Sentiment Analysis Using Natural Language Processing (NLP) & Deep Learning

Prepared by

Kazi Minhazul Islam
17101456
Md. Safkat Reza
17101106
MD. Samin Yeaser
17301063

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
BRAC University
September 2021

Declaration

It is hereby declared that

1. 1. We are the only authors of this thesis and that neither any part of this thesis nor the whole of the thesis has been submitted for a degree to any other university or any Institution. This paper is fully our own work while completing degree at BRAC University.
2. 2. The thesis does not contain material previously published or written, except where this is appropriately cited through full and accurate referring.
3. 3. The content of the thesis is the result of work which has been carried out since the date of approval of the research program.
4. 4. All the ethics procedure and guidelines have been strictly followed while preparing thesis

Student's Full Name & Signature:

Minhazul

Kazi Minhazul Islam
17101456

Md. Samin Yeaser
17301063

Md. Safkat Reza
17101106

Approval

The thesis/project titled Sentiment Analysis Using Natural Language Processing (NLP) & Deep Learning submitted by

1. Md Samin Yeaser(17301063)
2. Kazi Minhazul Islam(17101456)
3. Md. Safkat Reza (17101106)

Of Summer, 2021 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science and Engineering on September , 2021.

Examining Committee:

Supervisor:
(Member)

Saiful

MD Saiful Islam
Senior Lecturer
Department of Computer Science and Engineering
BRAC University

Co-Supervisor:
(Member)

Rafeed Rahman

Rafeed Rahman
Lecturer
Department of Computer Science and Engineering
BRAC University

Head of Department:
(Chair)

Dr. Sadia Hamid Kazi
Chairperson and Associate Professor
Department of Computer Science and Engineering
BRAC University

Program Coordinator:
(Member)

Dr. Md. Golam Rabiul Alam
Associate Professor
Department of Computer Science and Engineering
BRAC University

Ethics Statement

Assigning an individual based on a heuristic provided by a Machine Learning model without any further analysis or evaluation might provoke ethical dilemmas. Hence, we assured the transparency of the evaluation process. In addition to that provide visual interpretation of the output provided by the Machine Learning Model.

Abstract

It is an age of the Web and electronic media, and social media stages are one of the foremost frequently used communication mediums these days. But a few individuals utilize these platforms for a noxious reason and among those negative angles "Cyberbullying" is predominant. The way of monitoring user opinions throughout social media platforms such as Twitter and Facebook have been proven to be an effective way of learning practically all of the consumers' thoughts which can open the door of potential future implementations. General emotion inspection can give us important data. The examination of supposition on informal communities, for example, Twitter or Facebook, has become an amazing method for finding out about the clients' sentiments and has a wide scope of utilizations. Notwithstanding, the productivity furthermore, the exactness of notion examination is being blocked by the difficulties experienced in characteristic language handling (NLP). As of late, it is established that profound learning models are potential answers to the drawbacks of NLP. Natural language processing refers to a process that enables the machine to act like human and decreases the space between the person and the machine. Thus, NLP readily communicates with the computer in a straightforward sense. NLP has gained several uses in recent times. Each one of them are extremely effective in daily life. An example can be a device which can be handled by voice commands. Several research workers are putting effort on this idea in order to make even more real-life applications Natural Language Processing has tremendous potential to facilitate the use of computer interfaces for humans, as people will ideally communicate in their own language to the computer instead of learning an exclusive language based on computer instructions. In case of programming, traditional programming language's importance has always been underrated. This concept is questionable. We believe that modern Natural Language Processing techniques can make possible the use of natural language to express programming ideas, thus drastically increasing the accessibility of programming to non-expert users. Our team thinks that the implementation of natural language to convey programming concepts may be made possible by contemporary natural language processing techniques so that programming is accessible to inexperienced consumers substantially. The following paper surveys the most recent analysis that have utilized profound methodology how to take care of conclusion investigation issues, for example, assessment extremity. Models utilizing term recurrence opposite record recurrence (TF-IDF) and content insertion was implemented to an arrangement of datasets. At last, one similar examination of the exploratory outcomes was carried out in respect to several models and information highlights.

Keywords: NLP; Deep Learning; Machine Learning; Cyber Bullying; Racism; Social Media; Prediction; Bi-directional LSTM

Acknowledgement

Firstly, all praise to the Almighty Allah who has enabled us to complete our thesis without facing any hassle.
Secondly, thanks to our supervisor Md. Saiful Islam sir and co-supervisor Rafeed Rahman sir for their continuous help and consultation. They provided us with profound assistance at any moment.
last but not the least, we are grateful to our parents for their endless prayers and blissful support throughout the journey.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Commencement

The regulation of emotions, views, and texts from a particular point of view or abstract information is referred to as sentiment analysis. [1]. Sentiment analysis provides understanding information about popular attitudes and opinions through analyzing numerous tweets, posts, comments, and criticisms. This is a proven method in order to predict a several major events, such as movie reviews, posts regarding a selective topic or any comment of any users, and general elections [2]. Detecting racial behavior can also be done through SA which is a contemporary situation right now. The next element of internet modernization, the shift from static web sites to constantly changed or customized content, as well as the rise of social media, are examples of this. It has sparked the development of blogs, forums, and online social media that allow consumers to spread and establish their thoughts on any issue. For example, they might grumble about an item that they have purchased, share negative perspectives, or express their political perspectives. Many applications (such as recommender systems), organizational survey studies, and political campaign planning all rely on similar knowledge about consumers. In addition, breaking down general suppositions is likewise essential to authorities since they clarify human action and conduct and the way others' perspective may change them. The conjecture of user perspective may be particularly beneficial while recommending and personalizing to compensate whenever there is a lacking regarding clear consumer input on a supplied service. Moreover, in the field of machine learning, different strategies, for example, the beneficiaries' can be utilized solving the issue [3]. Online social media, whose members create a constantly enhancing volume of data, is a source of data for sentiment analysis (SA). As a result, these information banks must be regarded as part of the big data strategy, because extra challenges needs be acknowledged so that effective data storage, access, and processing can be achieved and the accuracy of the findings' is ensured. [4] Web-based media or social media is a stage that empowers clients to impart and associate with their companions on the web and permits them to share their photographs, recordings, and day-by-day activities. Nowadays, pretty much every individual is found via Social media. According to statistics nearly 2 billion users used social sites in 2015 and the figure has now increased to 3.196 billion [5]

1.2 Importance/usefulness:

Bullying someone using digital means is considered as Cyberbullying. Cyberbullying can occur on various communication platforms such as social network, texting websites, video-gaming platforms etc. It is repeated behavior, aimed at scaring, angering or shaming those who are targeted. It's a recurring action towards individuals who are being targeted, insulted or shamed. This kind of behavior includes impersonating someone, leaking someone's private or embarrassing contents online and not to mention attacking someone based on his/her physical attributes. Sometimes real-life bullying and online bullying can occur at the very same time. However, cyber bullying gives away some kind of digital foot-mark information that can help eradicate the abuse and offer proof. [21] A research conducted by Sameer Hinduja and Justin W. Patchin shows that more than a third of the 5,000 kids polled in middle and secondary schools in the USA had faced cyberbullying. Above that, around 750 students bullied others by some mean. Physical attributes, activities, disability, religion and other aspects of children are sometimes targeted on to bully them. Since 2016, the overall number of young individuals with cyberbullying is up by 3%. Reports originated in the United Kingdom demonstrates the severity of the issue as well. Around 2000 people from 4140 teenage persons questioned by DitchTheLabel.org were bullied while surfing online. [22] Machine learning brings up several ways of preventing online bullying. [23] The advantage such algorithms have over parental control software and keyword-spotting blockers is that they should recognize subtle and sarcastic comments—a task that the former solutions can't cope with. The characteristic that separates these algorithms from parental-control and message filtering software is that these algorithms can detect detailed and humoristic insults. This concept was mostly ignored by previous related technologies. Other than that, they handle the fact whether the insults or slurs are intentional or somewhat accidental. [22] An article titled "Automatic detection of bullying in social media texts" some finders discussed a method of such kind. While on the early stage, it managed to detect insulting actions online which were English and Dutch based. [24] It also predicts who is playing which role in that particular case , thus enabling a human moderator of the system doing their work sooner and more efficiently. Our research will help detect bullying words sentence or phrases from different social media platform. It will help us to lower the rate of cybercrime by rising peoples' awareness.

1.3 Problem Statement

In this current world person close to each other. People all over the world are much more connected than before. And this thing happened only because smartphones are now really available. Also, social media made people closer to each other. People can give their opinions through these platforms. People are using various social media platforms frequently. People use these for various purposes. Some share their feelings, lifestyle or some use it to connect with the world or to make friends.
But the most alarming thing which has been a great problem for a century is racism. As people are now more involved in social media, these racial things are spreading more quickly and for this reason, all over the world different chaos is occurring.
It is a matter of regret that some people use these for some mean reason or to

abuse others and to spread hatred. Racism is one of the biggest problems that is frequently occurring on various social media platforms. This is one kind of crime but it is hard to detect. So we are trying to detect this kind of hate speech which is creating discrimination among people.
We will collect data from various social media platforms and run NLP algorithms on that data to find out which comments, texts, or posts are connected with racism.

1.4 Current scenario and Motivation:

In Bangladesh, almost 49% of students have reported being a victim of cyberbullying. [25] All over the world people are facing all the time cyber bullying. The ratio is bigger in the case of woman. So, we wanted to make a secure platform for everyone use. As Facebook is the biggest social media platform, we will take it as an example. Facebook has introduced letting people report bullying or harassment on behalf of friends who themselves are not willing to report out of intimidation or some other reason. [26] The good news is that modern advancements in big data have helped create sophisticated AI speech recognition systems that can identify nuances in written speech. These machine learning tools detect abusive and damaging online behavior and alert authorities. This not only helps ensure the safety of likely victims but also helps rehabilitate potential cyber bullies. But we wanted to use the machine learning to solve this problem. People don't need to custom report for a comment or for particular thing, the AI will detect the bullying words and it will delete by itself. [27]. Most platforms are now using algorithms powered by AI to detect and censor abusive content. This ensures potential attempts at cyberbullying can be thwarted before it causes serious harm. [27]. An ancient and common implementation of NLP is Email or message filters. Starting out with spam filters, it gradually uncovered certain words or phrases that indicate a spam message. As filtration is being upgraded day by day, a more modern and effective application of NLP can be located in email arrangement of Google mail. Depending on their bodies, the algorithm determines that e-mails fall into one of three categories which are primary, social or promotional. [27]. By using the same algorithm, we can detect racial behavior or cyberbullying in social media platform. We can detect words and omit them or even we can ban the user who is using racist words. Any post can be removed or any comment can be paused or permanently delete it.

1.5 Research Objectives

We are willing to make as system which is able to detect racial comments, posts, and texts from different social media platforms through our research. Most of the existing research in this field tends to focus on simple bullying words or sentences. This is not ideal and has very little to no applications in everyday use.
The most typical reactions of online adults in the United States when confronted with racist postings on social media are depicted in a comic way. In a poll conducted in August 2017, 40% of contributors said they would disregard the message, while 31% said they would cut connection with the content creator. (reference)
There are different types of racial words in these worlds. White supremacist uses different types of vulgar words to taunt the black people or the brown people. Also,

different religions people use other words or taunting phrases to taunt other religious people.
Different NLP programs or social media detections algorithms search for only bullying words but the sake of the simplicity of our research, here in our project, we will fetch data about a particular part which is racism, racist words, or racial phrases. This type of data is not readily available, and so we will need to collect data and process them. Initially, we will collect approximately 447,883 data words for different racial words, phrases. Then we will process to detect racial behavior through NLP algorithms.

1.6 Thesis Outline

The subsections of the paper are arranged in the following way. Section II contains a Literature review of the most recent research papers of the relevant topic and the background information of the algorithm that we will use. Section III contains the details about our data sets and the architecture of the proposed model.

Chapter 2

Literature Review

Ali et al. [8] extracted features and applied them to preprocess to include the original text in the feature set as well. Preprocessing was performed on all Twitter and Form spring datasets. The preprocessing included removal of special characters, single characters left after removing special characters, substituting multiple spaces with single spaces and stop words. The text was converted into lowercase as well. After that classification was performed.
Text files turned into numerical then classification algorithms were applied then SVM, naïve Bayes, Random Forest, and then an ensemble approach (hybrid model consisting of all the algorithms as mentioned) was used.
Form Spring dataset contains 13,110 posts labeled as bullying and non-bullying. This dataset was created. These researchers performed various tasks on the datasets like text classification, role labeling, sentiment analysis, as well as topic modeling. They applied machine learning classifiers, i.e. SVM, Naïve Bayes, and Logistic Regression to identify bullying traces from the dataset. According to them, SVM performed better than the rest with an accuracy of 81.6%.
The table "Dataset 2–Twitter" depicts the accuracy achieved after applying classification on the Twitter dataset consisting of 13,420 tweets labeled as "offensive" and "not offensive" 79% Training (SVM) and 78% testing (Logistic Regression) accuracy of their proposed model.
Ain at el. [9] collected data from 2270 user reviews and BOO, KTT, ELE reviews. To address the lack of labeled and big corpus in strategies that exist already, the proposed study creates a Treebank for Chinese emotions of social information. The Recursive Neural Deep Model (RNDM) was suggested for portending the labels at the phrase level, whether they are supportive or critical, and it outperformed SVM, Nave Bayes, and Maximum Entropy. Around 2300 film criticisms were gathered from the webpage and segmented using ICTCLAS which origins from China and considered effective for segmenting words. Five divisions were structured for each sentence and in order to purse them the Stanford parser was applied. By completing around 14000 sentences and 15000 words written in Chinese, the suggested strategy made the omen of sentence emotion labels much more efficient. With opposite prediction strategy, ME and NB outperform baselines by a large margin.
While working with BOO, KTT, and ELE reviews, a modified strategy combining a Probabilistic Neural Network (PNN) and a Restricted Boltzmann network (RBM) with two layers was proposed. This modified deep learning strategy was introduced to improve sentiment categorization efficiency. To tackle this sort of problem, the

polarity, i.e., negative and positive ratings, varies depending on the context. Neutral reviews are not taken into account. Experiments were conducted using the datasets of Pang and Lee and Blitzer, et al., binary classification applied on each dataset. Because the suggested solution does not rely on external resources such as a POS tagger or a sentiment dictionary, it is also quicker than a rival. As a prior research utilized a complicated approach for feature selection, dimensionality reduction was implemented to get a decreased number of features. Recursive Neural Deep Model (RNDM) using 2270 movie reviews from websites works better (90.8%) than baselines with a massive difference. PNN and RBM using BOO, KTT, ELE reviews attain accuracies in the following manner: MOV=92% BOO=92.7% , DVD=93.1% ELE=93.2% , KIT=94.9% .

Dang at el. [10] used datasets from Sentiment140 which was gathered from Stanford University. There are 1.6 million tweets regarding items and brands on it. They also collected data from Tweets Airline which is a dataset from Tweeter that contains consumer feedbacks regarding U.S. airlines. This dataset originated on February 2015 including around 15000 samples, which was classified into negative, neutral, and positive divisions. And they also used Cornell and IMDB Movie Reviews to collect their data. They used DNN, CNN, and RNN methodological approaches on their dataset.

The DNN model's accuracy is about 75% -80% The CNN model's accuracy is about 80% The RNN model's accuracy is about 50% .

Sintaha at el. [11] used a Twitter dataset consisting of 11,500 tweets. Then they applied machine learning algorithms and sentiment analysis techniques to build the model. They collected data from Twitter API and tracked the collected data using particular keywords such as dyke, bitch, stop Islam. After applying the preprocessing, they folded the 80% data for training and 20% data for testing. Then, they used an automated training set classifier to label the training tweets positive, negative or objective. After extracting the feature, the training tweets were processed via several machine learning classifying strategies to analysis and contrast the performance of that algorithm. Then the pre-processed test tweets were processed via the classifier to find the polarity.

They used Naïve Bayes Classifier, Support Vector Machines (SVM), and Convolutional Neural Network for the comparison of machine learning techniques. 89.39% accuracy (SVM). 73.0328% accuracy (Naïve Bayes). 48.6404% accuracy (Convolutional Neural Network Approach).

Chapter 3

Methodology

This section provides details about the workflow plan. Coming to the implemented approach that is shown in Fig 1. It depicts how racial behavior was detected in the acquired data sets.

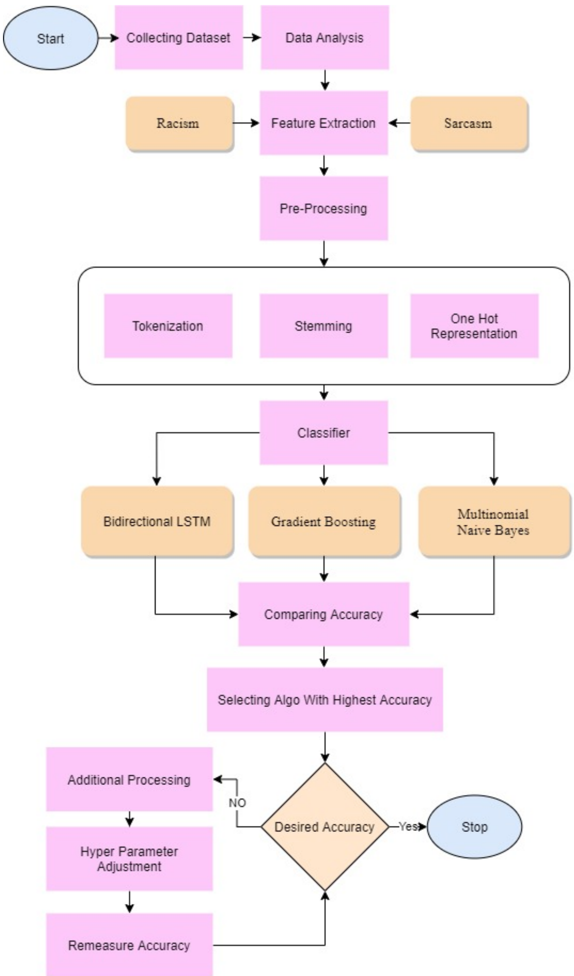


Figure 3.1: Implemented approach

Social network information is the biggest, most powerful dataset about human conduct. They provide researchers of social activities and trade specialists a universe of fresh scopes to get individuals, gatherings, and community. Sentiment analysis can be considered as the fundamental technique that machine learning is implemented in social networks. As an example, after a fresh commodity is introduced, consumers may post on Facebook about it or fill out an Amazon survey. Organizations may utilize machine learning to comprehend the overall population's response to their own or a rival's new item or plan.

Dataset	Quantity
aggression parsed data	115863
attack parsed data	115863
kaggle parsed data	8798
toxicity parsed data	159685
twitter parsed data	16852
twitter racism parsed data	13472
YouTube parsed data	3468

Table 3.1: Data-set and Quantity

In this dataset, there were 5 columns which are: Index, Text, ed_label_0, ed_label_1, and oh_label. We removed the unwanted columns from our datasets such as ed_label_0 and ed_label_1. After processing all the data, we achieved our desired dataset. Then we combined all of them into one dataset.

3.1 Data Pre-processing: Stemming:

Stemming is the procedure of decreasing bend in words (e.g., studies, studying) to originates (e.g., study). We have added a function for stemming our dataset, where we have taken an array named corpus. In the for loop, we implemented regular expression to our dataset which will run up to the length of the message. Then we made all the text to lower cases, following splitting all the text and finally storing it in a variable. After performing the stemming, we join all the splinted words back. Lastly, we append the whole review to the corpus array. One of the most important concepts for porter stemmer is the concept of m in the following formula: [15]

$$[C][VC]^m[V] \tag{3.1}$$

Porter‘s Stemmer Algorithm Consonant and Vowel Formula

3.2 Tokenization:

Tokenization is a technique of breaking a big portion text into smaller portions called tokens. Tokens can refer to word, character, or sub-word. As tokens are the basis of natural language, the usual way of processing the raw text occurs at the token level. Thus, Tokenization is done before applying machine learning models to text data. The tokens are used to construct a word repository. This repository indicates the set of unique tokens in the dataset. It can be prepared by taking each unique token in the dataset to account or by the top K frequently found words.

$$p(x) = \prod_{i=1}^{M} p(xi) \tag{3.2}$$

$$\forall i \, xi \, \epsilon \, V, \sum_{x \epsilon V} p(x) = 1 \tag{3.3}$$

x: Sentence; xi: subword sentence; V: forming vocubulary

3.3 One hot Representation:

A one-hot encoding is a category binary vector representation. This to start with, the category values must be translated to integer values. After that, every integer value is presented in a binary vector form with all zero values except the number's index being denoted with a 1.
One hot encoding permits the representation of categorical information towards being more obvious. Numerous machine learning calculations may not be able to perform with categorical information specifically. The categories require to be changed over into numerical form. It can frequently be needed for categorical input and output components. A number encoding can be used straightforwardly, reformed if needed. It can also solve issues regarding categories being in a normal fulfilling relationship and integer values being changed, like specifications for temperature 'cold', 'warm', and 'hot'. Issues might appear regarding no fulfilling relationship and permitting the representation to be biased on this kind of relationship may lead to harm towards learning to fathom this issue. The names 'dog' and 'cat' are two examples.
For such issues, we are willing to provide the network more direct control to memorize a probability-like number with respect to every probable label esteem. This might offer assistance in creating the issue simpler for the network to structure. While utilizing a one hot encoding in case of the output variable, it will hopefully provide with more improved collection of omens than only one name.

3.4 Manual One Hot Encoding:

We are going to expect that the whole collection of conceivable inputs is the total alphabet of lower-case characters and space. This will be utilized as a defense to illustrate we have one-hot encoding to demonstrate how to roll.
From char values to integer values, a mapping of all potential inputs is created. This map is used to encode the entry string at this point. The principal letter is encoded as 7 within the "h" input or as 7 within the set of input value (alphabet). The integer encoding is at that point changed over to one hot encoding. Ordinally one integer is encoded at once. A collection of 0 values is made the length of the alphabet in a way that representation of any anticipated character is assured. Following, the index of

the particular character is checked using a 1. Ready to check that a binary vector with the length 27 and the seventh index indicated 1. The main letter 'h' integers encoded as a 7. At last, the processing of the first letter is altered and the outcome is shown. It is done by finding the list of within the binary vector with the biggest value utilizing the NumPy argmax () method and after that utilizing the integer value in a traceback lookup table of character values to integers. The data using one-hot encoding is numerical, a machine learning model can effectively join such categorical include information by learning an isolated parameter, w, for each measurement. One of the issues with utilizing one-hot encoding in practice, be that as it may, is that the cardinality of factors can be large. Recommendation frameworks such as click-through expectation models can be constrained to bargain with million-, billion- or indeed trillion-dimensional feature spaces [16] In such settings, efficient processing and even storing of the data using one-hot encoding becomes a problem [17]

Also, the number of parameters to be learned become very large. If one additionally considers cross-product transformations - common in logistic regression models [18] - the problem is further exacerbated.

Presently that it has been showed how we roll our own one hot encoding from beginning, now we will see how the scikit-learn library can be utilized for conducting this mapping consequently cases regarding input sequence completely framing the expected range of input values.

There is also 2 other ways of One Hot representation which are:

1. Encoding with the help of scikit-learn:
2. Encoding using Keras

3.5 Embedding:

Word embedding refers to word representation for examination of text, usually in the shape of a real-valued vector that encodes the significance of the word so that the similar words in the vector space are expected to be closer in meaning. In different words, word embedding are word representation vectors that have not been supervised and have commonalities with semantical similarities. In the early 1950's, Zellig Harris, John Firth and Ludwig Wittgenstein introduced the theoretical representation of word embedding. [20]

In the section of Embedding, we took the sentence length 20, then we used a function called pad sequence where we have passed 3 parameters which are one hot, padding as 'pre', and maximum length as the sentence length. Then we printed the embedded doc.

3.6 Bi-directional LSTM:

LSTM was first proposed in [13]. Not at all like traditional RNNs, an LSTM network is enabled to improve itself monitoring previous cases, prepare and portend time series in case of exceptionally long-time slacks of obscure measure between significant occurrences. Other than that, LSTM too is able to solve the issue of vanishing angle and detonating slope that is unavoidable by RNN at the time of backpropagation optimization [14].

Long Short-Term Memory cell can be applied using these functions:

$$i_t = \sigma(W_{xi}X_t + W_{hi}h_{t-1} + b_i) \tag{3.4}$$
$$o_t = \sigma(W_{xo}X_t + W_{ho}h_{t-1} + b_o) \tag{3.5}$$
$$f_t = \sigma(W_{xf}X_t + W_{hf}h_{t-1} + b_f) \tag{3.6}$$
$$g_t = \tanh(W_{xc}X_t + W_{hc}h_{t-1} + b_c) \tag{3.7}$$
$$c_t = f_t \otimes c_{t-1} + i_t \otimes g_t \tag{3.8}$$
$$h_t = o_t \otimes \tanh(c_t) \tag{3.9}$$

Bidirectional LSTMs are the specialized version of traditional LSTMs. It is able to enhance model ability on sequence classification issues. In these issues, all of the time steps of the input sequence are accessible. In those cases Bidirectional LSTMs train two in place of one LSTMs on the input sequence. [12]

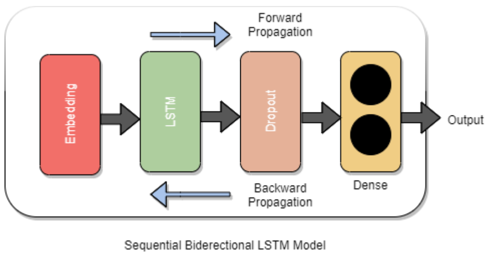


Figure 3.2: Bi-directional LSTM

We took the length of embedding-vector-features to 40. After then we pass the Sequential function to the model. Then adding Embedding function with the model passing voc-size, embedding-vector-features, input length=sent length as the parameters. Furthermore, we added the LSTM (100) to the model. We add the Dense function making the activation to the 'sigmoid'. Lastly we we compiled the model passing the parameters loss='binary-crossentropy', optimizer='adam', metrics= ['accuracy'].

3.7 Multinomial Naive Bayes:

Naive Byes is the most straightforward and fast classification algorithm and It is very efficient in the case of large data [29], [30]

Multinomial naive Bayes (MNB) considers that attributes (i.e., features) do not rely on each other given the context of the class, and it avoids all dependencies among attributes.

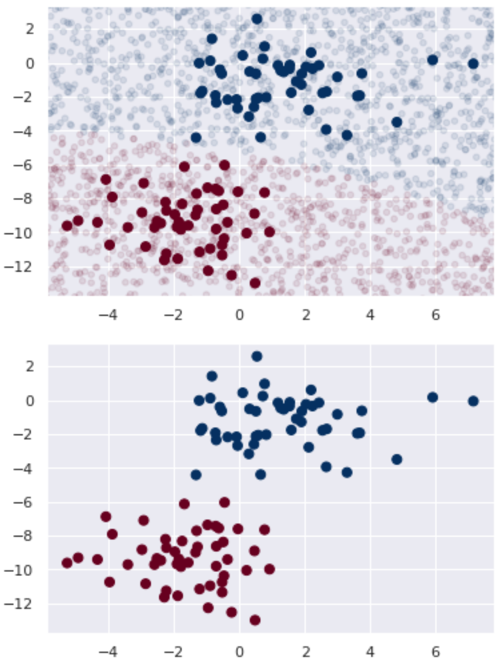
After importing Multinomial NB we pass it in a variable called a classifier. Then we fit the classifier passing the parameters of X_train and Y_train. For prediction, we used the X_test. After that, we found the accuracy score with Y_test.

As Naive Bayes works with probabilities of different data on classified labels. The probability of data by class is needed to be calculated. Bayes theorem is able to calculate posterior probability P(c $\mid x) from P(c), P(x), and P(x \mid c). The equation is :$

$$p(c \mid x) = \frac{p(x \mid c)p(c)}{p(x)} \tag{3.10}$$

$$p(c \mid X) = p(x \mid c) \times p(x_2 \mid c) \times ... \times p(x_n \mid c) \times p(x) \tag{3.11}$$

Here,
1. P(c | $x) indicates posterior probability of class (c, target) given predictor (x, attributes)$.
2. P(c) represents the previous probability of class.
3. P(x | $c) is the probability as in the probability of the predictor given class.$
P(x) is the previous probability of the predictor.
One greatly quick way to form a basic show is to accept that the data is described by a Gaussian conveyance with no co-variance between
measurements. This model can be fit by essentially finding the mean and standard deviation of the points inside each label, which is all you need to characterize such a distribution. The result of this naive Gaussian assumption is shown in the following figure: [19]



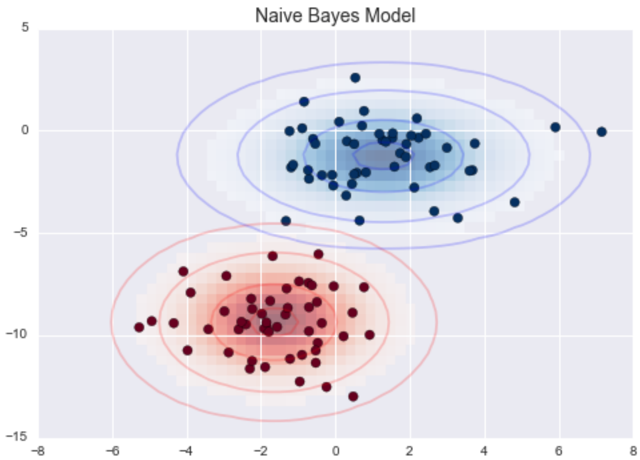


Figure 3.3: Multinomial Naive Bayes

Figure 3.3 shows our Naive Bayes model. Where we can find out the position of our train data set and test data set in the graph.

3.8 Gradient Boosting:

In the field of machine learning, Gradient Boosting algorithm has been playing an important role. Usually, errors are typically grouped into two types in machine learning procedures which are Bias Error and Variance Error. In order to decrease bias error of a model Gradient Boosting is mostly used. [31] Being the regressor, this algorithm can predict continuous target variable whereas being the classifier, it can also identify the categorical target variable. While being a regressor, Mean Square Error (MSE) is the cost function. On the other hand, while being a classifier Log loss is found as the cost function. [31] Gradient Boosting algorithm can be applied through many ways. Standard implementations in SciPy are one of them. Another application is efficient third-party libraries' implementation. Different interface and different identities are used in these implementations. [32]

$$gt(x) = Ey[\delta\psi(y, f(x))/\delta f(x) \mid x]f(x) = f^{(t-1)}(x) \tag{3.12}$$

$$(\rho, \theta t) = arg\,min\rho, \theta\Sigma i = 1N[-gt(xi) + \rho h(xi), \theta]2 \tag{3.13}$$

The training data and the test data that we have made for the Bi directional LSTM and Naive Bayes, these were used for Gradient boosting also. The result for the train data is: 87.7% accuracy The result for the test data is: 87.6%
One important aspect of gradient boosting is that it can provide predictive accuracy. There is a lot of freedom to optimize various loss functions and to offer many hyper parameter tuning which make the function highly versatile. In this case, no pre-processing data is necessary - typically works well with category and numerical values. Moreover, Handling missing data as in imputation is not needed. [34]

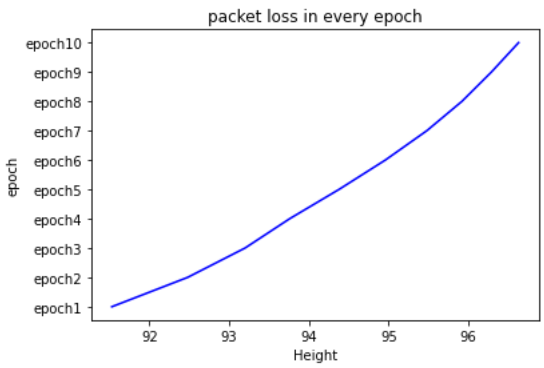


Figure 3.4: accuracy percentage for gradient boosting

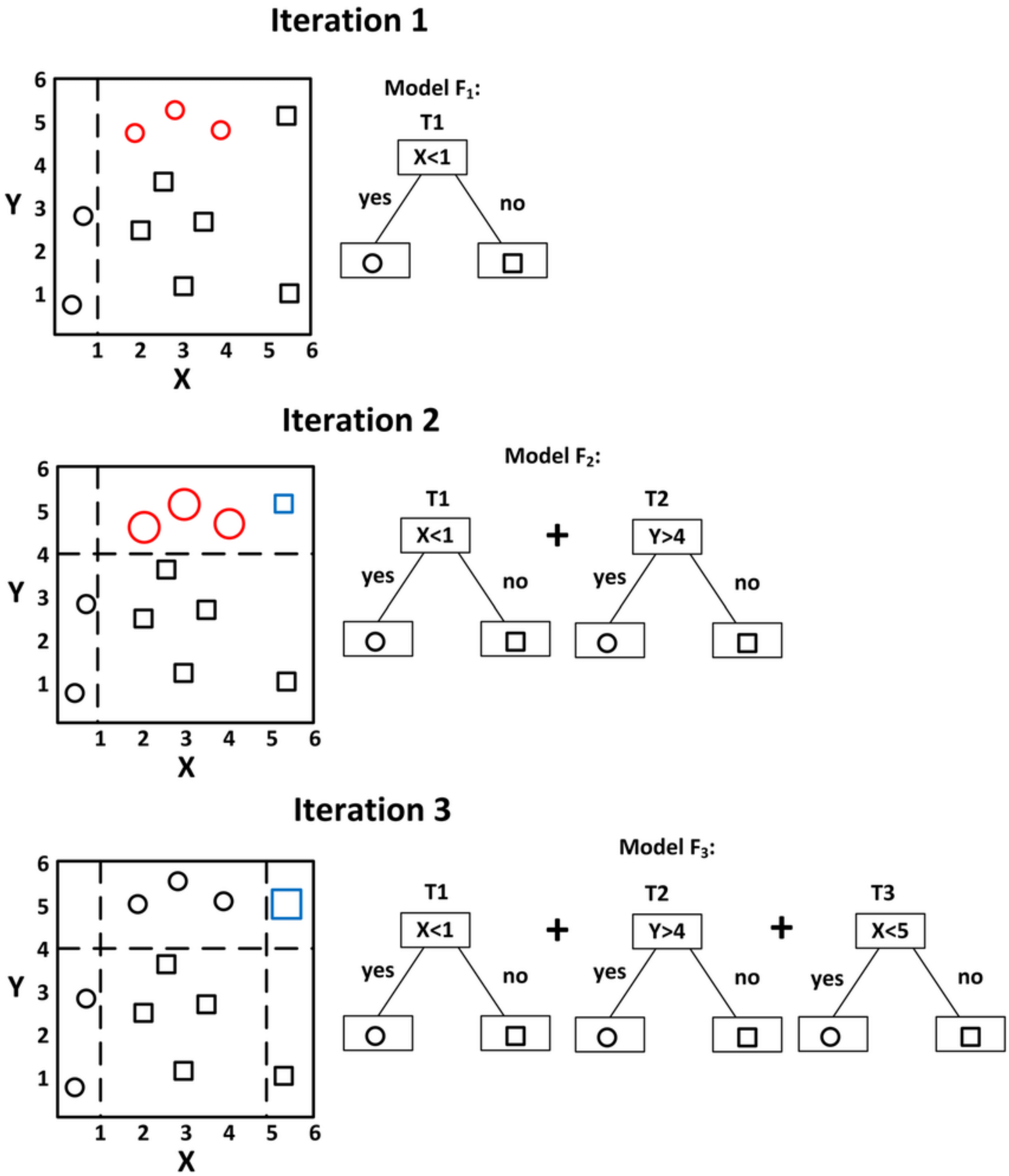


Figure 3.5: Gradient boosting graphs

The figure shows, in every epoch the accuracy rate is increasing. The lowest accuracy is in the epoch 1 and the highest accuracy is in the last epoch. By training our model more we can come to a best accuracy rate.

Chapter 4

Experimental Results and Analysis

4.1 Comparative Analysis of Supervised Models)

After fitting our data in the existing models described in the literature review,
- Bi Directional LSTM
- Naïve Bayes
- Gradient Boosting

We got the following accuracy-

Model	Accuracy
Bi Direction LSTM	92.44%
Naïve Bayes	61.40%
Gradient Boosting	87.6%

Table 4.1: Comparison of models

4.2 Comparison on Training Data:

We have used 90:10 spilt on our data. For that, we have achieved the following Loss and Accuracy.
In the Recurrent Neural Network, with the increase of every epoch the loss of the training data is decreasing. In the graph we see that when epoch is 1/10, the loss is more than 0.26 and when the epoch is 10, the loss is below 0.08. So, it is confirmed that with the increase of epoch, the loss of the training data decreases.
This figure 4.1 represent the data loss in very epoch in bidirectional LSTM model. The graph shows we loss the lowest level of data in our last epoch.

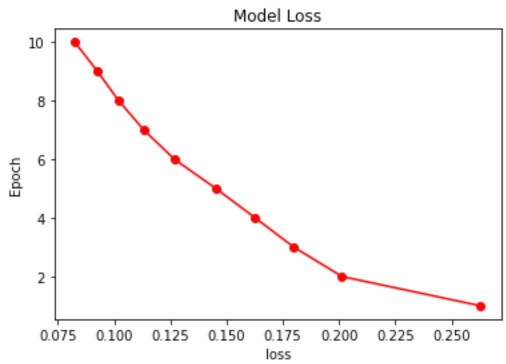


Figure 4.1: Data Loss

The figure 4.1 represent the data loss in very epoch in bidirectional LSTM model. It shows we loss the lowest level of data in the last epoch.

4.3 Graph:

The Validation accuracy of the Bi-directional LSTM which we have implemented is 92.44%. On the other hand, Multinomial Naive Bayes's accuracy is 61.40% which is not convenient.

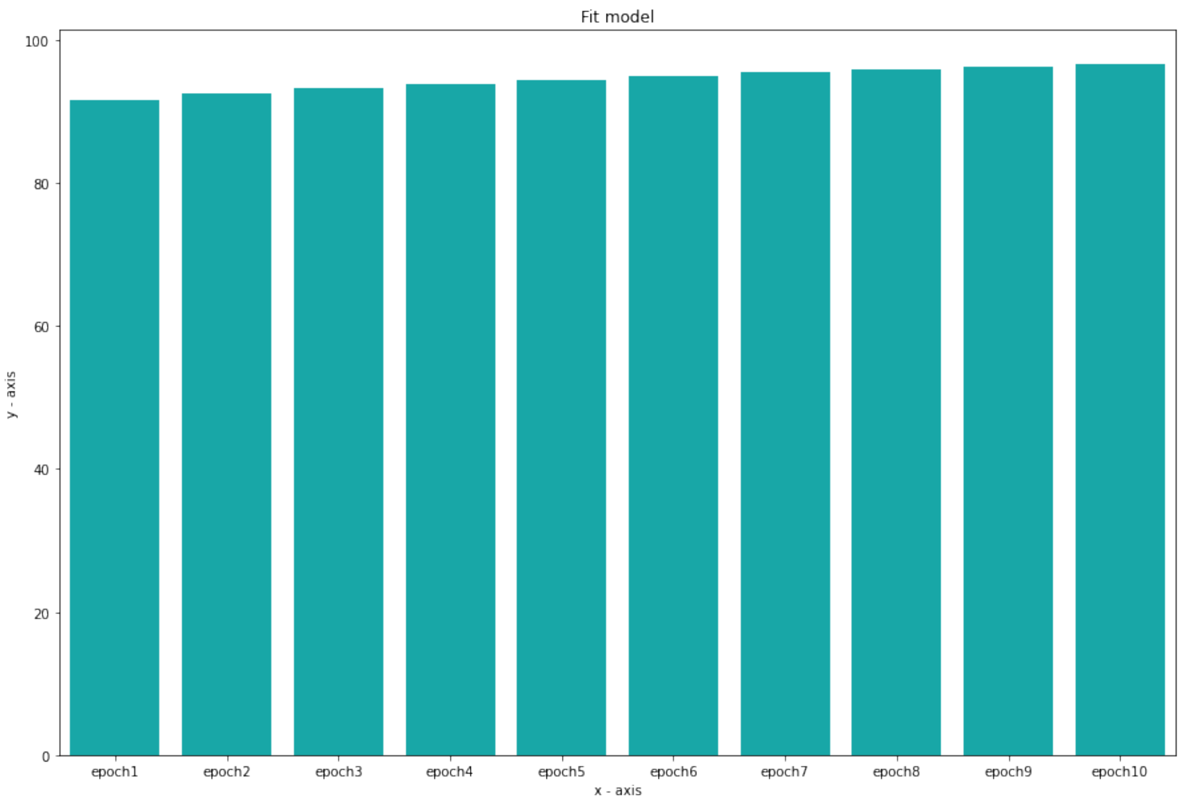


Figure 4.2: Accuracy Percentage

The figure 4.2 shows the accuracy rate in gradient boosting model in every epoch.

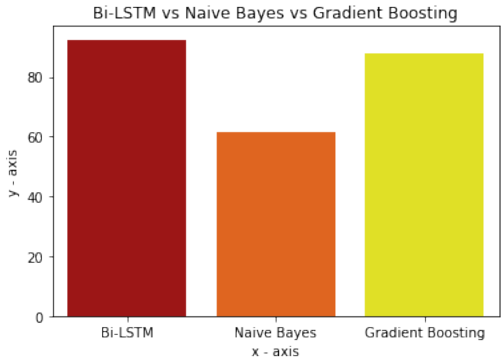


Figure 4.3: Bidirectional-LSTM vs Naive Bayes vs Gradient Boosting

Here in this figure we can determine which of our model gives us the highest percentage of accuracy rate. Bidirectional-LSTM gives the highest percentage of accuracy rate and on the other hand, Naive Bayes gives the minimum accuracy rate.

4.4 Challenge faced:

Initially, after a lot of brainstorming, we decided on a topic on which we wanted to conduct our thesis on. Since our topic was based on social media, that is why we relied on online data. We have collected about 300,000+ data from YouTube, twitter, Facebook. Also we performed an online survey. We asked people about their experience of going through racial behavior or facing any bullying words in social media. We then took an approach using the collected data set from Kaggle.

4.5 Conclusion:

Research regarding sentiment analysis is developing day by day. Our country is a developing country and if we achieve our goal of detecting racial behavior using sentiment analysis with deep learning, that will be a very big step in the world of science and technology. It will not only save time for various social media but also ensure a healthy and friendly platform of social media. We are only working with simple word detection so in the future we are looking forward to developing a system that can deal with simple and compound words. Sentences and phrases of different countries, races, and ethnicity people.

4.6 Future Work:

From our above analysis, we achieved a great extent of accuracy from our hybrid model and our future goal is to extend the research and use it for a wide range of malignancies. Furthermore, this type of research and analysis can be used for other social media platform. Most importantly, the objective is to afford people to have a healthy and friendly in the social media platform in the future.

Bibliography

[1] A. V. Yeole, P. V. Chavan, and M. C. Nikose, Opinion mining for emotions determination, ICIIECS 2015 - 2015 IEEE Int. Conf. Innov. Information, Embed. Commun. Syst., 2015.

[2] B. Heredia, T. M. Khoshgoftaar, J. Prusa, and M. Crawford, CrossDomain Sentiment Analysis: An Empirical Investigation, 2016 IEEE 17th Int. Conf. Inf. Reuse Integr., pp. 160165, 2016.

[3] Pouli, V.; Kafetzoglou, S.; Tsiropoulou, E.E.; Dimitriou, A.; Papavassiliou, S. Personalized multimedia content retrieval through relevance feedback techniques for enhanced user experience. In Proceedings of the 2015 13th International Conference on Telecommunications (ConTEL), Graz, Austria, 13–15 July 2015; pp. 1–8

[4] Thai, M.T.; Wu, W.; Xiong, H. Big Data in Complex and Social Networks; CRC Press: Boca Raton, FL, USA, 2016

[5] Chaffey Dave, "Global social media research summary 2019 — Smart Insights," 2019. [Online]. Available: https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/.[Accessed: 28-Sep-2019]

[6] 12 Best Social Media Datasets for Machine Learning — Lionbridge AI. (2021). Retrieved 8 January 2021, from https://lionbridge.ai/datasets/12-best-social-media-datasets/

[7] Kim, Y.; Sidney, J.; Buus, S.; Sette, A.; Nielsen, M.; Peters, B. Dataset size and composition impact the reliability of performance benchmarks for peptide-MHC binding predictions. BMC Bioinform. 2014, 15, 241. [CrossRef].

[8] Ali, A., Syed, A. (2020). Cyberbullying Detection Using Machine Learning. Pakistan Journal of Engineering and Technology, Pak JET.Volume: SI, Number: 01, Pages: 45- 50.

[9] Ain, Q., Ali, M., Riaz, A., Noureen, A., Kamran, M., Hayat, B., and Rehman, A. Sentiment Analysis Using Deep Learning Techniques: A Review. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 8, No. 6, 2017.

[10] Dang, N., Moreno-García, M., & De la Prieta, F. (2020). Sentiment Analysis Based on Deep Learning: A Comparative Study. Electronics, 9(3), 483. Doi: 10.3390/electronics9030483.

[11] Sintaha, M., Satter, S., Zawad, N., Swarnaker, C., Hassan, A. Cyberbullying detection using sentiment analysis in social media. http://dspace.bracu.ac.bd/xmlui/handle/10361/6420.

[12] Brownlee, J. (2017, June 16). How to Develop a Bidirectional LSTM For Sequence Classification in Python with Keras. Machine Learning Mastery.

https://machinelearningmastery.com/develop-bidirectional-lstm-sequence-classification-python-keras/: :text=Bidirectional%20LSTMs%20are%20an %20extension, LSTMs%20on%20the%20input%20sequence

[13] Sepp Hochreiter, Jurgen Schmidhuber. LONG SHORT-TERM MEMORY. In Neural Computation, 1997.

[14] Li, J., & amp; Shen, Y. (2017). Image describing based on bidirectional LSTM and improved sequence sampling. 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(. https://doi.org/10.1109/icbda.2017.8078733

[15] Bernardo, I. (2021, May 18). Stemming Text with NLTK - Towards Data Science. Medium. https://towardsdatascience.com/stemming-corpus-with-nltk-7a6a6d02d3e5

[16] Seger, C. (2018). An investigation of categorical variable encoding techniques in machine learning: binary versus one-hot and feature hashing (Dissertation). Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-237426

[17] A. Shrivastava. (2017) 2017 rice machine learning workshop: Hashing algorithms for large-scale machine learning. [Online]. Available: https://www.youtube.com/watch?v=tQ0OJXowLJA

[18] H. Cheng, L. Koc, J. Harmsen, T. Shaked, T. Chandra, H. Aradhye, G. Anderson, G. Corrado, W. Chai, M. Ispir, R. Anil, Z. Haque, L. Hong, V. Jain, X. Liu, and H. Shah, "Wide deep learning for recommender systems," CoRR, vol. abs/1606.07792, 2016. [Online]. Available: http://arxiv.org/abs/1606.07792

[19] Li, J., amp; Shen, Y. (2017). Image describing based on bidirectional LSTM and improved sequence sampling. 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA)(. https://doi.org/10.1109/icbda.2017.8078733

[20] Mandelbaum, A Shalev, A. (October 27, 2016). Word Embeddings and Their Use in Sentence Classification Tasks. Available: https://arxiv.org/pdf/1610.08229

[21] "Cyberbullying: What Is It and How to Stop It." UNICEF, www.unicef.org/end-violence/how-to-stop-cyberbullying.

[22] Dean Chester, et al. "How Ai Can Help Fight Cyberbullying." TechTalks, 6 Sept. 2019, bdtechtalks.com/2019/09/05/artificial-intelligence-online-bullying/.

[23] Ben Dickson, et al. "What Is Machine Learning?" TechTalks, 25 May 2019, bdtechtalks.com/2017/08/28/artificial-intelligence-machine-learning-deep-learning/.

[24] Van Hee, Cynthia, et al. "Automatic Detection of Cyberbullying in Social Media Text." PloS One, Public Library of Science, 8 Oct. 2018, www.ncbi.nlm.nih.gov/pmc/articles/PMC6175271/.

[25] Kobiruzzaman, Author M M. "Cyberbullying in Bangladesh Case Study- How to Prevent Cyberbullying." Educational Website For Online Free Learning, 26 May 2021, newsmoor.com/cyberbullying-in-bangladesh-case-study-how-to-prevent-cyberbullying/.

[26] staff, Science X. "Facebook Adds New Tools to Stem Online Bullying." Phys.org, Phys.org, 2 Oct. 2018, phys.org/news/2018-10-facebook-tools-stem-online-bullying.html.

[27] "Can Data Analytics and Machine Learning Prevent Cyberbullying?" Selerity, 11 June 2021, seleritysas.com/blog/2021/06/11/can-data-analytics-and-machine-learning-prevent-cyberbullying/.

[28] "8 Natural Language Processing (Nlp) Examples." Tableau, www.tableau.com/learn/articles/natural-language-processing-examples.

[29] M. Ontivero-Ortega, A. Lage-Castellanos, G. Valente, R. Goebel, and M.Valdes-Sosa, gaussian naıve bayes for searchlight classification analysis," Neuroimage, vol. 163, pp. 471-479, 2017

[30] H. Kamel, D. Abdulah, and J. M. Al-Tuwaijari, classification using gaussian naive bayes algorithm," in 2019 International Engineering Conference (IEC), IEEE, 2019, pp. 165-170.

[31] N. (2021, April 19). Gradient Boosting Algorithm — How Gradient Boosting Algorithm Works. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/04/how-the-gradient-boosting-algorithm-works/:

[32] Brownlee, J. (2021, April 26). Gradient Boosting with Scikit-Learn, XGBoost, LightGBM, and CatBoost. Machine Learning Mastery. https://machinelearningmastery.com/gradient-boosting-with-scikit-learn-xgboost-lightgbm-and-catboost/

[33] Natekin, A. (2013). Gradient boosting machines, a tutorial. Frontiers. https://www.frontiersin.org/articles/10.3389/fnbot.2013.00021/full

[34] Kurama, V. (2021, April 9). Gradient Boosting for Classification. Paperspace Blog. https://blog.paperspace.com/gradient-boosting-for-classification/