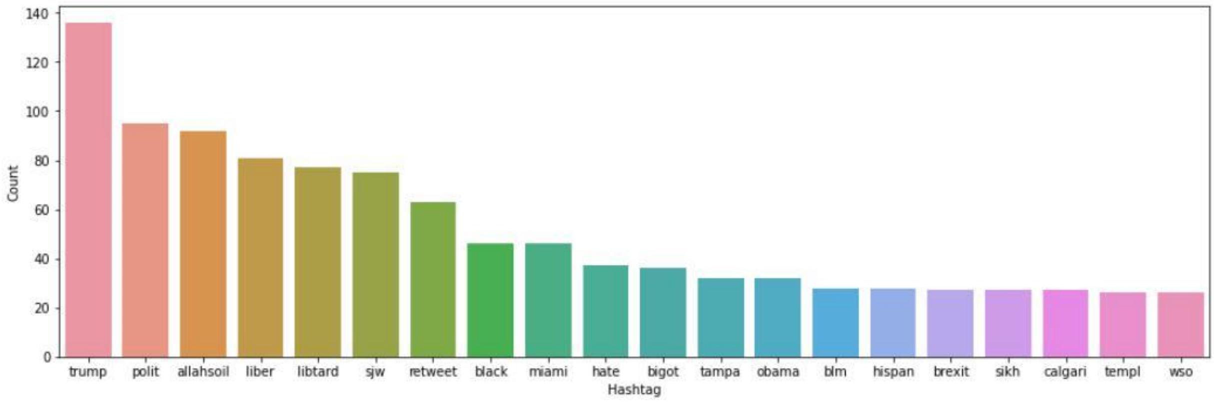

Figure 6.12: Quantity of top 20 negative word.

(ii) Text Processing

Clean Text:

The first step for any data analysis to first clean the dataset so that, we have come up for some simple technique for clean and prepare the text data to passes through the dataset into machine and deep learning algorithm.

We use those method for clean the dataset:

- Remove pattern

- Remove @
- Remove hashtag
- Remove short-word
- Remove article
- Split the white space
- Replace digit with spaces
- Remove username
- Removing Punctuations, Numbers, and Special Characters

Remove username:

We tried to analyzing the sentiment of a tweet. So that we don’t need any user name of the tweet. That is why we remove the username from all the tweet of our test dataset.

Removing Punctuations, Numbers, and Special Characters:

In the text processing we have to clear the punctuations, number and special characters to our dataset. That is why remove all of them form the test dataset.

Remove Pattern and @:

All the @ and pattern doesn’t need for the text processing that is why we tried to remove all of the things from our dataset.

Remove Hashtag:

Hashtag must exist in all the tweet but for sentiment analysis we don’t need that kind of things for our training model. That is why we remove all the hashtag from the dataset.

Remove Stop-words:

We have lots of tweet in our dataset, all of the tweet may contain some words that are too frequent and that do not contribute much to the overall meaning of the sentence. We've made a list of some of those words and removed them. For example: "a", "an", "the", "this", "that", "is", "it", "to", "and" in this example.

Tokenization:

Tokenization is very important things for sentiment analysis and passes through the dataset into the model. Because of we cannot passed the hole dataset into the model. We have to convert the dataset and combine it into the word. Tokenization doing that thing it removes the white space from the dataset and convert a sentence into several word.

Stemming:

Stemming is the process that reducing inflected words to their word stem, base, or root form— generally a written word form.

Example:

"cleaned" -> "clean"

It is the rule-based method sometimes its mistakenly removes suffix from the word.

Lemmatization:

Lemmatization is the process that has the clear identification about the noun, verb and adjective. Also, we use NLTK for text processing. Lemmatization converts the word to its base form that a word is noun, verb or adjective etc.

(iii)Analyzing the data

Splitting the Dataset into Train set and Test set:

The Data is divided into 2 sets of data:

• Training Data: The dataset which the model would be trained on. Contains 80% data.

• Test Data: The dataset which the model would be tested against. Contains 20% data.

(iv) Model Building

Algorithm for proposed approach

Input:

Labeled data

Output:

1.Load Data

2.Exploratory Data Analysis

3.Data Preprocessing

4.Tokenization

5.Stemmimg

6.NLTK

7.Lemmantization

8.Extract Feature

9.Spliting Dataset into test and train set

10.Combine the dataset

11.Model Implementation

12.Confusion Matrix

13.Accuracy

14.End

Accuracy:

1.Train Accuracy

2.Test Accuracy

Also measure the Precision, Recall, F1-score.

CHAPTER SEVEN

Result and Analysis

7.1: Result and Analysis:

In this analysis, we study about the sentiment analysis of twitter data. That one specific tweet is hatred tweet or non-hatred tweet. For analyzing the data set we found a dataset that has 29530 tweets from various person in a combine dataset. This data set has the information about the tweet is hatred tweet or non-hatred tweet. Also, it has a label that define, if the tweet is hatred, it shows 1 and if the tweet is non-hatred, it shows 0.

In our experiments we choose four algorithms from machine learning and one algorithm from deep learning. The algorithms are: Support Vector Machine, Naïve Bayes Classifier, Random Forest, Logistic Regression and from Deep Learning it is Bidirectional Long Short Term Memory (Bi-LSTM). We produce the best classification on twitter data and see the result according to our evaluation system. All the algorithms were applied to various selection process and find the confusion matrix with the help of TF-TP and FP-FN. All the algorithms give the best result and also, we find the precision, recall and f1-score for all the algorithms after that find the accuracy for all the algorithms.

7.2: Applying Machine Learning Algorithms:

Algorithm	**Train Accuracy (%)**	**Test Accuracy (%)**	**Precision**	**Recall**	**F1-Score**
Support Vector Machine	**82.36**	**94.86**	**0.95**	**0.89**	**0.91**
Naïve Bayes Classifier	**80.25**	**94.08**	**0.92**	**0.87**	**0.89**
Random Forest	**97.74**	**96.24**	**0.98**	**0.90**	**0.93**
Logistic Regression	**77.32**	**95.01**	**0.96**	**0.88**	**0.92**

Table 7.1: Experimental Result of Machine Learning Algorithms

From the experimental result we can see that, Random Forest model provides the best accuracy for train and test data set. Also, we find the best final accuracy with Random Forest model. The F1-score is also high for the Random Forest model. Random forest model is the privilege model for our experiment.

By using the train and test accuracy, precision, recall and f1 score we can see that how good is our model learned and classify every model clearly. Our each and every model gives the best result for our dataset. Each and every model produce the best result when it comes to train accuracy, test accuracy, precision, recall and F1 score. One models accuracy is very close to one another model. Random forest gives the best accuracy in all train, test, precision, recall and f1 score. We know that Random Forest is the best model for the sentiment analysis.

The Support Vector Machine, Naïve Bayes Classifier, Random Forest, Logistic Regression all the model produces some excellent result.

For Support Vector Machine model has the train accuracy is 82.86%, test accuracy 94.86% and f1 score is 0.91. This model also produces good result in terms of accuracy. This model works like that, it produces two classifiers from the dataset, that is 1 and 0. After that its draw a hyper plane between two classifiers. Then maximize the distance between the hyperplane of the two classifiers. After that choose the best result from that. Also produce the weighted vector and choose from that which is the best result. That is why Support Vector Machine has excellent result for our dataset.

Now the Naïve Bayes Classifier model has the lowest accuracy in our experiment. Because of this model works like that, it always counts the words after that it calculates the probability of every word while it's training the model. After that when the new sentence comes its divide the sentence into words and then match the probability that the model finds before. Then it says that the word is positive word or negative word if we want to compute the sentiment of the sentence. But sometimes it detects the wrong word for the specific sentiment. That is why this model does not gives the best result while detection something. That is why this model gives the worst result in our experiment. It has 80.25% train accuracy and 94.08% test accuracy and f1 score is 0.89 which is the lowest compare to the other model.

After seeing the train accuracy, test accuracy and f1 score compare to all the model. The Random Forest has produced the best result in our experiment. Random Forest classifier the data set has a score of 0.93 for the f1 score and accuracy of 96.24%, which is still high compared to the other following models. It has very interesting result compare to the other model's performance. Random Forest gives the best result because of we trained our dataset through bagging which improves the accuracy in machine learning algorithms. Also, Random Forest handling the missing data very well. The overall performance compare to other model Random Forest model represents the ideal model for our dataset.

Logistic Regression works like that, it works like the binary method. It only counts by using the 0 and 1. It will work like binary classification and works with binary methods. Our dataset also works like the binary methods. If the tweet is hatred its label is 1, if the tweet is non hatred its label is 0. That’s why Logistic Regression gives the second-best accuracy for our dataset. It has the train accuracy 77% and test accuracy 95.01% and f1 score is 0.92.

After all the discussion and seeing the experimental result, we can see that Random Forest model has the best result in our experiment. Random Forest has massive accurate power to beat the other model that we use in our experiment. The overall performance in our experiment we can say that Random Forest model gives the best result for our dataset.

7.3: Applying Deep Learning Algorithms:

In our experiment we apply Bidirectional Long Short Time Memory (Bi-LSTM) model from the deep learning model. We got **(95.22%)** accuracy by applying the Bi-LSTM model. It gives a excellent result in our experiment.

Algorithm	**Accuracy (%)**	**Precision**	**Recall**	**F1-Score**
Bi-LSTM	**95.22**	**0.83**	**0.79**	**0.81**

Table 7.2: Experimental Result of Deep Learning Algorithm

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
embedding (Embedding)        (None, 70, 40)            1200000
_________________________________________________________________
bidirectional (Bidirectional (None, 200)               112800
_________________________________________________________________
dropout (Dropout)            (None, 200)               0
_________________________________________________________________
dense (Dense)                (None, 1)                 201
=================================================================

Epoch 1/10

2022-02-23 14:13:43.959706: I tensorflow/stream_executor/cuda/cuda_dnn.cc:369] Loaded cuDNN version 8005

335/335 [==============================] - 10s 15ms/step - loss: 0.1922 - accuracy: 0.9411 - val_loss: 0.1341 - val_accuracy: 0.9554
Epoch 2/10
335/335 [==============================] - 4s 12ms/step - loss: 0.0764 - accuracy: 0.9730 - val_loss: 0.1362 - val_accuracy: 0.9553
Epoch 3/10
335/335 [==============================] - 4s 13ms/step - loss: 0.0422 - accuracy: 0.9860 - val_loss: 0.1584 - val_accuracy: 0.9585
Epoch 4/10
335/335 [==============================] - 4s 12ms/step - loss: 0.0327 - accuracy: 0.9891 - val_loss: 0.1871 - val_accuracy: 0.9527
Epoch 5/10
335/335 [==============================] - 4s 13ms/step - loss: 0.0170 - accuracy: 0.9953 - val_loss: 0.2152 - val_accuracy: 0.9511
Epoch 6/10
335/335 [==============================] - 4s 13ms/step - loss: 0.0103 - accuracy: 0.9977 - val_loss: 0.2527 - val_accuracy: 0.9521
Epoch 7/10
335/335 [==============================] - 4s 13ms/step - loss: 0.0074 - accuracy: 0.9982 - val_loss: 0.2615 - val_accuracy: 0.9522
Epoch 8/10
335/335 [==============================] - 4s 13ms/step - loss: 0.0059 - accuracy: 0.9986 - val_loss: 0.2973 - val_accuracy: 0.9509
Epoch 9/10
335/335 [==============================] - 4s 12ms/step - loss: 0.0064 - accuracy: 0.9981 - val_loss: 0.2842 - val_accuracy: 0.9502
Epoch 10/10
335/335 [==============================] - 4s 12ms/step - loss: 0.0053 - accuracy: 0.9988 - val_loss: 0.3066 - val_accuracy: 0.9522
```

The output of the embedding layer is a 2D vector with one embedding for each word in the input sequence of words. We get the final output by active the one hot encoding. By using that it covert categorical variables as binary vectors. After that each integer value is represented as a binary vector that is all zero values except the index of the integer, which is marked with a 1. Fit the model and run it for 10 epochs after that.

7.4: Applying confusion matrix:

For more observation in our experiment, we further implement the confusion, Matrix. In our experiment we use the confusion matrix function to show the confusion matrix for our all models that we use in our experiment.

We implement the confusion matrix because we want to more clear observation that our all models can detect how many values predict wrong or right. Confusion Matrix gives us the clear idea about the missing values or models detect how many values correctly. Also, it shown that which can also identify what kind of error our model is creating. We have applied confusion matrix on classifier models such as Support Vector machine, Naïve Bayes Classifier, Random Forest, Logistic Regression.

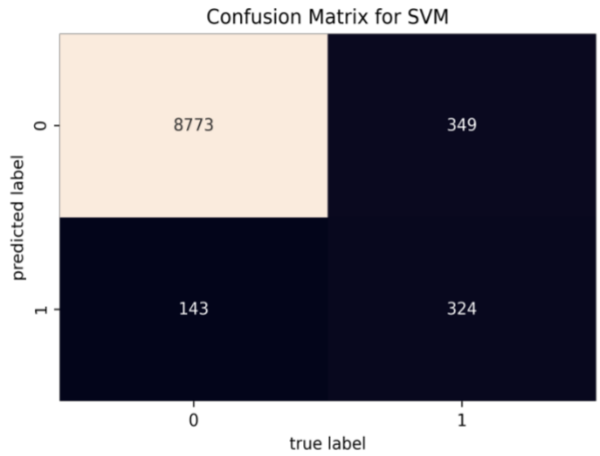

Figure 7.1: Confusion Matrix of SVM

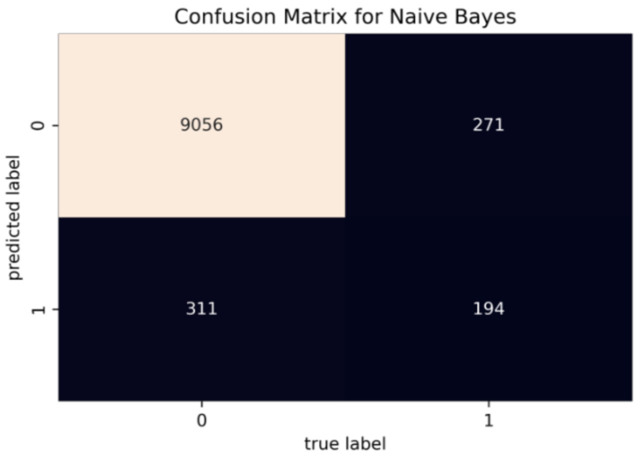

Figure 7.2: Confusion Matrix of Naïve Bayes

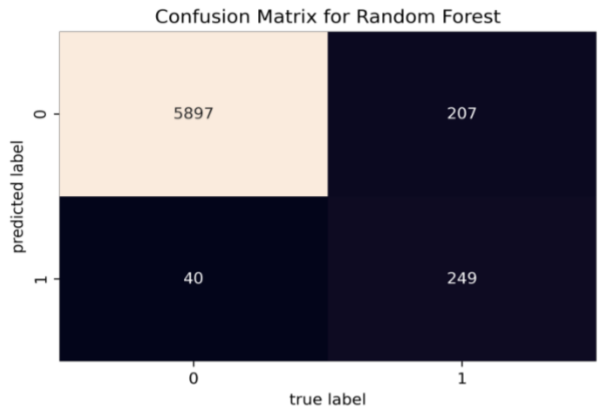

Figure 7.3:Confusion Matrix of Random Forest

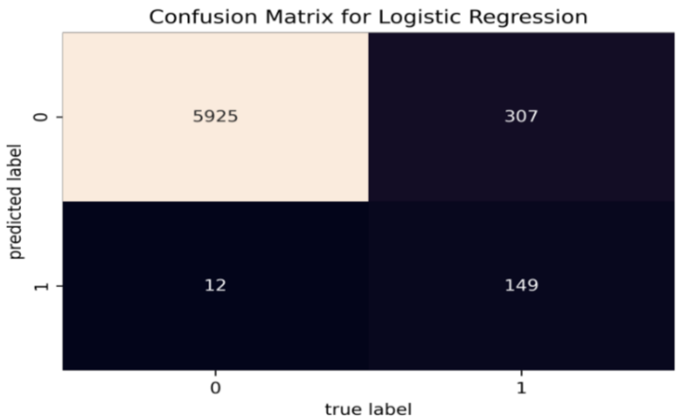

Figure 7.4: Confusion Matrix of Logistic

CHAPTER EIGHT

Comparison Between ML and DL

8.1: Advantages of Machine Learning:

Machine learning has already attracted a lot of interest as a result of its improved accuracy when it comes to predicting outcomes.

There are many advantages in machine learning and those are,

- Machine learning can easily identify latest trends and patterns. Machine learning can easily review large volumes of data and discover latest trends and patterns.
- In machine learning, there is no need for human intervention needed. This is automation.
- Machine learning keeps improving in accuracy and efficiency. So, ML have continued improvement.
- Machine Learning algorithms are good at handling data that are multi-dimensional and multi-variety, and they can do this in dynamic or uncertain environments.
- Machine learning have wide applications.

8.2: Challenges of Machine Learning:

Machine learning has some challenges and those are,

- Machine learning requires a large number of data sets to train the model. And this data set should be in good quality. So, machine learning has a challenge in data acquisition.
- Machine learning needs enough time to train the model. To fulfill or gain the accuracy, machine learning needs massive resources of function. So, there have challenges in time and resources.
- Another major challenge is the ability to accurately interpret results generated by the algorithms. We must also carefully choose the algorithms for our purpose.

8.3: Advantages of Deep Learning:

Deep learning has already attracted a lot of interest as a result of its improved accuracy when it comes to predicting outcomes. Large volumes of data were used to train the algorithm. There are various advantages of deep learning.

- Deep learning has the capacity to generate new features from a small set of inputs.
- One of its main benefits over all other machine learning algorithms is the inclusion of features in the training dataset.
- A modern way of education as a result, deep learning algorithms might create new tasks in order to improve performance to complete those that already exist
- Deep learning is highly scalable due to its ability to analyze massive amounts of data and perform multiple calculations at cheap cost and in a short amount of time. This has a direct impact on output, modularity, and portability.

8.4: Challenges of Deep Learning:

As new deep learning applications develop, so do the difficulties that must be solved. The Deep Learning Algorithm has a number of difficulties:

- Deep learning requires high-configuration technology to do such complicated computational calculations. As a result, creating a deep learning model is prohibitively expensive.
- There is no standard theory to help you choose the right deep learning tools since it demands understanding of topology, training technique, and other properties. As a result, it's quite tough for those with little skills to embrace it.
- One of the most fundamental difficulties with deep learning is a lack of transparency. On the inside, we don't get to see how the model works. As a result, discovering any model problems becomes more challenging.

8.5: Comparison Between Machine Learning and Deep Learning:

In our experiment, we used two types of models one is machine learning and another one is deep learning for sentiment analysis. We applied few algorithms from machine learning model and only one algorithm from deep learning model. First, we applied machine learning algorithms, its gives us good accuracy rate. After that we applied deep learning algorithm to find better accuracy rate than machine learning model, but surprisingly machine learning model gives the best accuracy rate in our experiment. We used both the model, so that we can give the best view on sentiment analysis.

Model	Applied Algorithms	Accuracy Rate (%)
Machine Learning	Support Vector Machine	94.86
	Naïve Bayes Classifier	94.08
	Random Forest	96.24
	Logistic Regression	95.01
Deep Learning	Bi-LSTM	95.22

Table 8.1: Comparison between the accuracy rates of ML & DL

From the table of accuracy difference, we see that not a huge difference between machine learning model compare to deep learning model. Both the model gives the fair enough accuracy rate. We just get 1% accuracy difference, if we compare those two models.

Now we give the reason why machine learning gives the best accuracy rate in our experiment:

- In machine learning model we can train the model with few amounts of data but in terms of deep learning we have to train the model with large amount of data. In our dataset we only have 29530 tweets, we cannot say that this dataset is a big dataset. That is why machine learning model give the best accuracy rate in our experiment.

- We all know that deep learning model has the higher accuracy rate than machine learning model, but surprisingly in our experiment we get the higher accuracy rate while applied machine learning model.

- When we applied the deep learning model, that time we used 10 epochs for the Bi-LSTM model. That epoch size is maybe too much for our dataset, the trained model may be becoming over fit, and maybe that's why we don't get the best accuracy rate by using deep learning model.

- To be qualified for the deep learning model there has to be at least three layers, but in machine learning model we just need one input and one output with barely one hidden layer. In our experiment we just give one input that is the tweet as a text and we show one output that the tweet is hatred or non-hatred. That is why we get the best accuracy rate in machine learning model.

- For using the deep learning model requires high performance hardware, but in machine learning model high performance hardware is not required. We run our experiment on a mid-performance hardware. Maybe that is why we don't get higher accuracy rate from the deep learning model.

- Machine learning cannot perform automatic feature extraction but on the other hand deep learning performs automatic feature extraction without human involvement.

- In machine learning model it divides all the task into small portions and then forms a combine effect. It helps to handle with the missing the data very well. If a model handle with the missing data very well that model can easily generate the result.

- NLP based on machine learning allows for more accurate sentiment analysis, which means it can detect whether the tweet is hatred or non-hatred.

- NLP based on machine learning allows a company to obtain a more complete understanding of their users.

Despite the fact that machine learning algorithms outperform deep learning in our experiment, we cannot overlook deep learning's performance and accuracy rate. Both machine learning and deep learning yielded high accuracy rates in our experiment.

CHAPTER NINE

Conclusion

9.1: Research Challenge:

The most important challenges in our research, we have not found a huge twitter dataset for sentiment analysis. In our dataset there is just 29530 tweets on that, and all the tweet is in English language, but when we discuss about research that time, we have decided that implement sentiment on some of the other languages.

After that when we have implemented the machine learning algorithms, we find that some of the libraries are older version and its takes lots of time to run the process.

While implement the Bi-LSTM deep learning model we faced lots of challenges, it takes lots of time for giving the result. Also showing error while we increase the epoch size. We faced lots of difficulties when we layering the dataset before passes it through the model.

Also, when we add more data to balance the dataset from imbalance, the build model becomes slower than before, so it takes about 20-25 minutes for the model to run. But that just limitation of hardware.

9.2: Future Scope:

In the above discussions we have showed the importance of analyzing sentiment as text, which can be useful for businesses and large corporations to know how they are doing to meet the demand of their customer. In this paper we have mainly focused on twitter data where people post their tweets and we analyzed those tweets if they contain any message that may cause the spread of hatred using both machine learning and deep learning. In our paper we illustrate a comparative analysis among various machine learning algorithms such as Naïve Bayes, Support Vector Machine and Logistic Regression and from deep learning model we have used RNN with Bi-LSTM. Finally, we have compared all these techniques in terms of their performance. However, even though we have successfully built our model to analyze text, there are still so many scopes where further work can be done to improve the working capability of our model. Some of the works that can be implemented in the future with our research is mentioned bellow.

- Our model only works with English sentiments. However, in the internet sentiments are available in various language. If we need to perform analysis over sentiment of different language we can implement machine translation with our model, though it will degrade the performance of our model.
- In our model we have only worked with supervised machine learning where the data set is already labeled. However, we may also need to work with unlabeled data most of the time as the size of labeled data is not that big, therefore, for designing a very big model we will need to work with
- Another feature that can be included in the future is deriving sentiment from audio and then convert is as a text, finally performing analysis over that text. As there are also audio feedback available on the internet.
- In the long term, we may be able to integrate machine learning with lexicon-based approaches for sentiment analysis, resulting in a hybrid approach. We can bypass several disadvantages and gain decent accuracy from the sentiment analysis model by using this strategy.

9.3: Conclusion:

After analyzing many papers on sentiment analysis, we can observe that various types of techniques and approaches are used to perform sentiment analysis. However, most of the work on this topic is done using machine learning and deep learning approaches. In this paper we have also tried to perform sentiment analysis using some popular machine learning and deep learning algorithms. For our research we have used Naïve Bayes, Support Vector Machine and Logistic Regression from machine learning and from deep learning model we have used RNN with Bi-LSTM. Booth of the techniques is implemented using the same train and testing dataset. The dataset we are using here contains 16 thousand tweets where all of them are labeled in one of two groups. Among the tweets there are some that contains hatred messages and some are neutral. As our dataset is labeled, we have used the supervised learning approach. After performing various classifications and feature extraction we found out that for our dataset Random Forest algorithm works the best for both the train and testing set (97.74% and 96.24% respectively) and the precession, recall and f1-score is also maximum for random forest. Though, the accuracy of other models is not so far behind. The accuracy of those models is just 1-2% shy of random forest technique. For better accuracy we wanted to implement RNN with Bi-LSTM from deep learning. However, in opposite of our expectation the deep learning model gave us lesser accuracy than random forest. Though the difference with machine learning models is insignificance. Our analysis is that deep learning works best with very big dataset where our dataset contains 16 thousand rows only which is considered a small size dataset. Therefore, the accuracy of Bi-LSTM model was lower than our machine learning models.

There are some limitations in our research. Deep learning approaches must have worked better in terms of accuracy, but using a small dataset Machine Learning model gives the best accuracy in our experiment, but overall, we got considerable accuracy from both machine learning and deep learning model.

References

1. M. R. Hasan, M. Maliha and M. Arifuzzaman, "Sentiment Analysis with NLP on Twitter Data," 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), 2019, pp. 1-4, doi: 10.1109/IC4ME247184.2019.9036670.

2. Bhowmik, Nitish Ranjan, et al. “Sentiment Analysis on Bangla Text Using Extended Lexicon Dictionary and Deep Learning Algorithms.” *Array*, Elsevier, 1 Jan. 2022, www.sciencedirect.com/science/article/pii/S259000562100059X.
3. T. Toma, S. Hassan and M. Arifuzzaman, "An Analysis of Supervised Machine Learning Algorithms for Spam Email Detection," 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI), 2021, pp. 1-5, doi: 10.1109/ACMI53878.2021.9528108.
4. L. Povoda, R. Burget and M. K. Dutta, "Sentiment analysis based on Support Vector Machine and Big Data," 2016 39th International Conference on Telecommunications and Signal Processing (TSP), 2016, pp. 543-545, doi: 10.1109/TSP.2016.7760939.
5. M. M. Nasr, E. M. Shaaban, and A. M. Hafez, "Building sentiment analysis model using graphlab," *International Journal of Scientific and Engineering Research,* vol. 8, p.1155–1160, 2017.
6. M. Moh, A. Gajjala, S. C. R. Gangireddy and T. -S. Moh, "On Multi-tier Sentiment Analysis Using Supervised Machine Learning," 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), 2015, pp. 341-344, doi: 10.1109/WI-IAT.2015.154.
7. A. Tripathy, A. Agrawal, and S. K. Rath, "Classification of sentimental reviews using machine learning techniques," *Procedia Computer Science,* vol. 57, p. 821–829, 2015.
8. S.-A. Bahrainian and A. Dengel, "Sentiment analysis using sentiment features," *in 2013*
   IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), vol. 3, p. 26–29, IEEE, 2013.
9. S. -A. Bahrainian and A. Dengel, "Sentiment Analysis Using Sentiment Features," 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2013, pp. 26-29, doi: 10.1109/WI-IAT.2013.145.
10. A. S. Rathor, A. Agarwal, and P. Dimri, "Comparative study of machine learning approaches for amazon reviews," *Procedia computer science,* vol. 132, p. 1552–1561, 201.
11. M. A. Badresiya and J. Teraiya, "Performance analysis of supervised techniques for review spam detection," *Int J Adv Netw Appl Special,* , no. 21-24, 2014.
12. N. Shrestha and F. Nasoz, "Deep learning sentiment analysis of amazon.com reviews and
    ratings," *arXiv preprint arXiv:1904.04096,* 2019
13. X. Fang and J. Zhan, "“Sentiment analysis using product review data," *Journal of Big Data,* Vols. 2, no. 1, p. 1–14, 2015.

14. Q. Huang, R. Chen, X. Zheng and Z. Dong, "Deep Sentiment Representation Based on CNN and LSTM," 2017 International Conference on Green Informatics (ICGI), 2017, pp. 30-33, doi: 10.1109/ICGI.2017.45.
15. A. J. Shamal, R. G. H. Pemathilake, S. P. Karunathilake and G. U. Ganegoda, "Sentiment Analysis using Token2Vec and LSTMs : User Review Analyzing Module," 2018 18th International Conference on Advances in ICT for Emerging Regions (ICTer), 2018, pp. 48-53, doi: 10.1109/ICTER.2018.8615600.
16. Naive Bayes Classifier in Machine Learning - Javatpoint", *www.javatpoint.com*, 2022. [Online]. Available: https://www.javatpoint.com/machine-learning-naive-bayes-classifier.
17. Support Vector Machine (SVM) Algorithm - Javatpoint", *www.javatpoint.com*, 2022. [Online]. Available: https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm
18. Logistic Regression in Machine Learning - Javatpoint", *www.javatpoint.com*, 2022. [Online]. Available: https://www.javatpoint.com/logistic-regression-in-machine-learning
19. Machine Learning Random Forest Algorithm - Javatpoint", *www.javatpoint.com*, 2022. [Online]. Available: https://www.javatpoint.com/machine-learning-random-forest-algorithm.
20. "Kaggle: Your Machine Learning and Data Science Community", *Kaggle.com*, 2022. [Online]. Available: http://kaggle.com/. [Accessed: 26- Apr- 2022].
21. Education, "What is Supervised Learning?", *Ibm.com*, 2022. [Online]. Available: https://www.ibm.com/cloud/learn/supervised-learning.
22. "Unsupervised Machine Learning: Algorithms, Types with Example", *Guru99*, 2022. [Online]. Available: https://www.guru99.com/unsupervised-machine-learning.html.
23. [11]"Reinforcement Learning Tutorial - Javatpoint", *www.javatpoint.com*, 2022. [Online]. Available: https://www.javatpoint.com/reinforcement-learning. [Accessed: 26- Apr- 2022].
24. "D01 Data Deliverable.docx - Deliverable 01 Research Paper Blake Wray and Luke Tuttle Definition of Deep Learning the Three Fathers of Deep Learning:: Course Hero." *D01 Data Deliverable.docx - Deliverable 01 Research Paper Blake Wray and Luke Tuttle Definition of Deep Learning The Three Fathers of Deep Learning: | Course Hero*, https://www.coursehero.com/file/133336981/D01-Data-Deliverabledocx/.
25. By: IBM Cloud Education. "What Are Neural Networks?" *IBM*, https://www.ibm.com/cloud/learn/neural-networks.
26. Marr, Bernard. "What Is Deep Learning Ai? A Simple Guide with 8 Practical Examples." *Forbes*, Forbes Magazine, 10 Dec. 2021, https://www.forbes.com/sites/bernardmarr/2018/10/01/what-is-deep-learning-ai-a-simple-guide-with-8-practical-examples/.
27. "Introduction to Recurrent Neural Network." *GeeksforGeeks*, 3 Oct. 2018, https://www.geeksforgeeks.org/introduction-to-recurrent-neural-network/.

28. Machine Learning Random Forest Algorithm - Javatpoint", *www.javatpoint.com*, 2022. [Online]. Available: https://www.javatpoint.com/machine-learning-random-forest-algorithm.
29. Bo, P., & Lillian, L., “Opinion mining and sentiment analysis”, Journal Foundations and Trends in Information Retrieval, Vol. 2, 2008.
30. Jeh, G., & Widom, J., “Mining the Space of Graph Properties”, Proceedings of the Tenth {ACM} {SIGKDD} International Conference on Knowledge Discovery and Data Mining, Seattle, Washington, USA, August, 2004.
31. Choudhari, and Rajankar, Introduction to Natural Language Processing With Python, International Journal of Modern Trends in Engineering and Research (IJMTER)Volume 02, Issue 04, [April – 2015].
32. Bespalov, D., Bai, B., Qi, Y., and Shokoufandeh, A., “Sentiment classification based on supervised latent n-gram analysis”, In Proceeding of the ACM conference on Information and knowledge management (CIKM- 2011), 2011.
33. Partha, C., and Amit, N., Software Innovations in Clinical Drug Development and Safety book, page 56.
34. Halleh, K., Comparing Feature Sets and Classifiers for Sentiment Analysis of Opinionated Free Text, MSc. Thesis, July 2012, page 13.
35. Bird, S., Klein, E., and Loper, E., Natural Language Processing with Python, 1st Edition, O’Reilly Media publisher, 2009.
36. B. Liu, "Sentiment analysis and opinion mining," Synthesis lectures on human language technologies, Vols. 5, no. 1, p. 1–167, 2012.
37. A. Tripathy, "Sentiment analysis using machine learning techniques," Ph.D. dissertation,2017.
38. Larsen, Peder Olesen, and Markus von Ins. “The Rate of Growth in Scientific Publication and the Decline in Coverage Provided by Science Citation Index.”,Scientometrics 84.3 (2010): 575–603. PMC. Web. 25 Sept. 2015.
39. Ian, H.W., Eibe, F., and Mark A. H., Data Mining: Practical machine learning tools and techniques. The Morgan Kaufmann series in data management systems, 3rd Edition, Morgan Kaufmann Publishers is an imprint of Elsevier, 2011.
40. S. Tammina and S. Annareddy, "Sentiment analysis on customer reviews usingconvolutional neural network," in 2020 International Conference on ComputerCommunication and Informatics (ICCCI), p. 1–6, IEEE, 2020.
41. Saifee, V., & Jay, T., “Applications and Challenges for Sentiment Analysis: A Survey”, International Journal of Engineering Research & Technology (IJERT), Vol. 2 Issue2, 2013.
42. https://www.kaggle.com/arkhoshghalb/twitter-sentiment-analysis-hatred-speech?fbclid=IwAR18OFwy52gI2vN756TPG98Ke23pjmJ8PueJFprJOb-JV9sdqwI7skUM6ds.