

## Research Article

# NNPCov19: Artificial Neural Network-Based Propaganda Identification on Social Media in COVID-19 Era

Akib Mohi Ud Din Khanday <sup>1</sup>, Bharat Bhushan <sup>2</sup>, Rutvij H. Jhaveri <sup>3</sup>,  
Qamar Rayees Khan <sup>1</sup>, Roshani Raut <sup>4</sup>, and Syed Tanzeel Rabani <sup>1</sup>

<sup>1</sup>Department of Computer Sciences, BGSB University, Rajouri, India

<sup>2</sup>Department of Computer Science and Engineering, Sharda University, Greater Noida, India

<sup>3</sup>Department of Computer Science and Engineering, Pandit Deendayal Energy University, Raysan, India

<sup>4</sup>Department of Information Technology, Savitribai Phule Pune University, Pune, India

Correspondence should be addressed to Rutvij H. Jhaveri; [rutvij.jhaveri@sot.pdpu.ac.in](mailto:rutvij.jhaveri@sot.pdpu.ac.in)

Received 7 February 2022; Revised 10 June 2022; Accepted 7 July 2022; Published 5 August 2022

Academic Editor: Sikandar Ali

Copyright © 2022 Akib Mohi Ud Din Khanday et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The latest trend of sharing information has evolved many concerns for the current researchers, which are working on computational social sciences. Online social network platforms have become a tool for sharing propagandistic information. This is being used as a lethal weapon in modern days to destabilize democracies and other political or religious events. The COVID-19 affected almost every corner of the world. Various propagandistic tweets were shared on Twitter during the peak time of COVID-19. In this paper, improved artificial neural network algorithm is proposed to classify tweets into propagandistic and non-propagandistic class. The data are extracted using multiple ambiguous hashtags and are manually annotated into binary class. Hybrid feature engineering is being performed by combining “Term Frequency (TF)/Inverse Document Frequency (IDF),” “Bag of Words,” and Tweet Length. The proposed algorithm is compared with logistic regression, support vector machine, and multinomial Naive Bayes. Results showed that improved artificial neural network algorithm outperforms other machine learning algorithms by having 77.15% accuracy, 77% of recall, and 79% precision. In future, deep learning approaches like LSTM may be used for this classification task.

## 1. Introduction

The 21st century has been marked by disinformation in the form of false news, fuelled by technology such as social media networks that allow information to spread rapidly and target individual views, prejudices, and emotions. Culture is one aspect that may be of significance. Cultures differ in many ways, including individualism vs. collectivism, the distance of influence (between those at the top of the social hierarchy and those at the bottom), masculinity vs. femininity, and avoidance of ambiguity [1]. It is fair to assume that different kinds of fake news in different cultures would spread more easily. For example, in a highly collectivist society, likely, a fake news story of a well-known person transgressing a social standard will spread faster than in an

individualistic community. If this is indeed, then the question arises as to whether the writers of fake news target particular communities with specific kinds of fake news that they think are optimal for that culture [2]. Making heuristics in decision-making is an essential factor that promotes fake news. Many distinct heuristics have been established, such as learning confidentiality, for example, which results in us believing that information presented as leaked or confidential is more likely to be true in some way. It is an under-researched area, but the relationship between culture and heuristics is possible. For instance, in cultures where there is a high power distance, the heuristic secrecy example may be more pronounced, as people are less used to learning details of those at the top of the power hierarchy [3]. In culmination, we would like to say that whether the news is fake or

real, we should go in the depth of the issues before believing or acting on it.

Online social networks have evolved as a communication tool, and social networks play a vital role in modern time. The social networks have made communication so easy that a person can communicate with other people who reside in different corners of the world. The posts that are posted on these platforms reach to every corner of the world in no time. With the numerous benefits of these platforms, it has its limitations. Some hatemongers and filthy users use these platforms for spreading hate and filthy information. The offensive information that is shared/posted can be categorized into two classes that are misinformation and disinformation. Misinformation is the type of information shared by a person without knowing its truthfulness and can harm the society in many ways [4]. The misinformation can be regarding various topics like elections, trending events, and religion, while disinformation is the type of information posted/shared intentionally. In this type of messages, the user knows that he/she is sharing false information. These types of posts are very harmful for the whole society/country. Recently, during COVID-19 times, numerous amount of misinformation is shared on social media platforms. Misinformation related to vaccines is posted by fear mongers, which produce a threat for the life of people across the globe [5].

Another type of information that is being shared on social media is known as propagandistic posts. Propaganda is being used as a weapon for reduplicating sentiments by shaping a person's behaviour [6]. The main aim of spreading propaganda is to create a mess in the mind of an ordinary person. These types of posts can be graphical, textual, and multimedia. The propagandistic information is related to the events related to politics, religion, sectarian discussion, and celebrities [7]. These posts are deliberately made so that a user will not understand that the message is propagandistic in nature. In some democratic states, propaganda is also shared through traditional media like news channels as political parties sponsor various news channels. These media channels are only used to spread hate and propagandistic news regarding the world's events as people now are aware of these paid channels, which spread propagandistic information. People shift their way of gaining knowledge from traditional media to many advanced media known as social media. But these platforms share more propagandistic messages as compared to traditional media. According to our knowledge, no such system is being made to identify the posts in real-time, whether a post is propagandistic or nonpropagandistic. The modern age is the age of artificial intelligence, and the AI has shown better results in every field, be it medical, education, robotics, etc. [8, 9]. Now machine learning is being used for analysing the data. Machine learning has shown better accuracy in classifying the text, images, etc. Machine learning is also used for identifying the spammers and rumours on social media [10]. Machine learning can be achieved using supervised and unsupervised learning. The machine learning algorithm is trained in supervised learning to be acquainted with the output class [11]. In unsupervised learning, we feed data to

machine learning algorithm without output class [12, 13]. There are various supervised machine learning algorithms, which can be used for classification tasks. Some of the algorithms are support vector machine, Naive Bayes, and decision tree. In this paper, we focus on classifying the tweets using artificial neural networks. Some of the novel contributions regarding this article are as follows:

- (i) A framework is proposed, which focus on machine learning algorithms for detecting the propaganda on online social networks.
- (ii) Novel dataset was prepared by annotating the tweets extracted using various keywords like #CORONAJIHAD #CORONAMUSLIM #CHINESEVIRUS #BIOWEAPON.
- (iii) Data annotation is done with the help of various techniques such as Name Calling, Bandwagon, and Testimonial.
- (iv) Feature engineering is improved by merging different features like "TF/IDF," "Bag of Words," and Tweet Length for training machine learning algorithms.
- (v) Artificial neural network is fine-tuned and is trained based on the features selected through proposed hybrid feature selection method.

This article consists of 6 sections that are briefly defined as Section 1 that describes the brief introduction about COVID-19 and the effect of misinformation and propaganda during these times. In Section 2, a detailed background is being provided related to misinformation and propaganda on social networks. In Section 3, detailed background of machine learning algorithms is provided. Proposed methodology is being discussed in Section 4. Section 5 discusses the results, while Section 6 concludes our work.

## 2. Related Work

The increasing allure and beauty of social media use has an effect on our daily lives, whether directly or indirectly by convincing us to believe other people's thoughts and recommendations while making little or large choices, such as purchasing goods and voting online for elections to establish a new government. It is unsurprising that social media has evolved into a tool for altering feelings through the dissemination of disinformation. On social media, false material and propaganda are widely utilized, and they must be recognized and fought. Gao et al. [14] examined the features of fraudulent accounts, with a particular emphasis on messages including URLs in textual format. With the rise in social media users, political events and government programmes are more debated, and social media misuse is becoming more frequent. Ratkiewicz et al. [15] proposed a technique to detect and track political misuse on social networks. The network structure exhibits slight deviation from the usual structure throughout the election. During election season, Halu et al. [16] looked at the structure of social networks. A model for opinion dynamics was

proposed to demonstrate the wide area parties survived by accumulating limited proportion of the vote throughout election season. Ramakrishnan et al. [17] studied and mined data from online social networks relating to civil unrest. They extracted data pertaining to the specific incident, identified contributing elements, and analysed the event's evolution. Liu et al. [18] discovered the inner sanctums of political power holders in governments beneath official job relationships and tracked the formation and evolution of the selected political groups across time. They accomplished their goal through three distinct techniques: constructing networks, identifying communities, and monitoring community evolution. Bhat et al. [19] developed a revolutionary technique for detecting stealthy accounts on online social networking sites. The processes include detecting communities at the node level, defining features, classifying them, and locating stealthy Sybils. Khanday et al. [20] detected coronavirus from clinical textual reports using machine learning techniques. Tweets, retweets, and re-tweeters were used to calculate political leanings. Wong et al. [21] calculated behaviour of Twitter users in the methodical process. Various occurrences occur simultaneously as a result of social media. Zarrinkalam and Bagheri [22] recognised social network's events. To find undefined events on social networks, three models were used: latent Dirichlet allocation (LDA) for topic modelling, *TF/IDF* for document clustering, and wavelet analysis for feature clustering. The impact of social bots on politics was investigated (political propaganda using social media bots). Lightfoot [23] believed that social bots have solely a negative impact on politics. They discovered that social bots play a critical role in the propagation of false news and that profiles who consistently disseminate disinformation are much more likely to be bots. People who hold extreme views solely use social media to promote hatred and terror. Ashcroft et al. [24] suggested a semantic graph-based approach to identifying radicalization via social media. Anti-ISIS users often discuss politics, geographical regions, and counter-ISIS activities, while pro-ISIS users frequently discuss religions, historic events, and ethnicity. They also posted the shady information on social media. Using cognitive psychology, Kumar et al. [25] were able to detect disinformation on online social networks. The cognitive process is made up of the message's consistency, coherence, source credibility, and general acceptance. Using the Twitter API, they gathered datasets using the hashtags #Syria, #Egypt, and others. Mazzoleni et al. [26] socially mediated form of popular communication, which was recognized as being significantly influenced by the nature of social media. "Appeal to the people," "attacking the Elite," and "ostracizing the others" were used to gauge the degree of populist communicative ideology. With the spread of COVID-19, various propagandistic textual posts were posted on Twitter using various hashtags. Mujahid et al. [27] analysed the sentiments of various topics that were shared during COVID-19 era through tweets and most of the tweets were related to the online education. With the help of artificial intelligence, various models were developed for predicting COVID-19. Ashraf et al. [28] analysed the published work that is related to COVID-19, and they

grouped such works under nine categories for age, gender, the outcome of the disease, method used for prediction, and role of underlying medical conditions. Table 1 summarizes recent works regarding propaganda, in Table 1. P, R, and F1 are precision, recall, and F measure, respectively.

The following conclusion can be drawn from the above literature reviews.

- (i) Clustering algorithms are used for identifying propaganda, classification algorithms may improve the accuracy, and it needs further investigations.
- (ii) Majority of work is being done on news-let datasets, and work can be explored to online social media platforms.
- (iii) Semantic nature of the propagandistic posts due to semantic nature of the posts and it is a challenge to detect propaganda for that manual annotation is necessary.
- (iv) Feature engineering can be explored by choosing various features such as semantic and emphatic.

### 3. Background Knowledge

Machine learning emerged as a separate discipline of computer science due to rapid development of data mining techniques and procedures. It can be thought of as a subfield of artificial intelligence, with the primary principle being that a system (computer program, algorithm, etc.) may learn from its own activities. Arthur Samuel first described it in 1959 as "a subfield of study that focuses on the ability of computers to learn without being explicitly programmed." "A computer program is said to learn from experience *E* with respect to a class of tasks *T* and a performance metric *P* if its performance on tasks in *T*, as measured by *P*, increases as a result of the experience *E*" [41]. The basic aim of machine learning algorithm is to perform a specific task based on its training. The model is developed using the input dataset as a basis for training, and it is then utilized to make predictions. Machine learning is being used in many areas of computer science such as the Internet of Things (IoT), image processing, and natural language processing (NLP). The machine learning nowadays is being used to detect various daily activities of humans [42]. With the eruption of COVID-19, machine learning plays a vital role in detecting the virus with the help of X-ray chest images [43]. In text mining, machine learning algorithms are used for classifying the sentiments of various text, and the text may be a document, a review of a product, a news, and a tweet [44].

Propaganda detection can be seen as a classification/clustering problem in machine learning; that is, unknown propaganda text should be clustered into different clusters depending on certain attributes detected by the algorithm. On the other hand, we can reduce this problem to classification after training a model on a large dataset of propaganda and nonpropaganda files. This challenge can be simplified down to classification only for known propaganda by having a small set of classes. One being propaganda class, and other being nonpropaganda class. In classification

TABLE 1: Summary of recent work.

Author	Year	Contribution	Results	Research gap
Morio et al. [29]	2020	The author used PoS with CNN	F1: 51.55, P: 56.54, and R: 47.37	More feature engineering can improve the performance.
Jurkiewicz et al. [30]	2020	The authors used CRF	F1: 49.15, P: 59.95, and R: 41.65	More data may improve performance.
Chernyavskiy et al. [31]	2020	The authors used embeddings with CRF	F1: 49.10, P: 53.23, and R: 45.56	Non-English language can be used.
Khosla et al. [32]	2020	The authors used Bag of Words with LSTM	F1: 47.66, P: 50.97, and R: 44.76	TF/IDF can be used with LSTM.
Paraschiv and Cercel [33]	2020	The authors used embeddings with LSTM	F1: 46.6, P: 58.61, and R: 37.94	More data can improve performance.
Dimov et al. [34]	2020	The authors used $n$ -Grams with LSTM	F1: 44.68, P: 55.62, and R: 37.34	The authors have only used $n$ -Gram feature other features may be considered in the future.
Blaschke et al. [35]	2020	The authors used PoS with SVM	F1: 43.86, P: 42.16, and R: 45.7	Different kernels of SVM can be explored in order to get more accuracy.
Verma et al. [36]	2020	The authors used ELMo with CNN	F1: 43.60, P: 49.86, and R: 38.74	More feature engineering is possible.
Singh et al. [37]	2020	The authors used PoS with BERT	F1: 42.21, P: 46.52, and R: 38.63	More data are required for BERT, and the authors only used 536 articles.
Ermurachi and Gifu [38]	2020	The authors used Bag of Words with MNB and with LR	F1: 33.21, P: 24.49, and R: 51.57	Other features like sentimental, and emphatic may improve the performance.
Dewantara et al. [39]	2020	The authors used embeddings with CNN	F1: 23.47, P: 22.63, and R: 24.38	More embeddings can be used in order to achieve better results.
Daval-Frerot and Yannick [40]	2020	The authors used embeddings with RF	F1: 18.18, P: 34.14, and R: 12.39	Bag of Words and TF/IDF may be used with RF.

problem, it is easier to identify the right class, and the result is more accurate than with clustering methods. Some of the machine learning algorithms are described below.

**3.1. Logistic Regression.** Based on the link between the numerical variable and the label, this method guesses the numerical variable's class [45]. The features that have been selected using feature engineering values are displayed in a table and given as an input. In general, the procedure determines the likelihood of being a member of a given class. We have two classes here.  $y \in 0, 1$ . Ref. (2) can be used to figure out the posterior probability.

$$P(y = k|x) = \frac{\exp^{\phi^{T\theta_k}}}{1 + \sum_{k=1}^3 \exp^{\phi^{T\theta_k}} \forall k = 1}, \quad (1)$$

$$P(y = 0|x) = \frac{\exp^{\phi^{T\theta_0}}}{1 + \sum_{k=1}^3 \exp^{\phi^{T\theta_k}}}. \quad (2)$$

**3.1.1. Multinomial Naïve Bayes (MNB).** Using the Bayes rule [46], MNB calculates the class probabilities of a given text. It can be used to solve problems with binary and multiclass categorization. The essential concept is that each feature should be treated separately. The Naive Bayes algorithm makes use of the Bayes theorem to determine the probability of each attribute independently of any relationships. This is why this strategy is referred to as "naive": characteristics in real-world tasks frequently have some degree of association. In our case, let  $A$  represent the collection of classes. There are two types of classes:  $A = 0$  and  $A = 1$ . Furthermore,  $N$  denotes

the set of features. Then, using the Bayes rule described in Ref. (3), MNB selects the class with the best probability  $P(a|ti)$  for the test text  $ti$ :

$$P(a|t_1i) = \frac{(P(a)P(t_i|a))}{P(t_i)}, \quad a \in A. \quad (3)$$

$P(a)$  is calculated by dividing the total number of textual records by the number of records labelled as class  $a$ .  $P(ti|a)$  is the likelihood of finding a textual data similar to  $t_i$  in class  $a$ . It can be calculated as follows:

$$P(t_i|a) = \frac{(\sum_n f_{ni})! \Pi_n P(w_n|a)^{f_{ni}}}{f_{ni}!}, \quad (4)$$

where  $f_n^i$  denotes the number of instances of the word "n" in our dataset.  $t_i$  and  $P(w_n|c)$  are the likelihood of the word "n" being presented in class  $a$ . The latter probability is derived using the training data using the Ref. (5):

$$P(w_n|a) = \left( \frac{1 + F_{na}}{N + \sum_{x=1}^N F_{xa}} \right), \quad (5)$$

where  $F_{xc}$  are the instances of the word "x" in all training data corresponding to the class  $a$ . The Laplace estimator is utilized for avoiding zero-frequency problem, which assigns a value of one to the count of each word.

The simplicity and ease with which this strategy can be understood are two of its main features. Furthermore, it performs effectively on datasets with irrelevant attributes since the probability of these features having an effect on the outcome is low. This means that when making projections, they will not be taken into consideration. Additionally, this strategy often results in low resource consumption because it

only involves calculating the probabilities of features and classes; no coefficients need to be identified, as is the case with other methods. As mentioned previously, its key shortcoming is that every features are treated independently, despite the fact that this is impossible in the vast majority of cases.

**3.2. Support Vector Machine (SVM).** The support vector machine (SVM) algorithm is a method for categorizing text using supervised machine learning [47]. The key concept is to discover such a hyperplane that best separates the classes. The term “support vector” refers to the points nearest to the hyperplane that, if removed, would alter the hyperplane’s location. The margin is the distance between the support vector and the hyperplane. It necessitates a predetermined amount of features for the given text with predetermined label. The training set’s data points are denoted by  $(y_k, x_k)_1^n$ , where “ $n$ ” denotes number of features extracted. The values associated with the features selected during feature engineering are represented in a table and offered as a source of input. SVM’s primary objective is building of classifier using the following equation:

$$y(x) = \text{sign} \left[ \sum_{(k=1)}^n \alpha_k y_k \psi(x, x_k) + b \right], \quad (6)$$

where  $\alpha_k$  is the positive real constant, and  $b$  is the real constant.

$$\psi(x, x_k) = \begin{cases} x_k^T, & x: \text{Linear SVM}, \\ (x_k^T x + 1)^d, & \text{Polynomial SVM with degree } d, \end{cases} \quad (7)$$

where  $k$  and  $\alpha$  are the constants.

The classifier is constructed on the basis of the following assumptions:

$$\begin{aligned} \omega^T \varphi(x_k) + b &\geq 1, & \text{if } y_k = +1, \\ \omega^T \varphi(x_k) + b &\leq -1, & \text{if } y_k = -1. \end{aligned} \quad (8)$$

This is equivalent to

$$y_k \omega^T \varphi(x_k) + b \leq -1, \text{ if } y_k = -1, \quad k = 1, \dots, n, \quad (9)$$

where  $\varphi(\bullet)$  is the nonlinear function that converts input spaces to a higher-dimensional space. The hyperplane is used to accomplish categorization. The hyperplane distinguishes the two classes, requiring the introduction of a new variable  $\xi_k$ . The hyperplane equation is

$$y_k [\omega^T \varphi(x_k) + b] \geq 1 - \xi_k, \quad k = 1, \dots, n \quad (10)$$

$$\xi_k \geq 0, \quad k = 1, \dots, n \quad (11)$$

SVMs are often capable of achieving high accuracy, particularly on “cleaned” datasets. Additionally, it works well with high-dimensional datasets, even when the number of dimensions exceeds the number of samples. However, it may be more successful for huge datasets with a high level of noise or overlapping classes. Additionally, training time can be lengthy when dealing with huge datasets.

## 4. Methodology

The proposed methodology consists of four stages: data collection and annotation, preprocessing, feature engineering, and classification. Figure 1 gives the graphical representation of proposed methodology.

**4.1. Data Collection and Annotation.** Data are the main component in every research. In our work, data are extracted from online social network platforms using various methods such as APIs and crawlers. We extracted data from Twitter, a social networking site using its REST API. Data are extracted based on some events using various hashtags.

**4.1.1. Data Collection.** The most important thing in every research is data. Since propagandistic text is of semantic in nature, the first thing is to perform keyword-based extraction. We used Twitter Application Program Interface (API) for the extraction of tweets from Twitter [48]. We identified five ambiguous keywords that were mostly used in COVID-19 times. We extracted using #CORONAJIHAD, #CORONATERRORSIM, #CORONAMUSLIM, #CHINESEVIRUS, and #BIOWEAPON. About 50 K tweets were extracted using these keywords.

**4.1.2. Annotation.** Due to propagandistic tweets semantic nature, we preferred manual annotation over crowd-sourcing like Monkey Learn. Tweets were labelled into two classes based on various propaganda techniques: Name Calling, Bandwagon, Testimonial, and Glittering Generality. Two journalists labelled about 5 k tweets. Figure 2 shows the length of annotated tweets in propaganda and non-propaganda class.

**4.2. Data Preprocessing.** Preprocessing plays a vital role in text classification problems. Since there are too much unnecessary spaces, links, etc., we need to clean the data. For this task, various preprocessing tasks are used. Various preprocessing techniques that are used for refining the text are as follows [49].

**4.2.1. Tokenization.** In this task, the tweets are divided into tokens such that we can easily refine the text by removing unwanted words and digits.

**4.2.2. Punctuation Removal.** In this phase, the punctuations such as comma, full stop, and semicolon are removed. Punctuation is mostly used by humans to communicate with each other.

**4.2.3. Number Removal.** Various numbers are posted in a tweets, and they does not have any sentiment. In this phase, the numbers are removed as numbers have the least preference in text analysis [50].

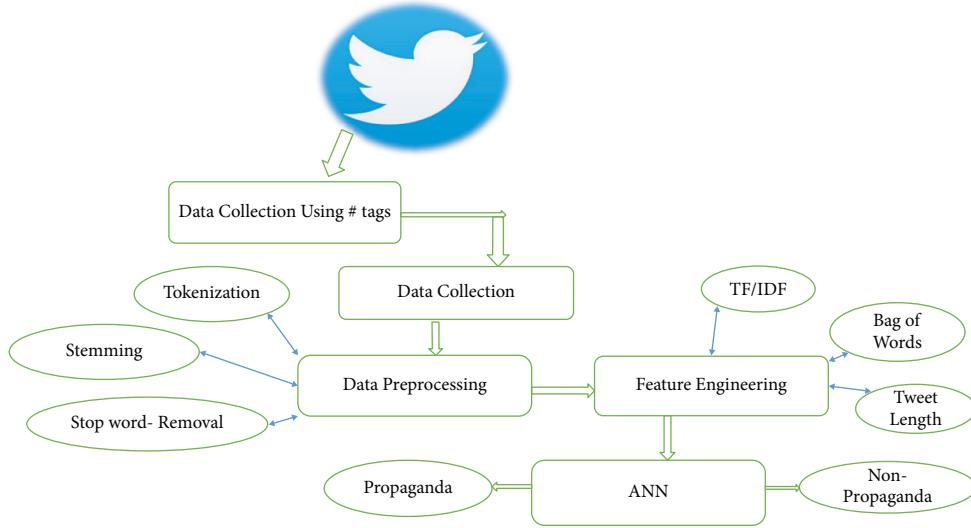


FIGURE 1: The proposed methodology.

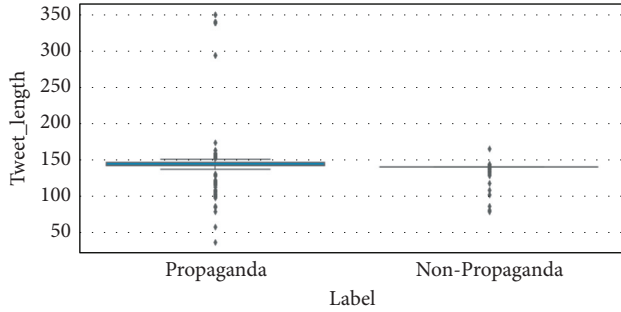


FIGURE 2: Propaganda and nonpropaganda class with the lengths.

**4.2.4. Stop Words.** This task removes all the stop words as they have a very least role in a classification task. We used the English stop words dictionary to remove the stop words as we have worked on tweets written in English only.

**4.2.5. Spell Checking.** Humans have the nature of making mistakes whether grammatical or spelling while writing a post/tweet. In this phase, various misspelled words are corrected by using “pyspellchecker,” a python library.

**4.2.6. Stemming.** Each word is stemmed to its root word such that the actual meaning of the word can be generated. It shaves off the ends of words with the expectation of accomplishing this goal correctly the majority of the time and frequently involves the omission of derivational affixes.

**4.3. Feature Engineering.** For every machine learning algorithm, we have to choose features on which we can train algorithms. We used different feature engineering techniques for selecting the most relevant features. In this work, we combined various methods to get the most relevant features. The methods that are used in this work are as follows.

**4.3.1. TF/IDF.** Term Frequency/Inverse Document Frequency mirrors a word’s significance in a tweet or over an entire corpus by providing quantitative insights. It is determined utilizing the accompanying condition.

$$\text{TFIDF}(t, w, D) = \text{TF}(t, w) * \text{IDF}(t, D),$$

$$\text{IDF}(t, D) = \log \frac{|D|}{1 + |(w \in D: t \in w)|} \quad (12)$$

where  $t$  denotes the word as a component,  $w$  denotes each tweet in the corpus, and  $D$  denotes dataset’s total number of tweets (document space).

**4.3.2. Bag of Words.** The process of extracting information from text for use in machine learning algorithms is referred to as text mining. Bag of Words is a representation of text that describes the presence of words inside a document’s textual information. To extract more information from the text, we used bi-grams and tri-gram words in this work.

**4.3.3. Length.** Another feature due to which accuracy got improved is the length of the tweet. As Tweet length is at most 280 characters, it is too difficult to detect the propaganda in this short message. We take this as a feature for our classifier such that it will perform better.

We combined all the above selected features out of which we choose the most relevant 40 features by using Gini coefficient.

**4.4. Classification.** The classification process is carried out in order to group the given substance into two distinct types of groups. The two classes are propaganda (a tweet which is propagandistic) and nonpropaganda (normal tweet). Artificial neural network algorithm is used to mastermind the substance into these classes.

**Require:** filtered tweets ( $T_{input}$ )  
**Ensure:** propaganda tweet ( $T_{Pr}$ ) and non-propaganda tweet ( $T_{Np}$ )

- (1) Filtered tweets  $\rightarrow T_{input}$
- (2) Propaganda tweets  $\rightarrow T_{Pr}$
- (3) Nonpropaganda tweets  $\rightarrow T_{Np}$
- (4) Tokenization  $\rightarrow T$
- (5) Stop word removal  $\rightarrow SW$
- (6) Stemming  $\rightarrow S$
- (7) Total Number of Tweets  $\rightarrow n$
- (8) Term frequency/inverse document frequency  $\rightarrow TF/IDF$ .
- (9) Bag of Words  $\rightarrow B$
- (10) **START**
- (11) **for**  $i$  from 1 to  $n$  **do**
- (12)  $C[i] = T_{input}[i] + \text{Label}$  //manual annotation
- (13)  $T[i] = \text{Tweetlength}(C[i])$
- (14) **end for**
- (15) **for**  $i$  from 1 to  $n$  **do**
- (16)  $Pro[i] = T(C[i])$
- (17)  $Pro[i] = SW(Pro[i])$
- (18)  $Pro[i] = S(Pro[i])$
- (19) **end for**
- (20) **for**  $i$  from 1 to  $n$  **do**
- (21)  $F[i] = B(TF/IDF(Pr[i]))$
- (22)  $F[i] = F[i] + T[i]$
- (23) **end for**
- (24)  $ANN(F[i])$  (CLASSIFIER)
- (25) **END**

ALGORITHM 1: Improved ANN for the classification of tweets (ANNBPI).

The multilayer perceptron classifier (MLPC) is a classification algorithm that is dependent on the framework. MLPC is utilizing backpropagation for learning [51]. The amount of hubs in the yield layer identifies with the number of classes. MLPC is comprised of several hubs, including the data layer, veiled layers (also known as centre layers), and yield layers, among others. A connection between each layer and the layer that it is connected to in the framework is established, where the data layer, the cover layer, and the yield layer can all be represented. Neurons in the data/input layer are responsible for perceiving information. The yield from these neurons is the same as the data pointers. Hubs in the data layer indicate the information. Every hub map commitments to yield by a straight blend of commitments with gesture's loads  $w$  and tendency  $b$  and applying for actuation work. The cross-sectional structure for MLPC with  $K + 1$  layers is composed. Hidden/middle layers are between the information and output layers. Usually, the number of concealed layers ranges from one to a significant number. The capabilities of the focused computing layer are responsible for mapping a node's contribution to its output. The sigmoid (strategic) work is utilized by the nodes in the middle levels.

The output layer is the final layer of a neural network, and it is responsible for returning the outcome to the original condition. With regard to the topology of a neural network, it also flags the previous layers based on how they have behaved in learning the input and enhancing their functions. The softmax function is used by nodes in the output layer.

TABLE 2: Classification report of algorithms.

Algorithm	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
LR	65	68	60	68.2
MNB	69	69	59	68.6
SVM	78	77	74	76.8
DT	68	70	63	69.65
ANNBPI	79	77	74	77.15

The output layer's node count  $N$  equals the number of classes (see Algorithm 1).

## 5. Results and Discussion

The experimentation is performed on workstation having configuration of 16 GB RAM and GPU. Dataset is split into many ratios. First, data were split into 50 : 50; that is, 50% of data was used for training and 50% for testing. It was found that the accuracy was less as compared to when data were split into 60 : 40. After performing classification on these datasets, it was concluded that the dataset. When data are split into 80 : 20, which means that 80% of data are used for training and 20% for testing, the best accuracy is achieved. The testing of the models was measured based on precision, recall, F-measure, and accuracy using equations (14), (15), and (16):

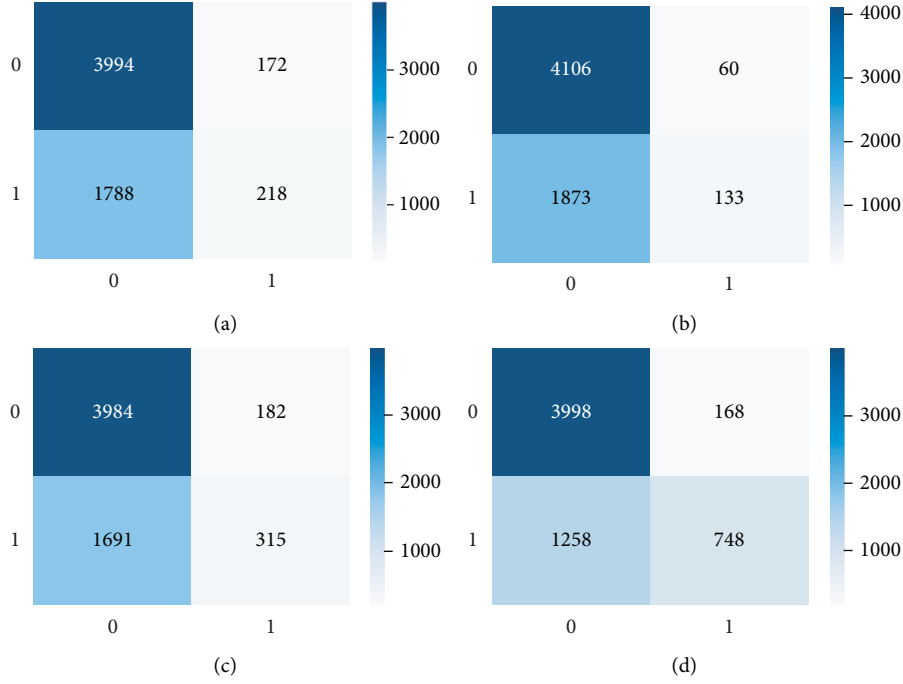


FIGURE 3: Confusion matrices of traditional machine learning algorithms.

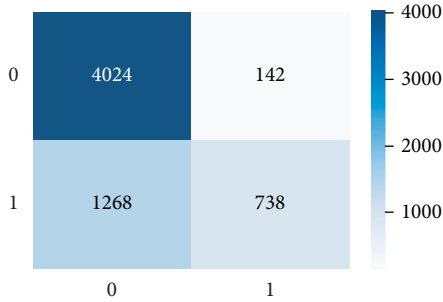


FIGURE 4: Confusion matrix of ANNBPI.

$$P = \frac{Tp}{(Tp + Fp)}, \quad (13)$$

where  $Tp$  is the true positive, and  $Fp$  is the false positive.

$$P = \frac{Tp}{(Tp + Fp)}, \quad (14)$$

where  $Tp$  is the true positive, and  $Fp$  is the false positive.

$$R = \frac{Tp}{(Tp + Fn)}, \quad (15)$$

where  $Tp$  is the true positive, and  $Fp$  is the false negative.

$$F = \frac{(2 * P * R)}{(P + R)}. \quad (16)$$

While training and testing the models, the ANN over performed other algorithms by showing accuracy of 77.15%. The objective of the factual examination is to gather as much information as possible about the data that will be used in the

TABLE 3: Comparison with existing work.

Algorithm	Precision (%)	Recall (%)	F1-score (%)	Accuracy
LR	64	67	59	67.8%
MNB	67	67	58	68.2%
SVM	76	75	72	74.8%
DT	67	69	62	68.65%
ANNBPI	78	76	75	76.15%
[29]	56.54	47.37	51.55	NA
[30]	59.95	41.65	49.15	NA
[31]	53.23	45.56	49.10	NA

classification task. The preceding work was completed using the Python programming language and the Spyder IDE. The data are presented in the form of a table with different qualities for the 40 attributes that were selected. The provided task is being carried out using several libraries. Pandas, scikit-learn, and NLTK, for example, are among the pre-owned libraries. The comparison of different methods for conducting binary classification is shown in Table 2. Figure 3 illustrates the confusion matrices of the conventional machine learning methods, whereas Figure 4 shows the confusion matrix of ANNBPI.

**5.1. Comparative Study.** For validating, the results generated using proposed approach with the existing work. The similar dataset, as used by other researchers, was experimented based on proposed methodology. After fine-tuning logistic regression, decision tree, support vector machine, multinomial Naive Bayes, and artificial neural network based on the proposed methodology, it was found that these algorithms performed better than other previous studies. The evaluation metrics showed that ANNBPI achieved highest



performance on the said dataset by having precision of 78%, recall of 76%, F1-score of 75%, and accuracy of 76.15%. Table 3 shows the comparison of the proposed work with existing approaches in this domain.

## 6. Conclusion

Propaganda is used as a lethal weapon on online social networks. Using Twitter's Application Program Interface, data were retrieved from one of the social media platform "Twitter." The data were labelled manually using various propaganda identification techniques. Hybrid feature engineering is being performed, which includes Bag of Words, TF/IDF, and Tweet length. About 40 most relevant features are chosen using these techniques. Machine learning classifiers are trained and tested on the labelled dataset by splitting it into 70 : 30 ratio. The results showed that artificial neural network showed better results as compared to other algorithms. ANN gave 77.15% accuracy, 74% *F*-measure, 79% precision, and 77% recall. In the future, more data may be supplied to these algorithms and there are also future directions in deep learning.

## Data Availability

The data used to support this study are available from the author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] M. Glenski, T. Weninger, and S. Volkova, "Propagation from deceptive news sources who shares, how much, how evenly, and how quickly?" *IEEE Transactions on Computational Social Systems*, vol. 5, no. 4, pp. 1071–1082, 2018.
- [2] J. Zhu, P. Ni, and G. Wang, "Activity minimization of misinformation influence in online social networks," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 4, pp. 897–906, 2020.
- [3] S. Alhazbi, "Behavior-based machine learning approaches to identify state-sponsored trolls on twitter," *IEEE Access*, vol. 8, pp. 195132–195141, 2020.
- [4] R. F. Sear, N. Velasquez, R. Leahy et al., "Quantifying COVID-19 content in the online health opinion war using machine learning," *IEEE Access*, vol. 8, pp. 91886–91893, 2020.
- [5] T. Chen, W. Liu, Q. Fang, J. Guo, D. Z. Du, and D. Du, "Minimizing misinformation profit in social networks," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 6, pp. 1206–1218, 2019.
- [6] S. Kausar, B. Tahir, and M. A. Mehmood, "ProSOUL: a framework to identify propaganda from online Urdu content," *IEEE Access*, vol. 8, pp. 186039–186054, 2020.
- [7] A. Badawy and E. Ferrara, "Rise of Jihadist Propaganda on social networks," *Journal of Computational Social Science*, vol. 1, no. 2, pp. 453–470, 2018.
- [8] B. Han, R. H. Jhaveri, H. Wang, D. Qiao, and J. Du, "Application of robust zero-watermarking scheme based on federated learning for securing the healthcare data," *IEEE Journal of Biomedical and Health Informatics*, 2021.
- [9] P. N. Srinivasu, A. K. Bhoi, R. Jhaveri, G. T. Reddy, and M. Bilal, "Probabilistic Deep Q network for real-time path planning in censorious robotic procedures using force sensors," *Journal of Real-Time Image Processing*, Springer, vol. 18, no. 5, pp. 1–13, 2021.
- [10] T. Fernando, C. Fookes, S. Denman, and S. Sridharan, "Detection of fake and fraudulent faces via neural memory networks," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1973–1988, 2021.
- [11] T. M. Khoshgoftaar, J. Van Hulse, and A. Napolitano, "Supervised neural network modeling: an empirical investigation into learning from imbalanced data with labeling errors," *IEEE Transactions on Neural Networks*, vol. 21, no. 5, pp. 813–830, May 2010.
- [12] M. Li and F. Li, "Survey on lie group machine learning," *Big Data Mining and Analytics*, vol. 3, no. 4, pp. 235–258, Dec. 2020.
- [13] R. Kumar, G. Dhiman, V. Joshi, R. H. Jhaveri, and A. K. bhoi, "The combined study of improved fuzzy optimization techniques with the analysis of the upgraded facility location center for the covid-19 vaccine by fuzzy clustering algorithms," *International Journal of Nanotechnology, Inderscience*, 2021.
- [14] H. Gao, J. Hu, B. Y. Zhao, S. Barbara, U. C. S. Barbara, and Y. Chen, "Detecting and characterizing social spam campaigns categories and subject descriptors," in *Proceedings of the 17th ACM conference on Computer and communications security*, pp. 681–683, Chicago, IL, USA, October 2010.
- [15] J. Ratkiewicz, M. Conover, M. Meiss et al., "Truthy: mapping the spread of astroturf in microblog streams," in *Proceedings of the 20th International Conference Companion on World Wide Web*, pp. 249–252, ACM, March 2011.
- [16] A. Halu, K. Zhao, A. Baronchelli, and G. Bianconi, "Connect and win: the role of social networks in political elections," *EPL*, vol. 102, no. 1, 2013.
- [17] N. Ramakrishnan, V. Tech, F. Chen, J. Arredondo, D. Mares, and K. Summers, "Analyzing Civil unrest through social media," *Computer (Long. Beach. Calif)*, vol. 46, no. 12, pp. 80–84, 2013.
- [18] J. S. Liu, K. C. Ning, and W. C. Chuang, "Discovering and characterizing political elite cliques with evolutionary community detection," *Social Network Analysis and Mining*, vol. 3, no. 3, pp. 761–783, 2013.
- [19] S. Y. Bhat and M. Abulaish, "Using communities against deception in online social networks," *Computer Fraud & Security*, vol. 2014, no. 2, pp. 8–16, 2014.
- [20] A. M. U. D. Khanday, S. T. Rabani, Q. R. Khan, N. Rouf, and M. Mohi Ud Din, "Machine learning based approaches for detecting COVID-19 using clinical text data," *International Journal of Information Technology*, vol. 12, no. 3, pp. 731–739, 2020.
- [21] F. M. F. Wong, C. W. Tan, S. Sen, and M. Chiang, "Quantifying political leaning from tweets, retweets, and retweeters," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 8, pp. 2158–2172, 2016.
- [22] F. Zarrinkalam and E. Bagheri, "Event identification in social networks," *Encyclopedia with Semantic Computing and Robotic Intelligence*, vol. 01, no. 01, Article ID 1630002, 2017.
- [23] S. Lightfoot, "Political Propaganda Spread through Social Bots," *Media, Culture, Global Politics*, vol. 8, pp. 1–22, 2018.
- [24] M. Ashcroft, A. Fisher, L. Kaati, E. Omer, and N. Prucha, "Detecting jihadist messages on twitter," in *Proceedings of the 2015 Eur. Intell. Secur. Informatics Conf. EISIC*, pp. 161–164, Manchester, UK, September 2015.

- [25] K. K. Kumar and G. Geethakumari, "Detecting misinformation in online social networks using cognitive psychology," *Human-centric Comput.*, *Information Science*, vol. 4, no. 1, pp. 1–22, 2014.
- [26] G. Mazzoleni and R. Bracciale, "Socially Mediated Populism: The Communicative Strategies of Political Leaders on Facebook," *Palgrave Communications*, vol. 4, no. 1, 2018.
- [27] M. Mujahid, E. Lee, F. Rustam et al., "Sentiment analysis and topic modeling on tweets about online education during COVID-19," *Applied Sciences*, vol. 11, no. 18, p. 8438, 2021.
- [28] I. Ashraf, W. S. Alnumay, R. Ali, S. Hur, A. Kashif Bashir, and Y. Bin Zikria, "Prediction models for covid-19 integrating age groups, gender, and underlying conditions," *Computers, Materials & Continua*, vol. 67, no. 3, pp. 3009–3044, 2021.
- [29] G. Morio, T. Morishita, H. Ozaki, and T. Miyoshi, "Hitachi at SemEval-2020 Task 11: an empirical study of pretrained transformer family for propaganda detection," in *Proceedings of the 14th International Workshop on Semantic Evaluation*, December 2020.
- [30] D. Jurkiewicz, L. Borchmann, I. Kosmala, and F. Gralinski, "ApplicaAI at SemEval-2020 task 11: on RoBERTaCRF, span CLS and whether self-training helps them," in *Proceedings of the 14th International Workshop on Semantic Evaluation, SemEval '20*, Barcelona, Spain, May 2020.
- [31] A. Chernyavskiy, D. Ilvovsky, and P. Nakov, "Aschern at SemEval-2020 Task 11: it takes three to tango: RoBERTa, CRF, and transfer learning," in *Proceedings of the 14th International Workshop on Semantic Evaluation*, August 2020.
- [32] S. Khosla, R. Joshi, R. Dutt, A. W. Black, and Y. Tsvetkov, "LTIatCMU at SemEval-2020 Task 11: incorporating multi-level features for multi-granular propaganda span identification," in *Proceedings of the 14th International Workshop on Semantic Evaluation*, August 2020.
- [33] A. Paraschiv and D.-C. Cercel, "UPB at SemEval-2020 Task 11: propaganda detection with domain-specific trained BERT," in *Proceedings of the 14th International Workshop on Semantic Evaluation*, July 2020.
- [34] I. Dimov, V. Korzun, and I. Smurov, "NoPropaganda at SemEval-2020 Task 11: a borrowed approach to sequence tagging and text classification," in *Proceedings of the 14th International Workshop on Semantic Evaluation*, January 2020.
- [35] V. Blaschke, M. Korniyenko, and T. Sam, "CyberWalle at SemEval-2020 Task 11: an analysis of feature engineering for ensemble models for propaganda detection," in *Proceedings of the 14th International Workshop on Semantic Evaluation*, January 2020.
- [36] E. Verma, V. Motupalli, and S. Chakraborty, "Transformers at SemEval-2020 Task 11: Propaganda fragment detection using diversified bert architectures based ensemble learning," in *Proceedings of the 14th International Workshop on Semantic Evaluation*, December 2020.
- [37] P. Singh, S. Sandhu, S. Kumar, and A. Modi, "newsSweeper at SemEval-2020 Task 11: context-aware rich feature representations for propaganda classification," in *Proceedings of the 14th International Workshop on Semantic Evaluation*, July 2020.
- [38] V. Ermurachi and D. Gifu, "UAIC1860 at SemEval-2020 Task 11: detection of propaganda techniques in news articles," in *Proceedings of the 14th International Workshop on Semantic Evaluation*, December 2020.
- [39] D. S. Dewantara, I. Budi, and M. O. Ibrahim, "3218IR at SemEval-2020 Task 11: Conv1D and word embedding in propaganda span identification at news articles," in *Proceedings of the 14th International Workshop on Semantic Evaluation*, January 2020.
- [40] G. Daval-Frerot and W. Yannick, "WMD at SemEval- 2020 Tasks 7 and 11: assessing humor and propaganda using unsupervised data augmentation," in *Proceedings of the 14th International Workshop on Semantic Evaluation*, December 2020.
- [41] J. G. Carbonell, R. S. Michalski, and T. M. Mitchell, "An overview of machine learning," in *Machine Learning, Symbolic Computation*, R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, Eds., Springer, Berlin, Germany, pp. 3–23, 1983.
- [42] F. Rustam and A. A. Reshi, I. Ashraf, A. Mehmood, S. Ullah, D. M. Khan, and G. S. Choi, "Sensor-based human activity recognition using Deep stacked multilayered Perceptron model," *IEEE Access*, vol. 8, pp. 218898–218910, 2020.
- [43] M. Umer, I. Ashraf, S. Ullah, A. Mehmood, and G. S. Choi, "COVINet: a convolutional neural network approach for predicting COVID-19 from chest X-ray images," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, no. 1, pp. 535–547, 2022.
- [44] M. Khalid, I. Ashraf, A. Mehmood, S. Ullah, M. Ahmad, and G. S. Choi, "GBSVM: sentiment classification from unstructured reviews using ensemble classifier," *Applied Sciences*, vol. 10, no. 8, p. 2788, 2020.
- [45] "Description of logistic regression algorithm," San Francisco, California, CA, USA, 2021, <https://machinelearningmastery.com/logistic-regression-for-machinelearning/>.
- [46] "Description of multinomial naive bayes algorithm," 2021, <https://www.3pillarglobal.com/insights/document-classification-usingmultinomial-naive-bayes-classifier>.
- [47] A. M. U. D. Khanday, Q. R. Khan, and S. T. Rabani, "SVMBPI: support vector machine-based propaganda identification," *Cognitive Informatics and Soft Computing: Proceeding of CISC*, p. 445, Springer Nature, Berlin, Germany, 2021.
- [48] P. Verma, A. M. U. D. Khanday, S. T. Rabani, M. H. Mir, and S. Jamwal, "Twitter sentiment analysis on Indian government project using R," *International Journal of Recent Technology and Engineering*, vol. 8, no. 3, pp. 8338–8341, 2019.
- [49] V. Rupapara, F. Rustam, H. F. Shahzad, A. Mehmood, I. Ashraf, and G. S. Choi, "Impact of SMOTE on imbalanced text features for toxic comments classification using RVVC model," *IEEE Access*, vol. 9, pp. 78621–78634, 2021.
- [50] R. Jamil, I. Ashraf, F. Rustam, E. Saad, A. Mehmood, and G. S. Choi, "Detecting sarcasm in multi-domain datasets using convolutional neural networks and long short term memory network model," *PeerJ Computer Science*, vol. 7, p. e645, 2021.
- [51] J. Yao, K. Wang, and J. Yan, "Incorporating label Co-occurrence into neural network-based models for multi-label text classification," *IEEE Access*, vol. 7, pp. 183580–183588, 2019.