

Sentiment Analysis of COVID-19 Vaccination from Survey Responses in Bangladesh.

Anjir Ahmed Chowdhury (✉ anjir.ahmed@primeasia.edu.bd)

Primeasia University <https://orcid.org/0000-0002-6544-1005>

Argho Das

American International University Bangladesh

Suben Kumer Saha

American International University Bangladesh

Mahfujur Rahman

American International University Bangladesh

Khandaker Tabin Hasan

American International University Bangladesh

Research Article

Keywords: Covid-19, Corona, Vaccine, BERT, Deep Learning, Text classification, Sentiment Analysis, LSTM, Transformer, Deep Neural Network, Naive Bayes

Posted Date: May 7th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-482293/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Sentiment Analysis of COVID-19 Vaccination from Survey Responses in Bangladesh

Anjir Ahmed Chowdhury · Argho Das · Suben Kumer Saha · Mahfujur Rahman · Khandaker Tabin Hasan

Received: April 26, 2021/ Accepted: date

Abstract

Objectives: The COVID-19 pandemic is among the most serious global threats, and it is still a significant concern. The people of Bangladesh are undergoing one of the world's largest vaccination drive. With the recent launch and introduction of the COVID-19 vaccines, many of us are curious about the general opinion or view of the vaccine. While the vaccine has ignited new hope in the battle against COVID-19, it has also sparked militant anti-vaccine campaigns, so the need to analyze public opinion on the COVID-19 vaccine has emerged.

Methods: Traditional machine learning methods were used to obtain a benchmark result for the experiment. The recurrent neural network (RNN) algorithm was used next. Several different types of recurrent neural networks were used, including simple RNNs, Gated Recurrent Units (GRUs), and LSTMs. Finally, to achieve

a more optimal result, small BERT models (Bidirectional Encoder Representations from Transformers) were used.

Results: Upon study and testing on several models and methods, it can be seen that BERT model was the most accurate of the bunch, which was 84%. On the other hand, Naive Bayes was able to obtain an accuracy of 81%. Naive Bayes and BERT produced similar results in F1- Score, but the performance of Naive Bayes can improve as the dataset size grows.

Conclusion: Knowing about public opinions on the COVID-19 vaccine is critical, and action must be taken to ensure that everybody understands the value of vaccination and that everybody receives the COVID-19 vaccine. Vaccination may help to develop immunity, which lowers the likelihood of contracting the disease and its consequences.

Keywords Covid-19, Corona, Vaccine, BERT, Deep Learning, Text Classification, Sentiment Analysis, LSTM, Transformer, Deep Neural Network, Naive Bayes

Anjir Ahmed Chowdhury (✉)
Primeasia University
E-mail: anjir.ahmed@primeasia.edu.bd

Argho Das
American International University-Bangladesh
E-mail: arghodas555@gmail.com

Suben Kumer Saha
American International University-Bangladesh
E-mail: subenksaha@gmail.com

Mahfujur Rahman
American International University-Bangladesh
E-mail: mahfuj@aiub.edu

Khandaker Tabin Hasan
American International University-Bangladesh
E-mail: tabin@aiub.edu

1 Introduction

COVID-19 is a novel human pathogen that virologists suspect originated from bats and ultimately jumped through an intermediate host to humans [1]. The COVID-19 pandemic has been one of the most serious global threats and is still a continuing threat to a large degree. The clinical manifestations vary from mild to no symptoms to more serious diseases, which can lead to lung failure and even death [2]. In equal measure, Bangladeshi people are in the midst of the biggest vaccination campaign in human history. Many of us are

1 curious about the general opinion or view of the vac-
2 cine with the recent launch and implementation of the
3 COVID-19 vaccines. Although the vaccine has renewed
4 optimism in the fight against COVID-19, radical anti-
5 vaccine campaigns have also been sparked. This appears
6 in both news outlets and social media to be a widely
7 discussed and questionable topic. Thus, evaluating the
8 public's view of the COVID-19 vaccine with sentiment
9 analysis would be interesting. In recent decades, deep
10 learning algorithms have been very popular in differ-
11 ent fields of study, such as natural language process-
12 ing (NLP), image classification, and computer vision
13 & pattern recognition (CVPR). The justification be-
14 hind being preferred to these models is deep learning
15 models perform well when there is a large amount of
16 data and automated feature extraction is needed. Deep
17 learning models mainly include automatic feature ex-
18 traction by training complex features with minimal ex-
19 ternal efforts to ensure constructive representation of
20 the data through deep neural networks. Besides, deep
21 learning methods such as deep neural networks, long-
22 term memory networks (LSTMs) [3], and a transformer
23 based model called Bidirectional Encoder Representa-
24 tions from Transformers (BERT) [4] are used in classi-
25 fication tasks in many aspects.

26 In this study, public sentiment analysis on the
27 COVID-19 vaccine situation is conducted using deep
28 neural networks, long-term memory networks (LSTMs),
29 and Bidirectional Encoder Representations from Trans-
30 formers (BERT). For this purpose, all the data is ac-
31 quired through a survey of Bangladeshi citizens. The
32 methods mentioned earlier will be compared to find
33 which model gives the best sentiment analysis result for
34 the collected survey data. After collecting the dataset,
35 all the outliers were filtered and then removed. The user
36 comments are classified as "support taking vaccine" or
37 "oppose taking vaccine" using deep neural networks,
38 LSTM network, and BERT. This will help resolve vac-
39 cine skeptics' concerns and build the requisite public
40 faith in immunization to recognize the aim of herd im-
41 munity. Sentiment Analysis refers to the use of text
42 analysis and the development of natural language to
43 recognize and extract textual meaning from subjective
44 knowledge. Facts and opinions on electronic media are
45 available in two forms which are user-generated con-
46 tent and consumer-generated content. It is important
47 to discover, evaluate, and consolidate their views for
48 improved decision-making due to the immense expan-
49 sion of user perceptions, reviews, feedback, and recom-
50 mendations accessible via web tools. This means factual
51 claims about certain subjects by evidence, while views
52 are user-specific statements that reflect positive or neg-
53 ative feelings. Typically, since facts can be classified us-

ing keywords, it is harder to categorize views. To mine
statements from a paper, different text analysis and
machine learning techniques have been used [5]. Much
research has been done on sentiment analysis using ma-
chine learning. A paper on sentiment analysis to classify
twitter data using machine learning was done, and the
model gave an accuracy of 89.47 % [6]. A study about
sentiment analysis on Uri attack has been conducted
to mine emotions and polarity on Twitter. Approxi-
mately 5000 tweets were recorded and pre-processed to
create the dataset. For mining emotional responses and
polarity, R has been used. Experimental results showed
that the Uri attack disgusted 94.3 percent of individuals
[7]. Research about sentiment analysis using machine
learning for business intelligence has also been con-
ducted in the past [8]. Another article on analytically
categorizing and evaluating the prevalent testing tech-
niques and deployment of Sentiment Analysis Machine
Learning techniques on different applications has been
done [9]. Machine learning techniques like Naïve Bayes
and OneR have been used for sentiment analysis [10].
Another similar paper also uses six different machine
learning algorithms, including Naïve-Bayes (NB), Sup-
port Vector Machine (SVM), Logistic Regression (LR),
Decision Tree (DT), K-Nearest Neighbor (KNN), and
Random Forest (RF) for sentiment analysis [11]. The
forthcoming distribution of COVID-19 vaccines implies
an urgent need to track and try to better understand
public opinion on an ongoing basis to establish baseline
levels of vaccine confidence and to detect early warn-
ings of confidence loss [12]. Research on public attitudes
on Facebook and Twitter towards COVID-19 vaccina-
tions using Artificial intelligence-enabled analysis in the
UK and US [13]. A study on Social Network Analy-
sis of COVID-19 Sentiments has been conducted using
machine learning methods with twitter data [14]. This
study states the creation of an analysis of sentiment,
obtaining a large number of tweets. This development
uses prototyping. Results categorize consumers' view-
point into positive and negative through tweets, de-
picted in a pie chart and application module [15]. In this
journal, sentiment analysis in Fighting COVID-19 and
Infectious Diseases can be found [16]. In another study,
public sentiment insights on COVID-19 using machine
learning for tweet classification are used. Here machine
learning (ML) classification methods like Naïve Bayes
and logistic regression are used [17]. The remainder of
the analysis is structured as follows: the Methodology
of the studies is given in Section 2. The dataset used
in the experiment is explained in section 3. After that,
Result & Analysis and Discussion will be discussed re-
spectively in sections 4 and 5. Threats to Validity will
be in section 6 and lastly Conclusion in section 7.

2 Methods

To get a benchmark result, traditional methods such as Naive Bayes, Support Vector Machine, K-nearest Neighbor, Random Forest, Gradient Boosting, Ada Boost, and Decision Tree were used. Before using deep learning methods, deep neural network was used. A deep neural network is a system of algorithms that attempts to identify underlying associations in a set of data using a method that mimics how the human brain works. Deep Learning is concerned with converting and extracting features that attempt to create a correlation between stimuli and associated neural behaviors found in the brain. At the same time, Deep Neural Networks utilize neurons to convey data in the form of input data and output data through connections. In this research recurrent neural network algorithm (RNN) was used. There are several different recurrent neural networks, such as simple RNNs, Gated Recurrent Units (GRUs), and LSTMs. Each network has its own collection of benefits and drawbacks. Simple RNN, for instance, has no gates, while GRU uses gates to determine whether or not to move the previous input towards the next cell. GRU has a memory unit as well. On the other hand, LSTM has two additional gates known as the "Forget gate" and "Output gate," which make LSTM more efficient. The LSTM also has a feedback relation, which helps it process both single and sequential data points. Recent advances in natural language processing have shown that transfer learning can help achieve state-of-the-art outcomes for new tasks by tuning pre-trained models rather than starting from scratch. Transformers have made considerable progress in producing new state-of-the-art results for various NLP tasks, including text classification, text generation, and sequence labeling, among others. For a small dataset, bidirectional LSTM models can achieve significantly better results than BERT models. These simple models can be trained in much less time than their pre-trained counterparts [18]. Because a model's performance is dependent on the task and the data, these factors should be considered before selecting a model, rather than simply selecting the most popular model. The main issues with LSTM are that words are passed in and generated sequentially; the neural network can learn in many timesteps. Even bi-directional LSTMs aren't particularly good at capturing the true meaning of words. Because they're learning the left to right and right to left contexts independently and then concatenating them, the actual context is lost. These issues can be solved in BERT by pretraining the model to understand language and context and fine-tuning BERT to learn how to solve a specific task [19]. BERT (Bidirectional En-

coder Representations from Transformers) is based on transformers, a deep learning model in which each output element is connected to each input element. The weightings between them are automatically calculated based on their relationship. Google unveiled transformers for the first time in 2017. To manage NLP tasks, language models mainly used recurrent neural networks (RNN) and convolutional neural networks (CNN) at the time of their introduction [20]. BERT architecture was a game-changer in natural language processing, allowing for the use of transfer learning in a variety of tasks. BERT is a machine learning framework for natural language processing that is open source (NLP). BERT is a program that uses adjacent text to assist machines in interpreting the message of ambiguous language in the text.

2.1 Traditional Machine Learning Methods

Traditional machine learning methods such as Naive Bayes, Support Vector Machine, K-nearest Neighbor, Random Forest, Gradient Boosting, Ada Boost, and Decision Tree were used to get a benchmark result for our model. First, Naive Bayes was used. The Bayesian Theorem is used to construct a set of classification algorithms known as Naive Bayes classifiers. It is not a single algorithm but rather a group of algorithms that all follow the same principle: each pair of features to be classified is distinct from the others. Support Vector Machine (SVM) is a Supervised Machine Learning Algorithm for classification and regression that is relatively straightforward. It is most widely used for classification, but it can also be useful for regression. The K-Nearest Neighbors (KNN) algorithm is a supervised machine learning algorithm used to resolve regression and classification tasks. The KNN algorithm assumes that identical items are close together. To put it another way, related things are close together. A Random Forest is an ensemble technique that uses several decision trees and a system called Bootstrap and Aggregation, also known as bagging, to perform both classification and regression problems. Instead of depending on individual decision trees, the basic concept is to combine several decision trees to decide the final production. Ada Boost was the first genuinely efficient boosting algorithm created specifically for binary classification. Adaptive Boosting, or Ada Boost, is a common boosting technique that combines several "weak classifiers" into a singular "solid classifier." A standard boosting algorithm is gradient boosting. Each predictor in gradient boosting corrects the error of its predecessor. Unlike Ada boost, the training instance weights are not adjusted; instead, each predictor is conditioned using

the predecessor's residual errors as labels. The most effective and widely used method for classification and prediction is the decision tree. A decision tree is a tree like structure in which each internal node represents an attribute test, every branch portrays a test result, and each leaf node represents a class mark.

2.2 Deep Neural Network (DNN)

A deep neural network's basic concept is to simulate a large number of densely interconnected brain cells within a computer to learn, identify patterns, and make choices in a humanlike manner [21]. A traditional deep neural network consists of a series of layers, each of which corresponds to the layers on the other side, ranging from a few dozen to hundreds, thousands, or even millions of artificial neurons known as units. Some of them, referred to as input units, are intended to receive different types of data from the outside world, which the network will try to learn about, identify, or otherwise process. Other units, known as output units, are located in the network's contrary direction and signal how the network reacts to the information it receives [22].

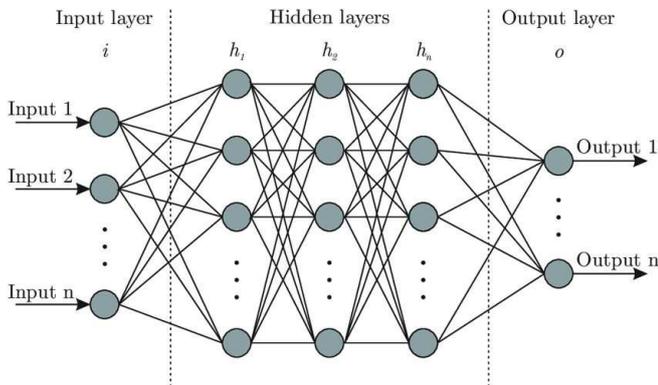


Fig. 1: Architecture of Deep Neural Networks (DNN) [23].

The majority of deep neural networks are fully connected, which means that each concealed unit and output unit is linked to every other unit in the layers on either side. A number named a weight, which can be either positive or negative represents the relationships between one unit and another. The greater the weight, the greater the sway of one unit over another. There are two ways for information to flow into a deep neural network. Patterns of data are transmitted to the network through the input units when it is learning or being trained or running normally, and these activate

the sections of hidden units, which then reach the output units. Over the last few decades, the deep neural network has been widely used in a wide range of applications, and it is likely to be the perfect solution [24]. One of the newest fields of deep neural network is stock price prediction. In the stock market, deep neural networks can predict prices [25].

2.3 Recurrent Neural Network (RNN) Variant

RNNs are excellent at processing sequence data and making predictions, but they have a short-term memory problem. LSTMs and GRUs were developed to solve short-term memory by using gate mechanisms. Gates are essentially neural networks that control the flow of data through the sequence chain. Speech recognition, speech synthesis, and natural language comprehension are examples of the state of the art deep learning applications that use LSTMs and GRUs.

2.3.1 Simple Recurrent Neural Network

Recurrent Neural Networks (RNN) are neural networks that operate with data in sequential order. Text, audio, and video can all be used to represent sequential data. RNN generates the current output by using the previous information in the series. The past is remembered by the recurrent neural network, and its decisions are guided by what it has learned in the past. Although RNNs learn similarly during training, they often remember what they've learned from previous inputs when producing outputs. It's a part of the system.

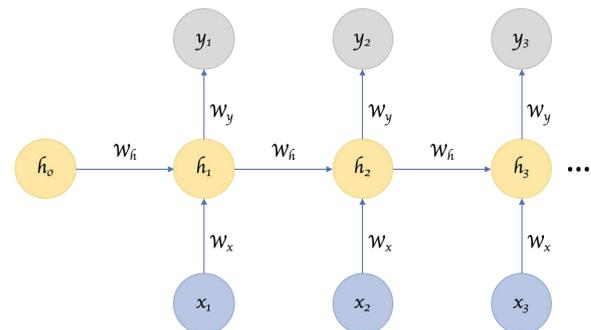


Fig. 2: Architecture of Recurrent Neural Networks (RNN) [26].

RNNs can take one or more input vectors and generate one or more output vectors, with the outputs influenced by weights applied to the inputs, as with a

standard NN and a "hidden" state vector representing the context based on previous inputs or outputs. As a result, depending on the earlier inputs in the sequence, the same input may generate different output. RNN keeps track of everything. All of the inputs in other neural networks are independent of one another. In an RNN, however, all of the inputs are connected. During preparation, the RNN remembers both of these relationships.

2.3.2 Gated Recurrent Unit (GRU)

The GRU workflow is identical to that of RNN, with the exception of the operations conducted inside the GRU unit. Let's take a look at the structure. GRU (Gated Recurrent Unit) is a recurrent neural network that aims to solve the vanishing gradient problem. Since both are constructed similarly and, in some cases, produce equally excellent results, GRU can be considered a variant of the LSTM.

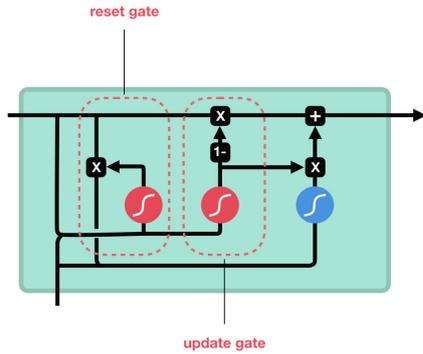


Fig. 3: Architecture of Gated Recurrent Unit (GRU) [27].

In equation (1) x_t is plugged into the network machine, its own weight is multiplied by it W_z . The same can be said for $h_{(t-1)}$, which stores information from previous $t - 1$ units and is multiplied by its weight U_z . Both results are added together, and the outcome is squashed between 0 and 1 using a sigmoid activation function.

$$z_t = \sigma(W^{(z)} * x_t + U^{(z)} * h_{t-1}) \quad (1)$$

This equation (2) is the same formula for the update doors. The difference is in the weights and how the gate is used. The update gate functions similarly to an LSTM forget and input gates. It determines what information should be discarded and what should be added. Another gate used to determine how much past

knowledge to forget is the reset gate. Since GRUs have fewer tensor operations, they are a little faster to train than LSTMs.

$$r_t = \sigma(W^{(r)} * x_t + U^{(r)} * h_{t-1}) \quad (2)$$

2.3.3 Long short-term memory (LSTM)

The control flow of an LSTM is analogous to that of a recurrent neural network (RNN). It processes data and passes the information on as it moves forward. The activities within the cells of the LSTM vary. The LSTM uses these operations to remember or forget information. The input gate, forget gate, control gate, and output gate are the four gates that make up the LSTM architecture [28].

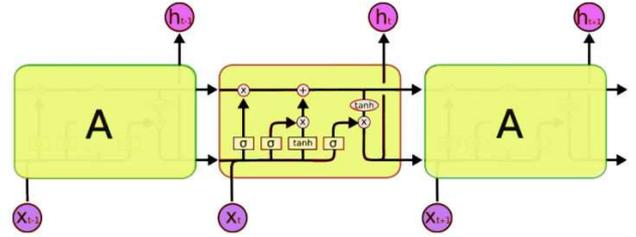


Fig. 4: Architecture of Long Short-Term Memory (LSTM) [29].

The gates of the LSTM are described by a set of equations. Before explaining the equation, one must first understand some of the variables used in these formulas. The activation function is regarded as the sigmoid, and also the weight matrices are W . The output of the prior LSTM block is represented by h_{t-1} , and the preference for the corresponding gates is represented by b_i . Finally, x_t is the existing timestamp's input. Now, The Input gate i_t is described as in equation (3),

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i) \quad (3)$$

This equation selected data that can be passed to the cell. The forget gate f_t determines which data from the input side of the previous memory should be ignored using the equation (4),

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f) \quad (4)$$

The following formula (5), where tanh is used to normalize the values into the range -1 to 1, and C is the candidate for cell state at the timestamp, controls the updating of the cell (t).

$$\tilde{C} = \tanh(W_c * [h_{t-1}, x_t] + b_c)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5)$$

The output layer (o_t) upgrades both the hidden layer h_{t-1} as well as the output layer according to the formula (6),

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t) \quad (6)$$

2.4 Bidirectional Encoder Representations from Transformers (BERT)

Transformer, an attention mechanism that learns contextual relationships among words in a text, is used by BERT. Only the encoder mechanism is required since BERT's goal is to produce a language model. The Transformer encoder reads the entire sequence of words simultaneously, unlike directional models that read the text input sequentially. As a result, it is classified as bidirectional, though it is more accurate to describe it as non-directional [30]. This property enables the model to deduce the context of a word from its surroundings [31]. The input consists of a series of tokens embedded into vectors before being processed by the neural network. The result is an H-dimensional chain of vectors, each of which refers to an input token with the same index. The task of determining a prediction goal when training language models is difficult. In order to overcome this challenge, the Masked Language Modeling (MLM) training strategy is used. In MLM, 15% of the words in each word sequence is replaced with a MASK token before being fed into BERT. Based on the context provided by the other non-masked words in the sequence, the model then attempts to predict the original value of the masked words [32]. For the prediction of the output words, there are some requirements. On top of the encoder output, a classification layer is added. Transforming the output vectors into the vocabulary dimension by multiplying them by the embedding matrix. Softmax is used to calculate the likelihood of each word in the vocabulary [33].

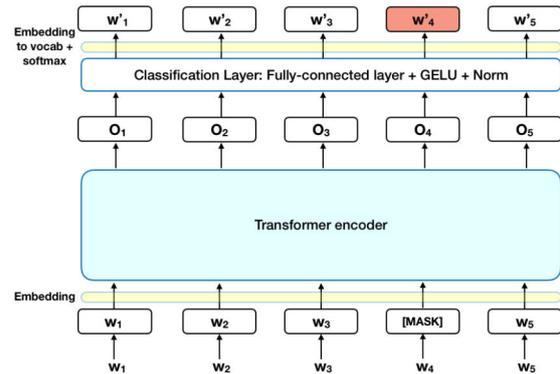


Fig. 5: Architecture of Bidirectional Encoder Representations from Transformers (BERT) [31].

Only the prediction of masked values is taken into account by the BERT loss function, which ignores the prediction of non-masked words. As a result, the model converges more slowly than directional models, but this is counter-balanced by its increased context-awareness. BERT is unquestionably a watershed moment in the application of machine learning to natural language processing. Because it's user-friendly and allows for quick fine-tuning, it'll likely have a wide range of practical applications in the future. For this research, pre-trained small BERT models are used because it is a simple yet effective method to maximize the utilization of all available resources [34].

2.5 Hardware and Software Setup

Naive Bayes, Support Vector Machine, K-Nearest Neighbor, Random Forest, Gradient Boosting, Ada Boost, and Decision Tree were not implemented from scratch. All the traditional machine learning methods mentioned above were implemented using sklearn. BERT, Neural Net, and LSTM, were implemented using TensorFlow 2 as a framework and Python 3.7 as a programming language. For API, Keras was used. Jupyter notebook, pip3, NumPy and Matplotlib were built as part of the supporting kit. These python packages all worked on a Linux operating system. The Intel Core i5 4th generation processor with a clock speed of 3.20 GHz was used in the hardware. It has 16 GB of DDR3 RAM with a bus speed of 1600 Mhz. The NVIDIA GTX 1060 6 GB graphics processing unit was also included.

3 Dataset

The method of researching with surveys that researchers submit to survey respondents is known as survey research. Survey results are then statistically evaluated to draw concrete research findings. The most important and fundamental justification for using surveys to perform research is to obtain answers to basic, important questions. Depending on the target audience and the purpose of the survey, the researcher will ask these questions in various formats. This is why the survey was done when conducting this research. On social media, people frequently share their likes and dislikes. These aren't all centered on the same thing. This distinguishes Twitter as a unique place to design a generic classifier, compared to domain-specific classifiers that could be created using datasets like film reviews. These features distinguish Twitter from other forms of media and make it difficult to manage in relation to other well-known text classification domains. At about the same time, it can also be said that it has a wide range of applications because it is one of the most important communication channels for gathering public opinion. The current project is designed to meet this need. The rapid exchange of user opinions on Twitter has allowed researchers to classify feelings about almost all, including feelings about goods [35], films [36], politics [37], digital technologies [38], and natural disasters [39]. During the COVID-19 pandemic, Twitter sentiment analysis was also used to measure the emotions of communities around the world and the increase in cases of racism and cyber racism after March 2020 [40, 41, 42, 43, 44, 45, 46]. As a result, in the age of deep learning, fast and efficient sentiment analysis techniques tend to be a must. But unfortunately, most Bangladeshi people do not use Twitter as their social media. Bangladesh is not even in the top 20 countries that use Twitter [47]. Only 7.73% of the population uses Twitter [48].

The survey was conducted through a google form. It consisted of 5 questionnaires. First, the participants were asked about their gender. After that, they were asked about their age group. There were five age groups in the survey. They are 0-20, 20-40, 40-60, 60-80 and 80+. Then it was asked whether they took the COVID-19 vaccine or not. Next, the participants were asked if they support taking the COVID-19 vaccine or not. Finally, they were asked to write supporting statements based on Yes or No from earlier sections about the COVID-19 vaccine situation. The survey was conducted from 16th February 2021 to 5th April 2021. A total of 1647 responses were recorded during that time. After finishing the dataset collection, it was found that around 57% of the responses were from males, and

around 43% were from females (figure 6). It was also discovered that only about 18% of survey participants received the vaccine, while the majority did not. In the end, it was discovered that roughly 53% of the participants endorse the COVID-19 vaccine, while the remaining 47% do not.

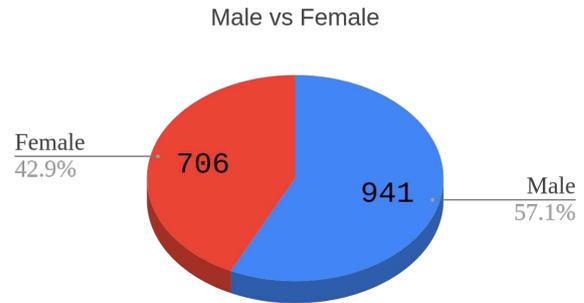


Fig. 6: Participants ratio of the survey.

The length of the comment was 2-3 sentences or 280 characters. The participants were asked to write their comment regarding their position on COVID-19 vaccine situation in English. They were also instructed to avoid writing just 1-2 words and YES or NO. They were also asked not to leave the comment section blank or write just numbers.

Class	Count	Example
Yes	751	I can accidentally spread the disease to friends, family and others. So in my view vaccination is a safer way to prevent this kind of situation.
No	657	This vaccine isn't enough trust worthy as it wasn't trailed or examined on mass people to be claimed as a perfect one. currently It's more like a pilot project to me. So in my opinion it would be better to not have this improper vaccine.

Table 1: Examples of both "Yes" and "No" classes.

Since any API is not being used in this research, it is tough to collect data. The 4th question of the survey wanted to know whether the participants support the COVID-19 vaccine or not. Using the response to this

question to label the entire dataset was labeled automatically. There are a few ways to collect surveys, such as asking over the phone, physically asking them, and sending the survey form online. Due to the outbreak of the corona pandemic, it was impossible to conduct the survey physically. The data collection was only conducted over the phone and online. But many people did not fill the survey form even though they said they would. The dataset contains various responses from 1647 people. But not all of the responses were used. Unfortunately, some outliers were found in the dataset. Because of this, 238 responses were removed. There were a few grammatical and punctuation errors in some responses. Those responses were manually corrected later on. The participants often used special characters in their comments. People often use repeating characters while using colloquial language, like “Vaccine is bad-ddddd,” “I got the vaccine, yaaayyyy!” Here characters were substituted that repeat more than twice as two characters as our final pre-processing stage. Although it was clearly stated that the language used to write the comment should be in English, some of the participants used different languages such as Bangla.

No.	Example
1	No Comments
2	.
3	Yes
4	I am agree
5	Ami vaccine ke bissas kori na karon eta te onk side effects ache

Table 2: Examples of outliers in the dataset.

4 Results

Cross-validation is a resampling technique for evaluating machine learning models on a small sample of data. The procedure has only one parameter, k , which specifies the number of groups into which a given data sample should be divided. As a result, the technique is often referred to as k -fold cross-validation. When a particular value for k is selected, it can be substituted for k in the model’s relation, for example, $k=10$ for 10-fold cross-validation. Cross-validation is a technique used in applied machine learning to estimate a machine learning model’s ability on unknown data. To use a small sample to approximate how the model would do in general when used to make predictions on data that was

not used during the model’s training. It’s a common method because it’s easy to understand and provides a less biased or positive estimation of model ability than other approaches, such as a simple train/test split. In this research, k -fold validation was used instead of an 80-20 split. This was done to check if the dataset is acceptable or not. Since the accuracy was almost similar in all the folds, the dataset is acceptable for this particular research. Another reason is the dataset used in this research is imbalanced. It has an uneven class distribution. The most intuitive success metric is precision, which is essentially the number of correctly expected observations to all observations. One might believe that if a model has high accuracy, it is the best. Accuracy is an important metric, but only when the datasets are symmetric, and the values of false positives (FP) and false negatives (FN) are nearly equal. Consequently, true positive (TP) and true negative (TN) must be considered when assessing the model’s results. The Accuracy formula is given below,

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (7)$$

The ratio of correctly predicted positive observations to the total predicted positive observations is known as precision. The low false-positive rate is related to high precision. The Precision formula is given below,

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

The ratio of correctly expected positive observations to all observations in the actual class - Yes, is known as recall. The Recall formula is given below,

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

The weighted average of Precision and Recall is the F1 Score. As a consequence, this Score considers both false positives and false negatives. F1 is typically more valuable than accuracy, mainly when class distribution is uneven. When false positives and false negatives have identical costs, consistency works better. It’s best to look at both Precision and Recall if the cost of false positives and false negatives is very different. The F1 Score formula is given below,

$$F1 = \frac{2 * (Recall * Precision)}{Recall + Precision} \quad (10)$$

4.1 Traditional Machine Learning Methods

In this research, a total of 7 traditional machine learning algorithms were used, those are: Naive Bayes(NB), Support Vector Machine (SVM), K-Nearest Neighbor

(KNN), Random Forest (RF), Gradient Boosting (GB), Ada Boost (AB), and Decision Tree(DT). Out of the seven traditional methods, Naive Bayes, Support Vector Machine, and Random Forest gave the best test results. Naive Bayes has a better Accuracy than both Support Vector Machine and Random Forest. Naive Bayes has an Accuracy of 81%. It also has the best result in Precision than the other two methods, which is 91%. On the other hand, Support Vector Machine and Random Forest both did considerably better than Naive Bayes in Recall. Both of them achieved 86% in Recall. Finally, in F1-Score, Naive Bayes did better than both Support Vector Machine and Random Forest. Naive Bayes achieved 84% in F1-Score.

Models	K-Fold 1	K-Fold 2	K-Fold 3	K-Fold 4	Test
SVM	0.79	0.76	0.79	0.77	0.79
RF	0.74	0.74	0.76	0.77	0.80
NB	0.77	0.76	0.80	0.71	0.81
KNN	0.76	0.72	0.73	0.70	0.75
GB	0.71	0.73	0.69	0.73	0.76
AB	0.72	0.71	0.71	0.72	0.78
DT	0.70	0.63	0.68	0.67	0.67

Table 3: K-Fold and Test accuracy of the traditional machine learning algorithms.

Models	K-Fold 1	K-Fold 2	K-Fold 3	K-Fold 4	Test
SVM	0.80	0.76	0.81	0.76	0.80
RF	0.75	0.74	0.77	0.76	0.81
NB	0.82	0.79	0.83	0.75	0.84
KNN	0.80	0.75	0.77	0.72	0.79
GB	0.73	0.70	0.72	0.74	0.77
AB	0.73	0.72	0.73	0.72	0.78
DT	0.71	0.63	0.70	0.63	0.69

Table 4: K-Folds and Test F1 Score of the traditional machine learning algorithms.

From the figure 7 all the traditional method's precision results can be seen. The graph contains precision

results of all k-fold and test results which was generated from the conventional methods.

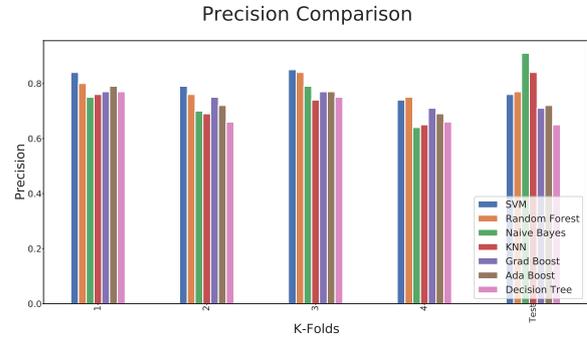


Fig. 7: K-Folds and Test precision comparison of the traditional machine learning algorithms.

The traditional method's recall results can be seen in the figure 8. The graph displays the recall of all k-fold and test results generated using traditional methods.

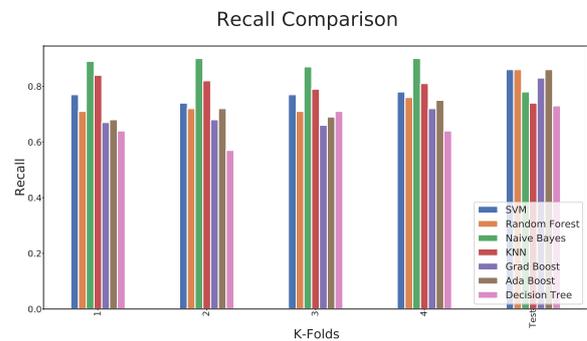


Fig. 8: K-Folds and Test recall comparison of the traditional machine learning algorithms.

4.2 Deep Neural Network (DNN)

At first, the deep neural network's minimal architecture was used to get a benchmark result shown in figure 9a. Then more layers and combinations were added to optimize the model as depicted in figure 9b. Finally, in figure 9c even more layers and combinations were included; however, the accuracy was poor. After getting poor accuracy, no more layers were added to the architecture.

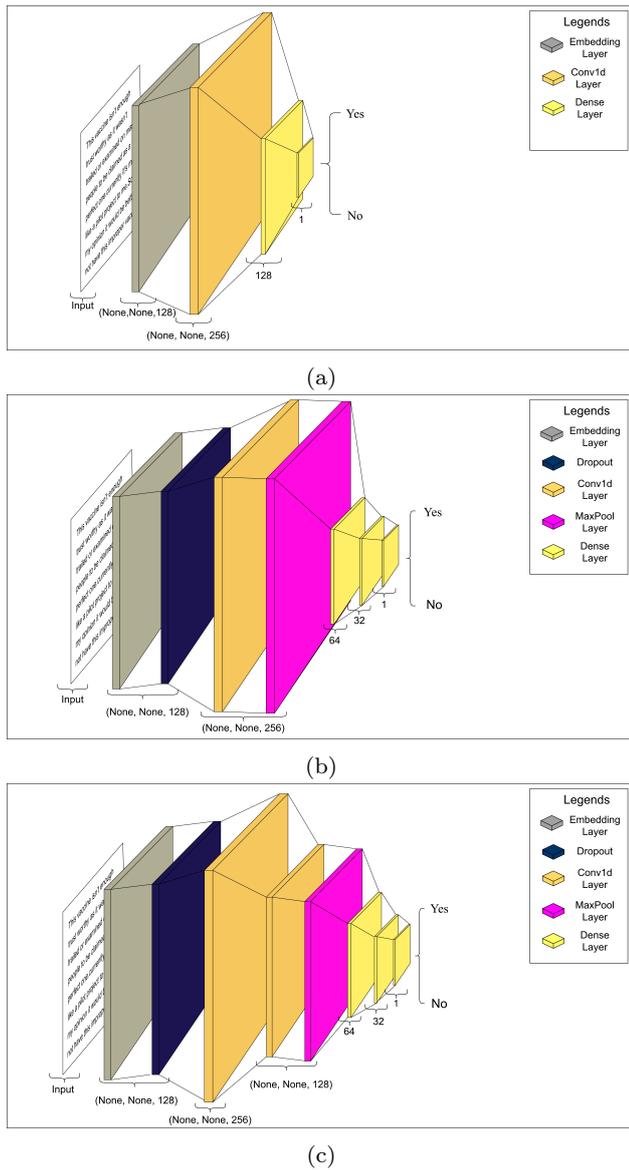


Fig. 9: a) Bare b) Moderate and c) Deep architectures of the Deep Neural Networks.

After testing all the combinations, three architectures were chosen. From table 5 it can be seen that Moderate gave the best accuracy among Bare and Deep. Moderate gave an accuracy of 76%, whereas Bare gave 62% and Deep gave 72%. Moderate's precision result was better than both Bare and Deep. Moderate got 77%, Deep got 76%, and finally Bare got 66%. Moderate and Deep achieved the same result in Recall, which is 81%, but Bare achieved the only 63%. As for F1-Score Moderate has the best outcome with 82%. Deep achieved 78%. Bare got the lowest result which is 65%. To summarize, the bare model is under-fitted, the deep

model was overfitted, and the moderate was well balanced.

Architectures	Accuracy	Precision	Recall	F1
Bare	0.62	0.66	0.63	0.65
Moderate	0.76	0.77	0.81	0.82
Deep	0.72	0.76	0.81	0.78

Table 5: Test results of the different architectures of Deep Neural Networks.

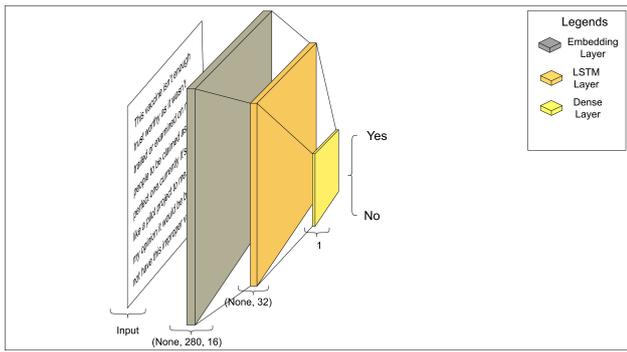
4.3 Long short-term memory network(LSTM)

From the table below, it can be seen that LSTM has the best accuracy among Simple RNN and GRU. LSTM achieved an accuracy of 79%, whereas Simple RNN got 71% and GRU got 78%.

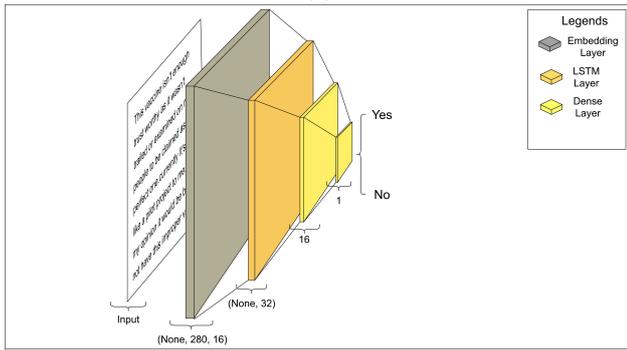
RNN Variants	Accuracy	Precision	Recall	F1
Simple RNN	0.71	0.69	0.89	0.78
LSTM	0.79	0.83	0.80	0.81
GRU	0.78	0.82	0.78	0.8

Table 6: Test results of the different variants of Recurrent Neural Networks.

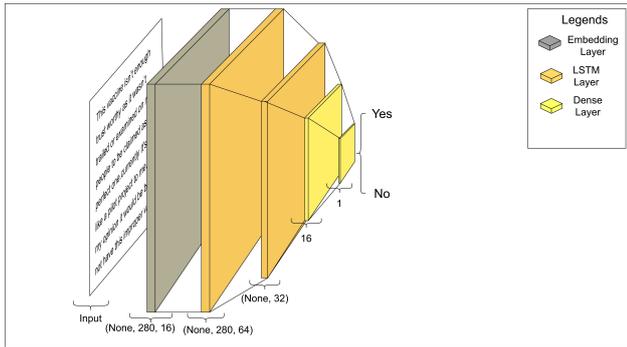
Several combinations of experiments were conducted in this research using parameter tuning. But no significant changes were recorded. Such as in [49] a large-scale analysis of eight LSTM variants on three representative tasks such as speech recognition, handwriting recognition, and polyphonic music modeling were conducted. Even after 5400 experimental runs (≈ 15 years of CPU time), none of the variants can significantly outperform the standard LSTM architecture, demonstrating that the forget gate and output activation function are the most important components.



(a)



(b)



(c)

Fig. 10: a) Bare b) Moderate and c) Deep architectures of the LSTM Networks.

After implementing, experimenting, and trying several combinations, three LSTM architectures were chosen, as seen in Figure 10. Moderate offered the best accuracy, as seen in table 7. Moderate had an 81% accuracy while Bare had a 79% and Deep had a 77%. Bare has the best result in precision, which is 83%, whereas Moderate got 81% and Deep got 77%. Moderate achieved the best result in recall. Moderate obtained 88%, Deep obtained 87%, and finally Bare obtained 80%. In F1- Score Moderate got better results than both Bare and Deep. Moderate acquired 84%, Bare acquired 81% and Deep acquired 82%.

Architectures	Accuracy	Precision	Recall	F1
Bare	0.79	0.83	0.80	0.81
Moderate	0.81	0.81	0.88	0.84
Deep	0.77	0.77	0.87	0.82

Table 7: Test results of the different architectures of LSTM Networks.

4.4 Bidirectional Encoder Representations from Transformers (BERT)

From the table 8 all results of small BERT models [34] are given. In this research, all the models were used, as seen in the table. From the above results, we can see that the change is insignificant. The L stands for hidden layers, the H for hidden size, and the A for attention heads. From the table, it is seen that L-4 H-768 A-12 and L-6 H-768 A-12 models gave the most optimal result. L-4 H-768 A-12 achieved an accuracy of 86%, and L-6 H-768 A-12 achieved 85%. On the other hand, both L-6 H-512 A-8 and L-10 H-512 A-8 gave an accuracy of 83%. On the other hand, when it comes to precision, L-4 H-768 A-12 and L-6 H-768 A-12 have the same result of 83%, whereas L-10 H-512 A-8 got 81% and L-6 H-512 A-8 got 80%. In recall, the L-4 H-768 A-12 model got the best result of 85%. But L-6 H-768 A-12 acquired 82%, L-6 H-512 A-8 acquired 81%, and finally, L-10 H-512 A-8 acquired 79%. The L-4 H-768 A-12 model got the best result in F1-Score, whereas L-6 H-768 A-12 got 82%. Both L-10 H-512 A-8 and L-6 H-512 A-8 achieved 80%.

BERT Models	Accuracy	Precision	Recall	F1
L-2 H-128 A-2	0.57	0.55	0.59	0.57
L-2 H-256 A-4	0.80	0.79	0.81	0.80
L-2 H-512 A-8	0.82	0.80	0.83	0.81
L-2 H-768 A-12	0.81	0.79	0.80	0.79
L-4 H-128 A-2	0.68	0.71	0.69	0.70
L-4 H-256 A-4	0.81	0.82	0.83	0.82
L-4 H-512 A-8	0.82	0.81	0.79	0.80
L-4 H-768 A-12	0.86	0.83	0.85	0.84
L-6 H-128 A-2	0.81	0.79	0.82	0.80
L-6 H-256 A-4	0.80	0.81	0.80	0.80
L-6 H-512 A-8	0.83	0.80	0.81	0.80
L-6 H-768 A-12	0.85	0.83	0.82	0.82
L-8 H-128 A-2	0.78	0.76	0.77	0.76
L-8 H-256 A-4	0.76	0.73	0.75	0.74
L-8 H-512 A-8	0.81	0.78	0.79	0.78
L-10 H-128 A-2	0.47	0.45	0.49	0.47
L-10 H-256 A-4	0.80	0.81	0.81	0.81
L-10 H-512 A-8	0.83	0.81	0.79	0.80
L-12 H-128 A-2	0.58	0.55	0.59	0.57
L-12 H-256 A-4	0.81	0.82	0.79	0.80
L-12 H-512 A-8	0.80	0.81	0.80	0.80

Table 8: Test results of the different small BERT models.

4.5 Analysis

Only the results of models with the best accuracy are shown in table 9. Among all the other models in this research, the models as mentioned earlier gave the most optimal results. From the table 9 it is seen that Deep Neural Network (DNN) gave the poorest accuracy of only 76%. DNN also got 77% in precision, 81% in recall, and 82% in F1- Score. On the other hand, Naive Bayes (NB) and LSTM gave almost similar results. Both of the models achieved an accuracy of 81%. However, Naive Bayes obtained better results in precision than LSTM. LSTM got better results in recall than Naive Bayes. Naive Bayes achieved 91% in precision, whereas LSTM got 81% and LSTM achieved 88% whereas Naive Bayes got 78% in recall. Both Naive Bayes and LSTM got the same result in F1- Score. Both the models got 84%. For this research, Naive Bayes is

the best choice because LSTM requires a lot of power to process. Learning LSTM models is computationally expensive for tasks that require a large number of output units and a large number of memory cells to store temporal contextual information [50]. Aside from that, LSTM is difficult to run on a GPU since it takes each input one at a time and is not parallelizable like a GPU. Another reason is sentiment analysis is primarily based on word meaning; word meaning is statistically independent. LSTM and BERT would work best on semantically dependent sentence structure, whereas Naive Bayes works best on sentiment analysis. BERT gave the highest accuracy of all the models. It achieved 86% in accuracy, 83% in precision, 85% in recall, and 84% in F1- Score. Although BERT is good in semantical analysis, it still gave a better result in sentiment analysis than Naive Bayes. But what if the dataset increases, and how it will affect the models' performance since BERT is much more resource-consuming than Naive Bayes.

Methods	Accuracy	Precision	Recall	F1
Naive Bayes	0.81	0.91	0.78	0.84
DNN	0.76	0.77	0.81	0.82
LSTM	0.81	0.81	0.88	0.84
BERT (L-4 H-768 A-12)	0.86	0.83	0.85	0.84

Table 9: The test results of different models with the best accuracy.

For this reason, another experiment was conducted with different size datasets. From the table 10 it can be seen that when a small dataset is used, BERT achieved much better results than Naive Bayes. BERT obtained 86%, and Naive Bayes got 71%. When the medium dataset was used, BERT still achieved a better result than Naive Bayes. For the medium dataset, BERT got 86.2%, and Naive Bayes got 73%. Finally, the entire dataset was used to test the models. BERT achieved slightly better results than Naive Bayes. Here BERT got 86.4%, and Naive Bayes obtained 81%. From the table below, it can be seen that as the dataset increased, the performance result of Naive Bayes increased as well. On the other hand, BERT did not achieve any mentionable change in its results. From this experiment, it can be said that Naive Bayes and BERT gave similar results, and both achieved the same in F1- Score, but if the size of the dataset increases, the performance of Naive Bayes may also increase.

Datasets	Naive Bayes	BERT (L-4 H-768 A-12)
Small (450)	0.71	0.86
Medium (900)	0.73	0.862
Full (1408)	0.81	0.864

Table 10: The overall accuracy of both models on different sized dataset.

5 Discussion

This study uses deep neural networks, long-term memory networks (LSTMs), and Bidirectional Encoder Representations from Transformers to analyze public opinion on the COVID-19 vaccine situation (BERT). This will help resolve vaccine skeptics' concerns and foster the public trust in immunization needed to understand the goal of herd immunity. All of the data is gathered for this reason through a survey of Bangladeshi people. There were a lot of outliers in the dataset. The outliers consisted of using other languages except for English, a mixture of two languages, spelling mistakes, and truncated versions of words. The outliers were handled manually. The algorithms as mentioned earlier' results are compared to see which model provides the best sentiment analysis result for the survey data. After comparing the results of all the models, BERT gave the best output, but Naive Bayes showed promising result. If the dataset is increased, then Naive Bayes may provide a better result than BERT.

6 Threats to validity

The validity of the research depends on the input the model gets. Computers are designed to process numerical data in a structured format. As a result of the inconsistency and volatility of human language, NLP becomes a difficult task, making language processing even more difficult. The meaning of words is largely determined by context and does not adhere to strict rules. For this model, other languages that use characters that are not in the English alphabet such as, Bangla, Hindi cannot be inserted, or this will cause serious problems in the model. Mixtures of two languages will also cause problems because the model will not recognize the context behind that sentence. Any spelling mistakes in the input will also raise many issues in the model. Almost everyone uses short form of some words in today's world, such as LOL and BRB. However, the model used in this study will be unable to accept these words as input.

7 Conclusion

It's worth noting that the average middle-class person is now using social media to get information and advice. A lot of research is going on sentiment analysis using various Natural Language Processing (NLP) prediction models. In this research, the deep neural network, LSTM, and BERT are used to predict public sentiment analysis on the COVID-19 vaccine in Bangladesh. The dataset of this research was collected by conducting a survey. But not all the responses were used as there were many outliers. All the outliers had to be removed manually. Traditional methods such as Naive Bayes, Support Vector Machine, K-nearest Neighbor, Random Forest, Gradient Boosting, Ada Boost, and Decision Tree were used to get a benchmark result for our model. After this deep neural network was used and then LSTM was implemented. Finally, BERT was used to get the best accuracy. Learning about the public sentiment toward the COVID-19 vaccine is essential because necessary steps must be taken to ensure that everyone knows the importance of vaccines and take the COVID-19 vaccine. As a result of developing an immune response to the SARS-Cov-2 virus, the COVID-19 vaccines provide protection against the disease. Immunity can be developed through vaccination, which reduces the risk of developing the disease and its effects.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors and Contributors: This work was carried out in close collaboration between all co-authors. First, we defined the research theme and contributed an early design of the system. We further implemented and refined the system development and wrote the paper. All authors have contributed to, seen, and approved the final manuscript.

Compliance with Ethical Standards: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent: Informed consent was obtained from all individual participants included in the study.

References

- Zu ZY, Jiang MD, Xu PP, Chen W, Ni QQ, Lu GM, et al. Coronavirus Disease 2019 (COVID-19): A Perspective from China. *Radiology*. 2020;296(2).
- Casella M, Rajnik M, Cuomo A, Dulebohn SC, Di Napoli R. Features, Evaluation, and Treatment of

- 1 Coronavirus. StatPearls Publishing, Treasure Island (FL); 2020. Available from: <http://europepmc.org/books/NBK554776>.
- 2
- 3
- 4 3. Kilimci ZH, Akyokus S. Deep Learning- and Word Embedding-Based Heterogeneous Classifier Ensembles for Text Classification. *Complexity*. 2018;2018:1–10.
- 5
- 6 4. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding; 2019.
- 7
- 8 5. Pang B, Lee L. Opinion Mining and Sentiment Analysis. *Found Trends Inf Retr*. 2008 Jan;2(1–2):1–135. Available from: <https://doi.org/10.1561/15000000011>.
- 9
- 10 6. Naresh A, Venkata Krishna P. An efficient approach for sentiment analysis using machine learning algorithm. *Evolutionary Intelligence*. 2020;
- 11
- 12 7. Kawade D, Oza K. Sentiment Analysis: Machine Learning Approach. *International Journal of Engineering and Technology*. 2017 06;9:2183–2186.
- 13
- 14 8. Chaturvedi S, Mishra V, Mishra N. Sentiment analysis using machine learning for business intelligence. In: 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI); 2017. p. 2162–2166.
- 15
- 16 9. Shathik A, Karani KP. A Literature Review on Application of Sentiment Analysis Using Machine Learning Techniques; 2020.
- 17
- 18 10. Singh J, Singh G, Singh R. Optimization of sentiment analysis using machine learning classifiers. *Human-centric Computing and Information Sciences*. 2017;7(1).
- 19
- 20 11. Raza H, Faizan M, Hamza A, Mushtaq A, Akhtar N. Scientific Text Sentiment Analysis using Machine Learning Techniques. *International Journal of Advanced Computer Science and Applications*. 2019;10(12). Available from: <http://dx.doi.org/10.14569/IJACSA.2019.0101222>.
- 21
- 22 12. Figueiredo A, Simas C, Karafillakis E, Paterson P, Larson H. Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study. *Lancet (London, England)*. 2020 09;396.
- 23
- 24 13. Hussain A, Tahir A, Hussain Z, Sheikh Z, Gogate M, Dashtipour K, et al.. Artificial intelligence-enabled analysis of UK and US public attitudes on Facebook and Twitter towards COVID-19 vaccinations; 2020.
- 25
- 26 14. Hung M, Lauren E, Hon ES, Birmingham WC, Xu J, Su S, et al. Social Network Analysis of COVID-19 Sentiments: Application of Artificial Intelligence. *J Med Internet Res*. 2020 Aug;22(8):e22590.
- 27
- 28 15. Sarlan A, Nadam C, Basri S. Twitter sentiment analysis; 2014. p. 212–216.
- 29
- 30 16. Alamoodi AH, Zaidan BB, Zaidan AA, Albahri OS, Mohammed KI, Malik RQ, et al. Sentiment analysis and its applications in fighting COVID-19 and infectious diseases: A systematic review. *Expert Systems with Applications*. 2021;167:114155. Available from: <https://www.sciencedirect.com/science/article/pii/S0957417420308988>.
- 31
- 32 17. Samuel J, Ali GGMN, Rahman MM, Esawi E, Samuel Y. COVID-19 Public Sentiment Insights and Machine Learning for Tweets Classification. *Information (Switzerland)*. 2020 06;11:1–22.
- 33
- 34 18. Ezen-Can A. A Comparison of LSTM and BERT for Small Corpus; 2020.
- 35
- 36 19. Sood S. LSTM & BERT models for Natural Language Processing (NLP). *OpenGenus IQ: Learn Computer Science*; 2020. Available from: <https://iq.opengenus.org/lstm-bert-nlp-model/>.
- 37
- 38 20. Bouchard LWA. Introduction to Transformer Networks-How Google Translate works: Attention Is All You Need. What is Artificial Intelligence; 2020. Available from: <https://medium.com/what-is-artificial-intelligence/introduction-to-transformer-networks-how-google-translate-works-attention-is-all-you-need-309827c9b942>.
- 39
- 40 21. Basic Learning Principles of Artificial Neural Networks. *International Series in Operations Research & Management Science*. 2007;p. 27–37.
- 41
- 42 22. Zayegh A, Bassam N. In: *Neural Network Principles and Applications*; 2018. .
- 43
- 44 23. Bre F, Gimenez J, Fachinotti V. Prediction of wind pressure coefficients on building surfaces using Artificial Neural Networks. *Energy and Buildings*. 2017 11;158.
- 45
- 46 24. Manickam MRM, Mohanapriya M, Kale S, Uday M, Khandagale Y, Patil SP. Research study on applications of artificial neural networks and e-learning personalization. *International Journal of Civil Engineering and Technology*. 2017 08;8:1422–1432.
- 47
- 48 25. Vaisla DK, Bhatt DA. An Analysis of the Performance of Artificial Neural Network Technique for Stock Market Forecasting. *International Journal on Computer Science and Engineering*. 2010 09;.
- 49
- 50 26. Venkatachalam M. *Recurrent Neural Networks. Towards Data Science*; 2019. Available from: <https://towardsdatascience.com/recurrent-neural-networks-d4642c9bc7ce>.
- 51
- 52 27. Phi M. Illustrated Guide to LSTM's and GRU's: A step by step explanation. *Towards Data Science*; 2020. Available from: <https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21>.
- 53
- 54 28. Tomar A, Gupta N. Prediction for the spread of COVID-19 in India and effectiveness of preventive measures. *Science of The Total Environment*. 2020;728:138762.
- 55
- 56 29. am a Senior Undergraduate at IIT (BHU) PS. Long Short Term Memory: Architecture Of LSTM; 2020. Available from: <https://www.analyticsvidhya.com/blog/2017/12/fundamentals-of-deep-learning-introduction-to-lstm/>.
- 57
- 58 30. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention Is All You Need. *CoRR*. 2017;abs/1706.03762. Available from: <http://arxiv.org/abs/1706.03762>.
- 59
- 60 31. Horev R. BERT Explained: State of the art language model for NLP. *Towards Data Science*; 2018. Available from: <https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270>.
- 61
- 62 32. Explanation of bert model - nlp;. Available from: <https://www.geeksforgeeks.org/explanation-of-bert-model-nlp/>.
- 63
- 64 33. Classify text with BERT ; TensorFlow Core;. Available from: https://www.tensorflow.org/tutorials/text/classify_text_with_bert.
- 65
34. Turc I, Chang M, Lee K, Toutanova K. Well-Read Students Learn Better: The Impact of Student Initialization on Knowledge Distillation. *CoRR*. 2019;abs/1908.08962. Available from: <http://arxiv.org/abs/1908.08962>.
35. Jagdale R, Shirsat V, Deshmukh S. In: *Sentiment Analysis on Product Reviews Using Machine Learning Techniques: Proceeding of CISC 2017*; 2019. p. 639–647.
36. Ramya V, Komati TR. Sentiment Analysis of Movie Review using Machine Learning Techniques. *International Journal of Engineering & Technology*. 2018 03;7:676.

- 1 37. J AM, Ravikumar P. Article: Survey: Twitter data Anal-
2 ysis using Opinion Mining. *International Journal of Com-*
3 *puter Applications*. 2015 October;128(5):34–36. Pub-
4 lished by Foundation of Computer Science (FCS), NY,
5 USA.
- 6 38. Maindola P, Singhal N, Dubey A. Sentiment Analysis of
7 Digital Wallets and UPI Systems in India Post Demoneti-
8 zation Using IBM Watson. 2018 International Conference
9 on Computer Communication and Informatics (ICCCI).
10 2018;p. 1–6.
- 11 39. Wang Y, Taylor JE. Coupling sentiment and human
12 mobility in natural disasters: a Twitter-based study of
13 the 2014 South Napa Earthquake. *Natural Hazards*.
14 2018;92(2):907–925.
- 15 40. Dubey A. Twitter Sentiment Analysis during COVID19
16 Outbreak. *SSRN Electronic Journal*. 2020 01;.
- 17 41. H Manguri K, N Ramadhan R, R Mohammed Amin
18 P. Twitter Sentiment Analysis on Worldwide COVID-
19 19 Outbreaks. *Kurdistan Journal of Applied Research*.
20 2020;p. 54–65.
- 21 42. Bhat M, Qadri M, ul Asrar Beg N, Kundroo M,
22 Ahanger N, Agarwal B. Sentiment analysis of so-
23 cial media response on the Covid19 outbreak. *Brain,*
24 *Behavior, and Immunity*. 2020;87:136–137. Avail-
25 able from: [https://www.sciencedirect.com/science/](https://www.sciencedirect.com/science/article/pii/S0889159120307091)
26 [article/pii/S0889159120307091](https://www.sciencedirect.com/science/article/pii/S0889159120307091).
- 27 43. Pastor CK. Sentiment Analysis of Filipinos and Effects
28 of Extreme Community Quarantine Due to Coronavirus
29 (COVID-19) Pandemic. *SSRN Electronic Journal*. 2020;.
- 30 44. Dubey A, Tripathi S. Analysing the Sentiments to-
31 wards Work-From-Home Experience during COVID-19
32 Pandemic. *Journal of Innovation Management*. 2020 04;8.
- 33 45. Rich M. As Coronavirus Spreads, So Does Anti-Chinese
34 Sentiment. *The New York Times*; 2020. Avail-
35 able from: [https://www.nytimes.com/2020/01/30/](https://www.nytimes.com/2020/01/30/world/asia/coronavirus-chinese-racism.html)
36 [world/asia/coronavirus-chinese-racism.html](https://www.nytimes.com/2020/01/30/world/asia/coronavirus-chinese-racism.html).
- 37 46. Dubey AD. The Resurgence of Cyber Racism During
38 the COVID-19 Pandemic and its Aftereffects: Analysis
39 of Sentiments and Emotions in Tweets. *JMIR Public*
40 *Health and Surveillance*. 2020;6(4).
- 41 47. Tankovska H. Twitter: most users by country; 2021.
42 Available from: [https://www.statista.com/](https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/)
43 [statistics/242606/number-of-active-twitter-](https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/)
44 [users-in-selected-countries/](https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/).
- 45 48. Social Media Stats Bangladesh;. Available from:
46 [https://gs.statcounter.com/social-media-](https://gs.statcounter.com/social-media-stats/all/bangladesh)
47 [stats/all/bangladesh](https://gs.statcounter.com/social-media-stats/all/bangladesh).
- 48 49. Greff K, Srivastava RK, Koutník J, Steunebrink BR,
49 Schmidhuber J. LSTM: A Search Space Odyssey. *CoRR*.
50 2015;abs/1503.04069. Available from: [http://arxiv.](http://arxiv.org/abs/1503.04069)
51 [org/abs/1503.04069](http://arxiv.org/abs/1503.04069).
- 52 50. Sak H, Senior AW, Beaufays F. Long Short-Term
53 Memory Based Recurrent Neural Network Architec-
54 tures for Large Vocabulary Speech Recognition. *CoRR*.
55 2014;abs/1402.1128. Available from: [http://arxiv.](http://arxiv.org/abs/1402.1128)
56 [org/abs/1402.1128](http://arxiv.org/abs/1402.1128).

Figures

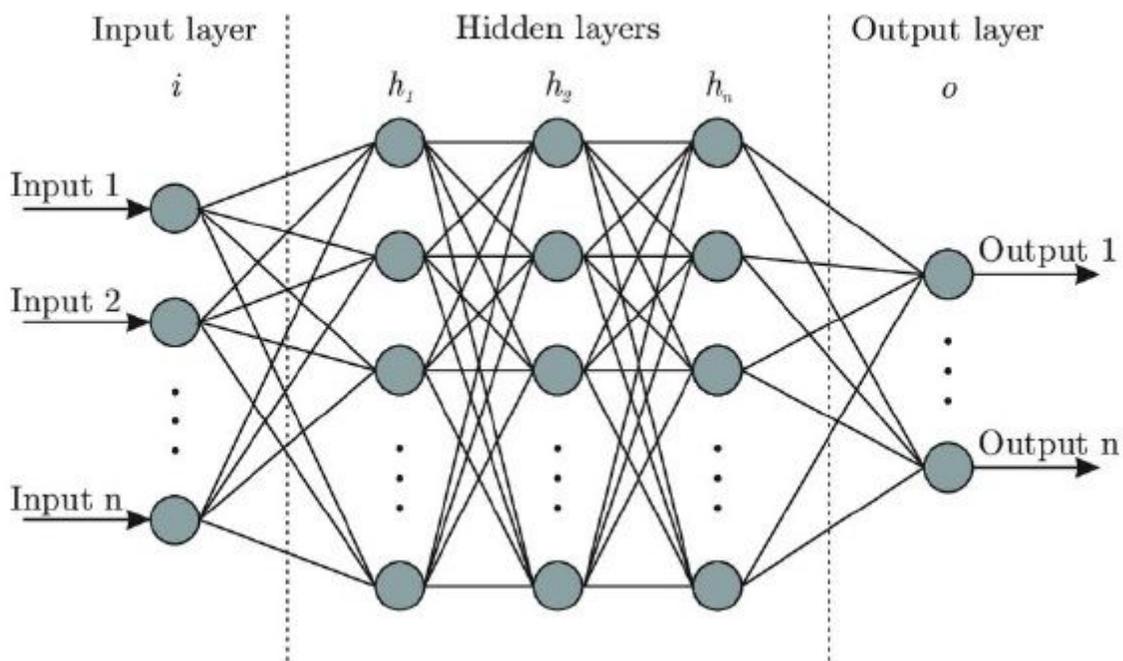


Figure 1

Architecture of Deep Neural Networks (DNN) [23].

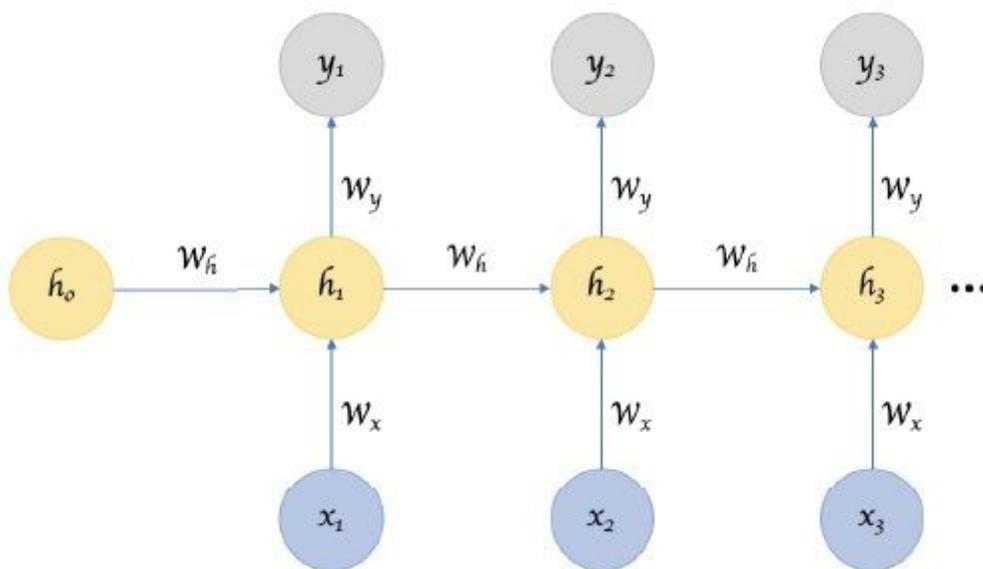


Figure 2

Architecture of Recurrent Neural Networks (RNN) [26].

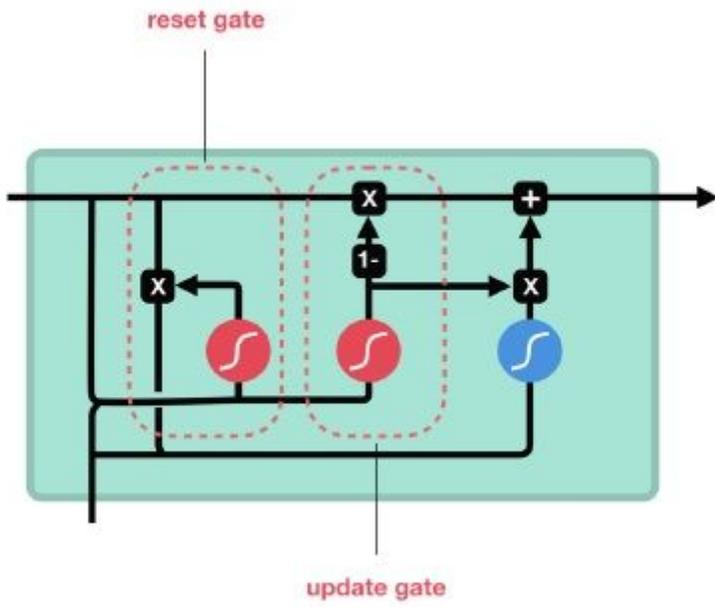


Figure 3

Architecture of Gated Recurrent Unit (GRU) [27].

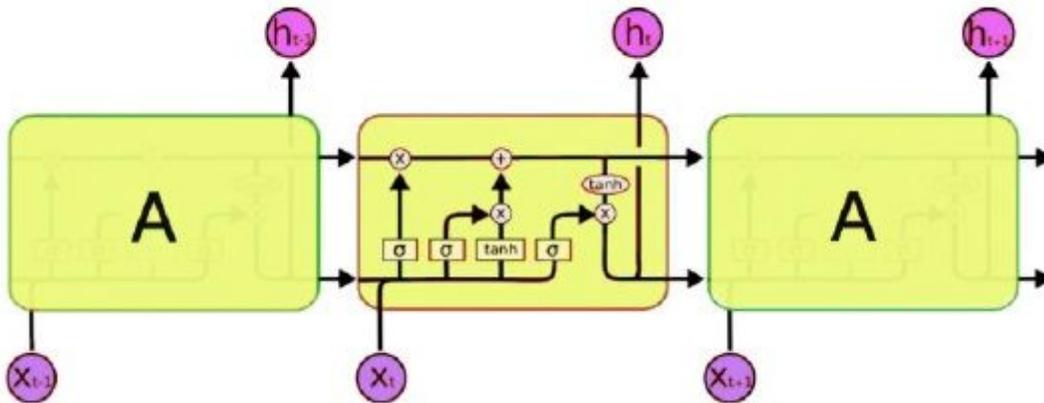


Figure 4

Architecture of Long Short-Term Memory (LSTM) [29].

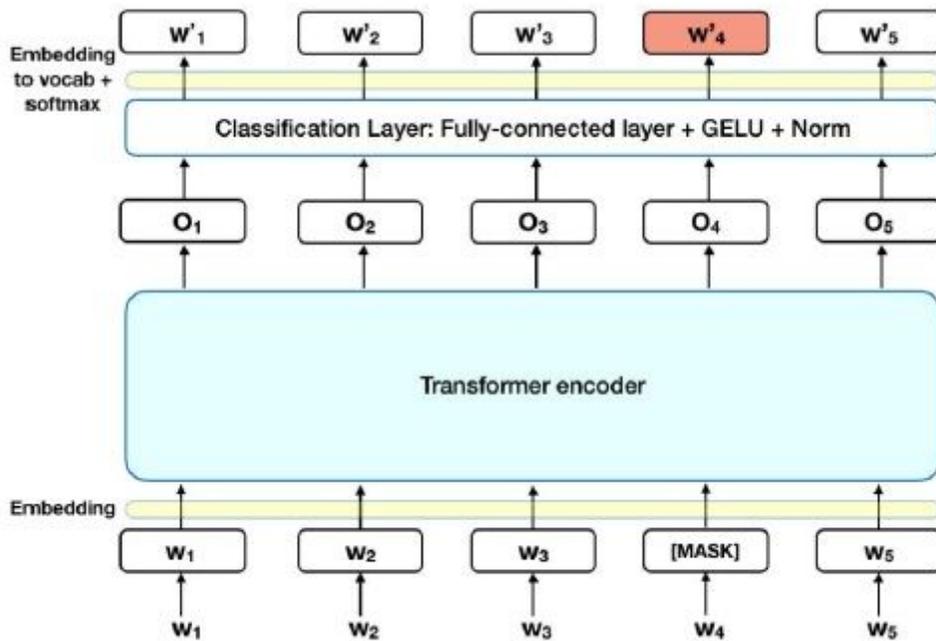


Figure 5

Architecture of Bidirectional Encoder Representations from Transformers (BERT) [31].

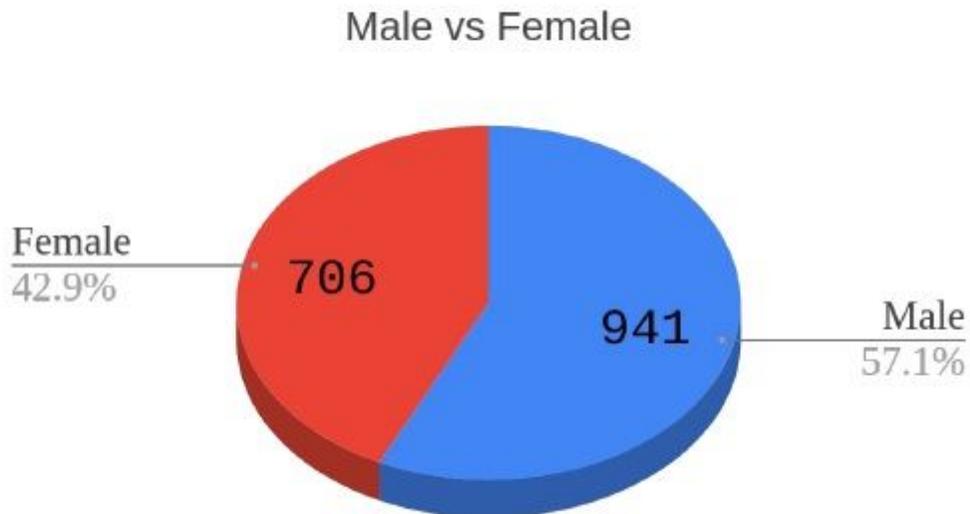


Figure 6

Participants ratio of the survey.

Precision Comparison

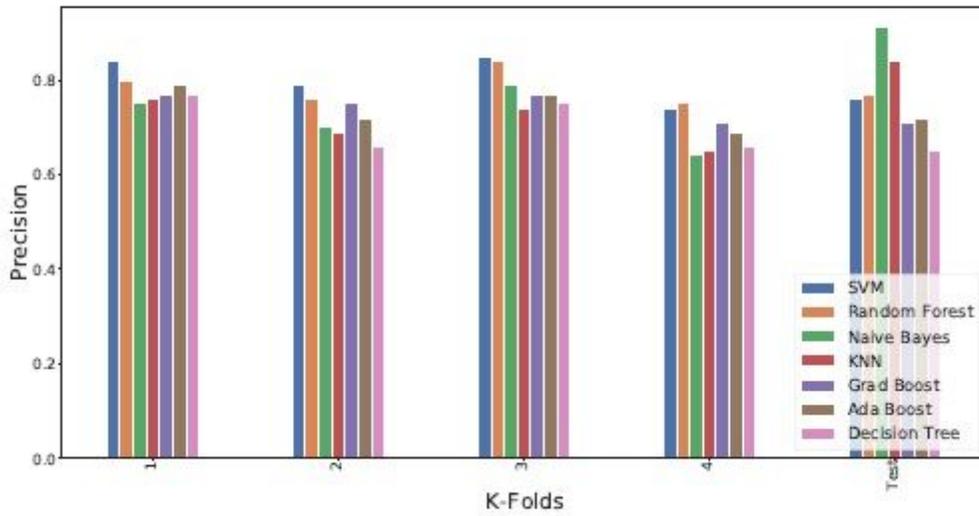


Figure 7

K-Folds and Test precision comparison of the traditional machine learning algorithms.

Recall Comparison

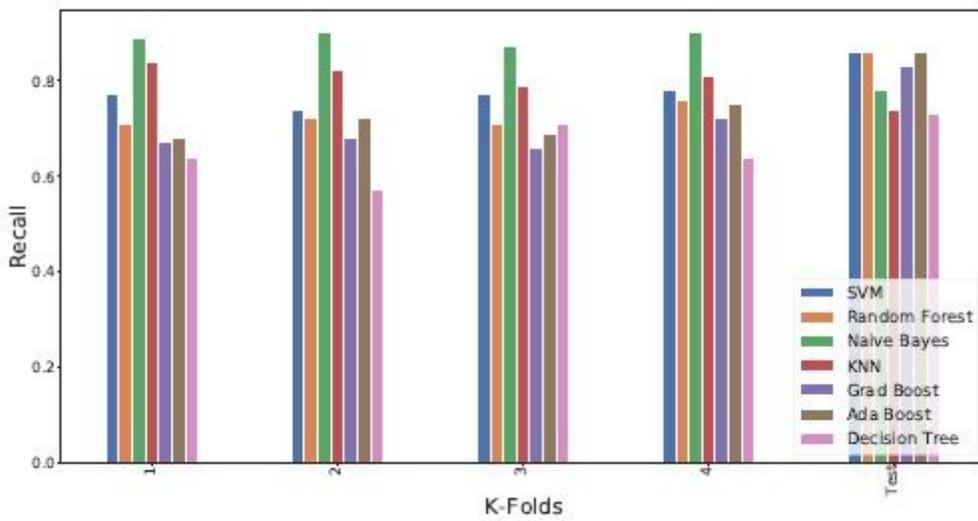
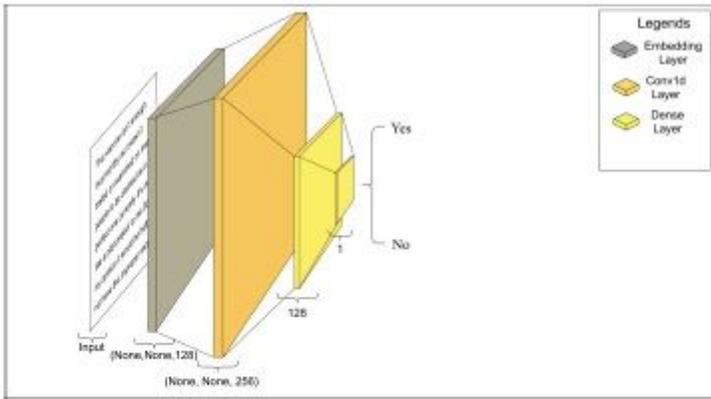
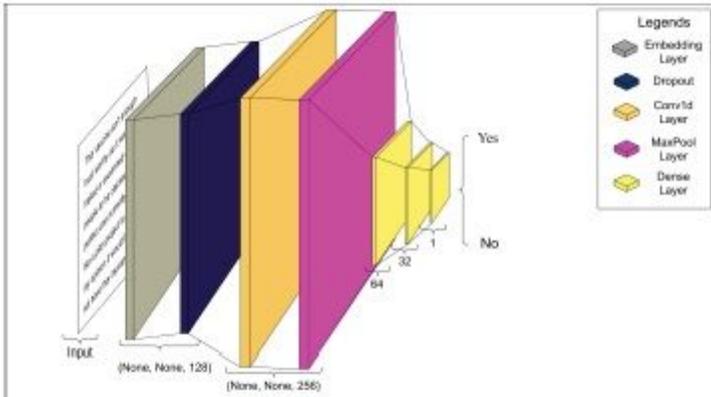


Figure 8

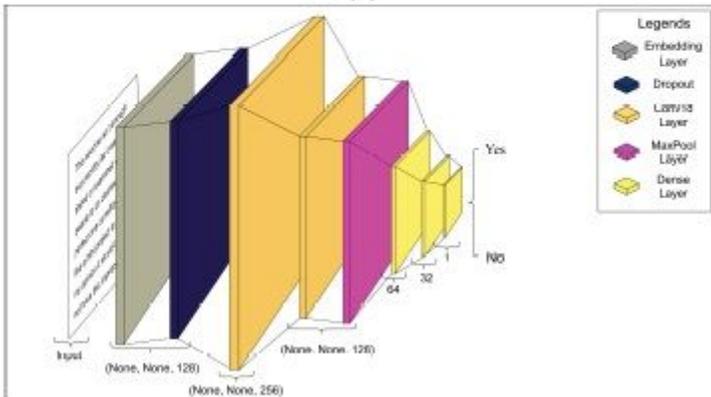
K-Folds and Test recall comparison of the traditional machine learning algorithms.



(a)



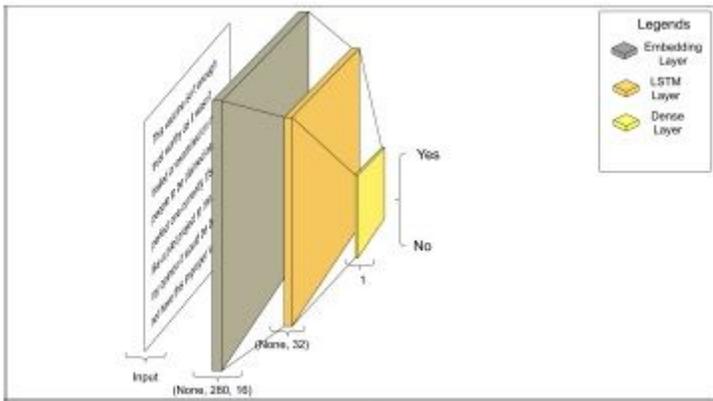
(b)



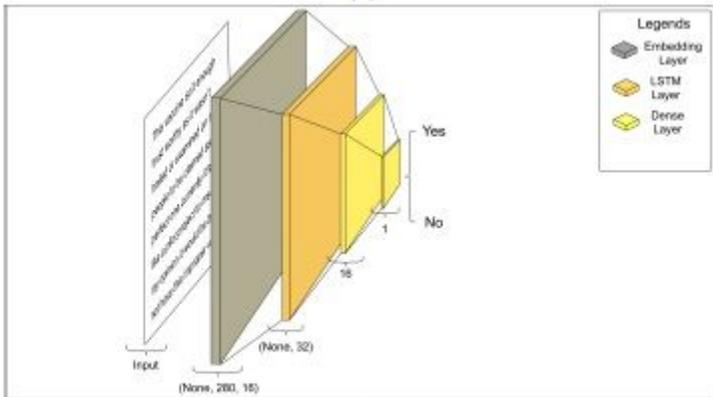
(c)

Figure 9

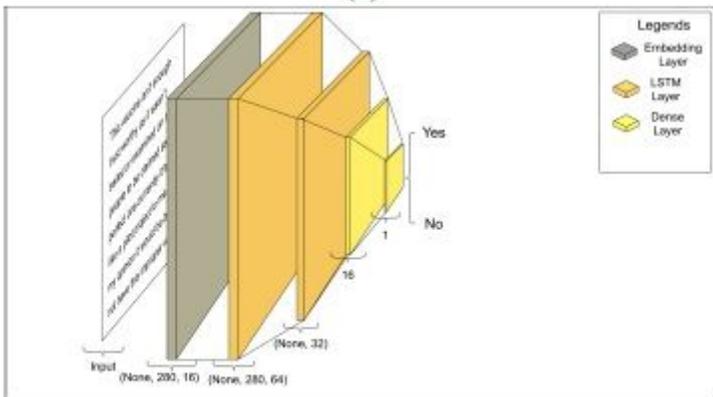
a) Bare b) Moderate and c) Deep architectures of the Deep Neural Networks.



(a)



(b)



(c)

Figure 10

a) Bare b) Moderate and c) Deep architectures of the LSTM Networks.