Wireless Personal Communications
https://doi.org/10.1007/s11277-021-08720-9

Certain Investigation of Fake News Detection from Facebook and Twitter Using Artificial Intelligence Approach

Roy Setiawan[1] · Vidya Sagar Ponnam[2] · Sudhakar Sengan[3] · Mamoona Anam[4] · Chidambaram Subbiah[5] · Khongdet Phasinam[6] · Manikandan Vairaven[7] · Selvakumar Ponnusamy[7]

Accepted: 29 June 2021

Abstract

The news platform has moved from traditional newspapers to online communities in the technologically advanced area of Artificial Intelligence. Because Twitter and Facebook allow us to consume news much faster and with less restricted editing, false information continues to spread at an impressive rate and volume. Online Fake News Detection is a promising field in research and captivates the attention of researchers. The sprawl of huge chunks of misinformation in social network platforms is vulnerable to global risk. This article recommends using a Machine Learning optimization technique for automated news article classification on Facebook and Twitter. The emergence of the research is facilitated by the strategic implementation of Natural Language Processing for social forum fake news findings in order to distort news reports from non-recurrent outlets. The relent from the study is outstanding with text document frequency words, which act as extraction technique's attribute, and the classifier is acted upon by Hybrid Support Vector Machine by achieving 91.23% accuracy.

Keywords NLP · Hybrid SVM · Machine Learning · Fake News

1 Introduction

The considerable cost of digital time in surfing social networking sites prone people to throng social media for news items rather than mainstream news agencies. At present, FND investigates the understanding of people to express an idea on social media. The Portrayal of this idea is known as Stance Classification, a Natural Language Preprocessing (NLP) [1] task that tries to categorize the stance headed towards a few claims. The primary function of NLP, a research-oriented area, is preprocessing human language with the help of language models and computational approaches such as Machine Learning [2]. The thriving of ML tools and techniques unfold different ways of designing the algorithm for stance

✉ Sudhakar Sengan
sudhasengan@gmail.com

Extended author information available on the last page of the article

Published online: 07 July 2021

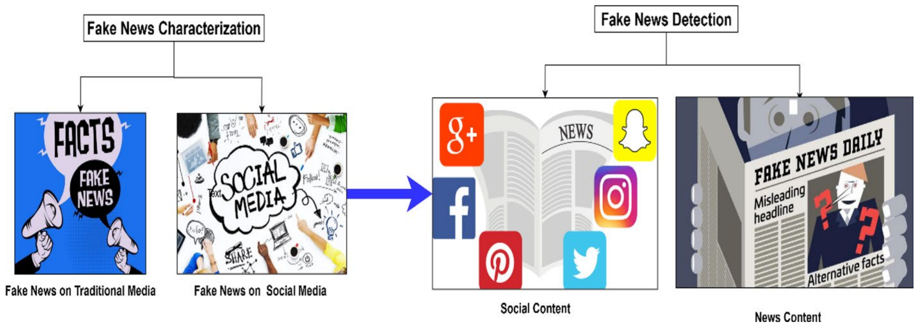

Fig. 1 Fake news on facebook social media

classification. A fruitful investigation of this progress and the acquisition of insightful knowledge of recent avant-garde approaches is quite motivating.

Fake news and failure of trust in media are severe issues in our culture. This paper has proposed a model that would find out the genuineness of the news. The word 'Fake News' [3] is perceived as a group discussion, especially to elucidate the inaccurate articles. The definition of fake news is represented in Fig. 1 as authenticity and intent. The word 'Authenticity' means that fake news can be checked as such and 'Internet' implies fake news to mislead the readers. The justification for the change of user habits is given besides such social networking: (a) time-consuming and economical, getting accessibility to social network content comparing to mainstream news sources such as newspaper or television and (b) further reflecting the same by discussing the news items with acquaintances and other social media users [4]. There is no feasibility of manual FND in a considerable volume of online text contents. Hence, a Smart tool or system that skillfully performs automatic FND is insisted by reliable agencies. The identification of security threats in the cyber-world transmitted by text contents is easier. But inadequate resources and standard ML text datasets are the significant obstacles for a suspicious text detection system to be created. Comparing to other automated computer languages, the execution is challenging.

To find a solution for this research question, a dataset of fake and non-fake news was developed, considering numerous renowned natural language data sources like Facebook and Twitter. For the sake of textual data processing, the characteristics of unigram, bigram, trigram with the help of Term Frequency (TF) and Term Frequency-Inverse Document Frequency (TF-IDF) [5] and a bag of words featuring extraction technique are considered. After the completion of feature extraction, we use the eminent ML classifiers to find out if a given text is reliable or not. A comparative analysis of ML models was also done by availing our dataset collection. The primary motive of this proposed work is to avoid the distribution of false information through popular platforms like Facebook and Twitter. This will provide great promotions towards the public and affect various business processes, even in political elections. Sometimes it creates emotions among the people. So, AI tools create a wide variety of technology to resolve these issues in popular platforms—this research aims to investigate the efficiency and weaknesses of language-based techniques for fake news detection. The accomplishment level of the task is decided upon by the outcome of the paper.

Moreover, this is not proposed to be a back-to-back solution for the classification of fake news but to help people detect fake news. This smart tool combines multiple tools and could be used in future applications. This paper explains the ML classifier for FND.

2 Related Works

The term "reality" presents [6] common knowledge resulting from a social experiences network. Discovering reality using Investigative journalism is a costly way as it involves vast online content. "Fake News" is a term that emerges from the core of social networks. Similar to System of Systems, online social networks activate emerging properties resulting in complex authentication processes, says the researcher.

The authentication ofthe news item in social media is very crucial from an ethical point of view. The results of the web-scale restrict the researchers from spotting and rectifying fake news. The proposed work devised an automatic fake news detector on Twitter [7]. We wrap up with a discussion that compares validity and integrity and why non-expert models outsmart journalists' models on Twitter to identify fake news. A walkthrough of the so far discussion reveals that the author of this research paper explores the structural problem of fake news escalation in social media platforms against factual news.

The terrifying growth of fake news in online media motivated the researchers to discover how the latest social networks like maximization effects, knowledge dissemination, and epidemiological models paved the way for creating and spreading fake news and the author's recommendation for future research [8].

This work [9] thoroughly studies the recent developments that expose and establish fake news identification techniques. The role of "Fake News Identification" is to classify news items using a veracity scale, which has some degree of certainty. The presence of intentional deception causes shortfalls in veracity. The availability of bountiful user-generated documents besides Computer-Mediated Communication technology such as forums, Twitter, and other social media has excellent prospects for mass-scale news distribution mechanisms. Maintenance of knowledge-based online tasks is very challenging, yet it is essential. The four long decades of extensive research to identify fake information facilitated the finding of how frequently lies are detected by humans using a meta-analysis conducted on more than 200 studies, just 3.97% better than chance.

Altmetric and qualitative data are synchronous to standard and citation-based metrics: the Altmetric and qualitative data for a Research Output [10, 11]. The score is calculated with an automated algorithm, which decides the weighted count of the medication amount estimated for research output. The output of this research is an Altmetric attribution score of 9500 for the article. Altmetric shows a tracking history of 13,600,200 research outputs till 2020 and, on the whole, 8,50,600 research outputs from all the sources. This Altmetric Attention Score can be compared with 290,000 tracked outcast published in 6 weeks. This research was published to 2500 others from a similar origin and in 6 weeks on either side. At the same time, conventional citation-based metric follows the Altmetric metric and qualitative data.

This paper intends to examine the disparity in the language of fake news and satirical stories. A profound linguistic analysis is not the aim of this study. This study utilized a dataset of 300 fake news articles and 200 satirical tales. Preliminary results recommended that the viability of tracing the theme relies on its word vector. Fetching additional information based on the prediction model is quite fascinating [12].

A volunteer group from industry and academia organized a “Fake News Challenge” in December 2019. This competition’s criteria are to develop fact-checker tools for supporting people to detect intentionally fictitious information in news items. The teams that won the first three positions achieved 80% accuracy in the task set for them. The model was constructed based on a weighted average between gradient-boosted decision trees and a deep CNN [13]. A method proposed to detect fake news in social media uses an ensemble approach to automatically classify text from new articles through different textual properties [14]. It was evaluated by using real-world datasets. Fake news detection multi-task learning is designed for detecting fake news having a high percentage and the authors who are having high intention to spread fake news [15]. It consists of multi-tasking and representation learning on fake news and classifies whether it is a new topic or integrating two or more tasks. To characterize more than hundreds of fake news, fake websites and publishers of real news are identified through domain reputation and understanding of contents [16]. The performance of fake news detection is evaluated by a novel stacking method with various ML and DL models and document frequency and embedding techniques for obtaining text representation. It was evaluated through precision, accuracy, F1-score, and recall [17] [18] [19]. A new methodology is proposed to address the employing dimensionality reduction methods to reduce passing fake news to the classifier by using a hybrid neural network like CNN-LSTM with dimensionality reduction through Chi-square and Principle component analysis (PCA).

Rumour classification [20] is analogous to fake news in which the integrity of information flow is yet to be checked at the time of diffusion. For example, on Twitter, fake news articles are devised deliberately and denounced as fake on verification. The tweet contains a claim and many responses, and then the task is to find whether the claim is True/Fake.

The author explains three kinds of fake news. Each one of them represents inaccurate or fake reporting. He has also analyzed the merits and demerits of each fake news using diverse text analytics. (a) Fabrication of serious news is not available in conventional or participant media. (b) Hoaxes on a massive scale are inventive and exclusive and frequently appear on several platforms. (c) The features of this kind of fake news style could diminish the efficiency of text classification techniques.

3 Proposed Methodology

3.1 Natural Language Processing

NLP is allied to Artificial Intelligence, which deals mainly with humans’ computers and languages (natural) [21]. The purpose of constructing a machine-based ML model on text data is to clean and convert text-oriented data into a machine-readable format. The era of pre-neural network NLP predominantly aims at improving domain-specific traits. This research presented exposure to learning words and representation of sentence-level rather than utilizing human-made input attributes. It educates on how to represent lexis and sentences as vectors to comprehend the situation in which it is used. It applies the integrated neural network architecture and algorithms for the various functions of NLP [22]. NLP creates more consideration towards automatically detected fake news in social media or the essential foundation of various tasks circulating throughout the world [23].

The news articles from various sources are stored in a repository, including their mark like True/Fake. This dataset has already been categorized as a Train and Test dataset consisting of the following attributes: *"Statements"* and *"Name"* info. The step-by-step process to forecast whether the news statement is factual or not is given below:

- Data Pre-processing
- Extraction of Features
- Machine Learning Training Dataset and Classification

Data Preprocessing is the storage of news statements, and the essential terms are regained using NLP. Translating the words to vectors is performed by function extraction techniques associated with Count Vectorizer and TF-IDF converter [24, 25]. Hybrid SVM has been tailored to train with the help of vector representation of news statements. The exact process is performed by test dataset after fixing Hybrid SVM with trained function vectors, and the subsequent feature vectors Hybrid SVM would distinguish test data into True/Fake claims. This design is optimized for the user-like Interface utilizing the inter-module Python *tk* [26].

The block diagram for fake news detection is represented in Fig. 2 because the dataset, qualified data, and reliable data were fragmented. Such statistical data were the collections of several press accounts.

3.1.1 Pre-Processing

Pre-processing signifies the fundamental transformation to the earlier data, feeding it into the model. Data Preprocessing is used to change unrefined historical data into a refined data set. The sci-kit-learn preprocessing is intact with a versatile sci-kit-learn library in Python. More options in pre-processing will be explored. NLP develops software that deciphers human languages. The Python NLTK (Natural Language Toolkit) is incorporated in this NLP tutorial. NLP.x makes use of NLTK, a benchmark library in Python [27, 28].

3.1.2 Extraction of Features

Text categorization faces the challenge of imbibing data of high dimensions. A pool of terminologies, lexis, and word groups in documents causes an increased computational burden in the process of learning. Moreover, precision and performance are affected by inappropriate and redundant features. Hence, executing a feature reduction for lessening the size of text features and prohibiting features of huge dimensions is the ideal way. Term Frequency (TF) [29] and Term Frequency-Inverted Document Frequency (TF-IDF) [30] are the two selection methods featured by this outcome-based research, and it is explained in the below section.

a. Term Frequency (TF)

    The TF approach avails the appearance of word counts in the documents to derive the analogy between documents. The characterization of each document is performed using an equal length vector, which embraces the word counts. After that, every single vector is brought to normalcy to be added to the totality of elements. Consequently, the conversion of each word count into the possibility of such a word exists in the documents. For instance, a word will be represented as one when it is present in a specific document, and it will be fixed as ‘0’ when the document has no word. Hence, a group of words is represented in each document.

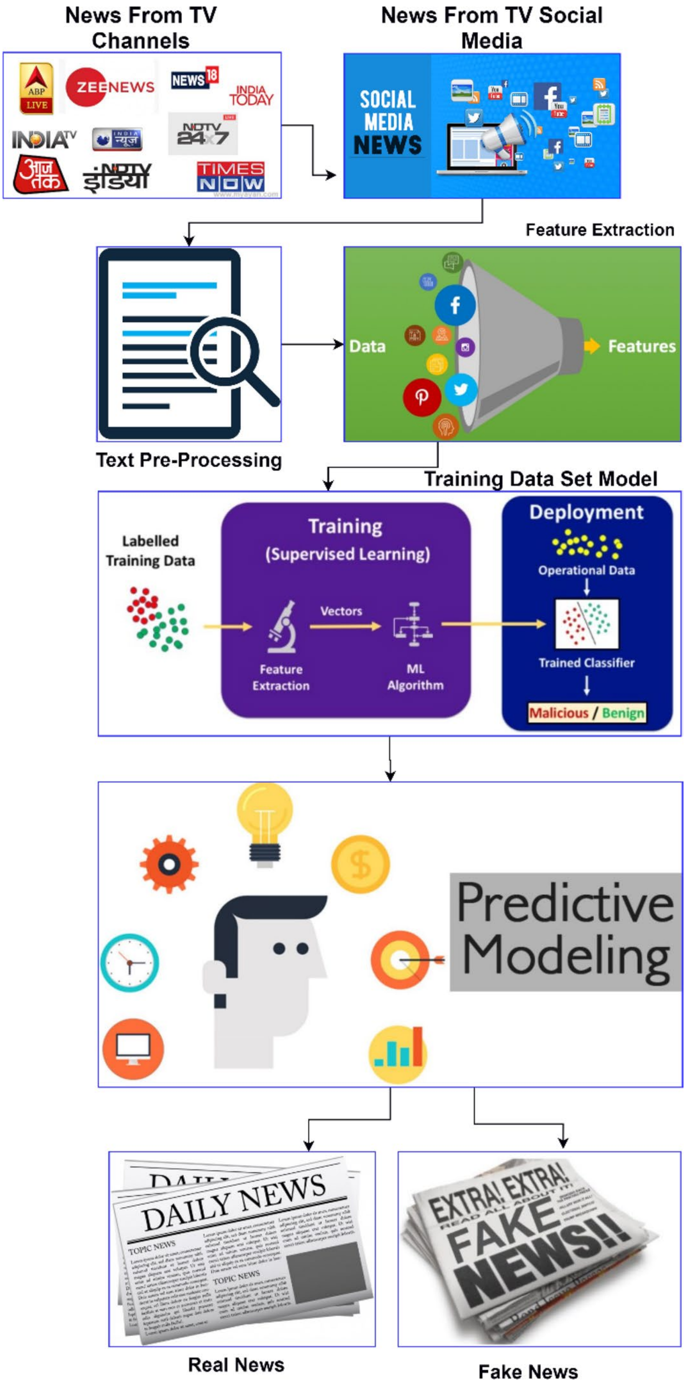

Fig. 2 Proposed Fake News Detection System

b. Term Frequency-Inverted Document Frequency (TF-IDF)

Information retrieval and NLP frequently use TF-IDF, a weighting metric. It's a statistical metric that determines the term's significance to a document in a dataset. The frequency of word appearance in the document depends on the increase of terms. But, the frequency of word appearance in the corpus counteracts this. One of IDF's significant features is the weighting down of the term frequency during the scaling up of unique ones. An illustration is the frequent appearance of words like "the" and "then," and if only TF is used, these terms will be controlled by the frequency count. But, the effect of these terms is minimized by IDF.

c. Fake News Detection

FND performs classification of a text fn_i€FN from a set of texts FN = {$fn_{1,}$ $fn_{2,}$ $fn_{3\ldots}$ fn_m} into a class c_i€C from a set of two classes C = {C_s,C_{ns}}, FND automatically assign fn_i to ci: < fn_i, c_i > .

Most cited definitions spotlight common topics like provoking violence, inducing hatred and terrorism, and intimidating an individual/group (Table 1). These definitions cover the practical aspects of fake substance from video, image, text, cartoon, graphics, and illustrations. However, this work concentrates on detecting the disturbing content in the text alone. An intensive study of these definitions from a distinct viewpoint helped us to give the following report for fake news: "Fake News is such content that disturbs the inner self of an individual or entity by tarnishing his/their reputation and provoking anti-social activities like terrorism, communal riots, political turmoil, racial prejudices, sexual harassment, etc., to make money. [31]".

d. Algorithm for Fake News Text Processing

In this summary, well-defined samples and features are given as follows:

Table 1 Fake News positing of Social Forum [40]

Source	Definition
Facebook	The subject matters which provoke or encourage grave violence that is dangerous to the safety of the public or individual, instructions for weapon-making which may hurt or take away the life of folks and intimidation leading to physical harm caused to laymen or leading personalities
Twitter	Terrorism or violent extremism, harassment or bullying people, antagonize fury towards an individual or group of people may not be nurtured by anyone

d. Algorithm for Fake News Text Processing

Step 1. Set Variable

Text Document of News Article → TD

Set of News Logs→NL

Text End Label→TEL

Text Start Label→TSL

Step 2. FOR i<TD **DO**

Word Count Fake News =0

Word Count True News=0

Step 3. IF NL== Word Count Fake News **THEN**

Word Count Fake News=0+1

Step 4. END IF

Step 5. ELSE IF NL= Word Count True News=0 **THEN**

Word Count True News=0+1

Step 6. END IF

Step 7. IF TEL >0 && TSL<0 **THEN**

TEL = i

Step 8. ELSE IF

TSL=i+1

Step 9. END IF

Step 10. END FOR

Step 11. END

- The high-frequency individual token incidence is considered as a function.
- The matrix of all token frequencies is called a multivariate sample.

Hence, characterization is done on a corpus of documents by a matrix existing in the corpus, with one row for every document and one column for every symbol (e.g., word). The vectorization method is used for converting the set of text documents into vectors with numerical functionality. The tokenization, numbering, and normalization are named the Bag of Terms' portrayal or Bag of n-grams. The word phenomenon classifies the document ignoring the relative word location of word details in the text. Stop terms are inclusive of terms like "and", "the", "he" that is supposed to be non-informative in recounting the meaning of the text and prevented from being perceived as a predictive warning. Yet, similar terms also render its support to derive the conclusion, for example, while classifying the writing style. Many problems are visible in the given 'English' stop word list. Hence, care should be taken while selecting a stop word list. The prominent word stop list possesses critical terms for numerous activities, like computers. The word stop list has employed the same preprocessing and tokenization, similar to those used in the vectorizer. Hence, the model locates the stop words; however, 've' is not; it can be retained from the converted text that has. The vectorizer and alert detect these kinds of abnormalities.

e. TFIDF Term Weight

Several terms rise to prominence right through a text corpus of a wide range (e.g., "the," "a," "is") by possessing very little tangible knowledge about the specific contents of the documents. The constant terms can shadow frequencies for unusual and motivating terms to directly feed direct-count data to the classifier. TFIDF transform is used forever in the absence of re-weigh count functionality to a floating-point appropriated by classifier use. TF stands for Term Frequency, and TFIDF stands for Term Frequency-Inverted Document Frequency.

4 Machine Learning Training Data Set and Classification

Figure 3 diagrammatically represents the classification process and begins data preprocessing commissioned by eliminating unwanted nature data characters and words. The derivation of N-gram's characteristics and a feature matrix is structured to represent its documents. Training the classifier is the final step in the classification process. Forecasting the class of documents is done by scrutinizing various classifiers, especially six distinct ML algorithms, i.e., Hybrid Support Vector Machines (SVM) [32] [33]. These classifiers are implemented in the Python Natural Language Toolkit (P-NLTK) [34, 35]. Testing and training sets are the two classifications of the dataset. For example, fivefold cross-validation was used in the experiments done after that. Therefore, every validation was done with a dataset usage of 80% for training and 20% for testing.

Assume that our training set $\Delta = DataSet_i, 1 \leq i \leq TD$ contains TD text documents $DataSet_i$. We estimate the feature values using TF/TF_IDF, one of the feature extraction techniques, identical to the whole terms/words used in the training corpus's entire documents, and choose the most significant feature values of x terms $TD_i(1 \leq i \leq p)$. The features matrix $FM = \left[FM_{ij}\right]_{1\leq i\leq TD, 1\leq j\leq p}$ is constructed in the next step.

where,

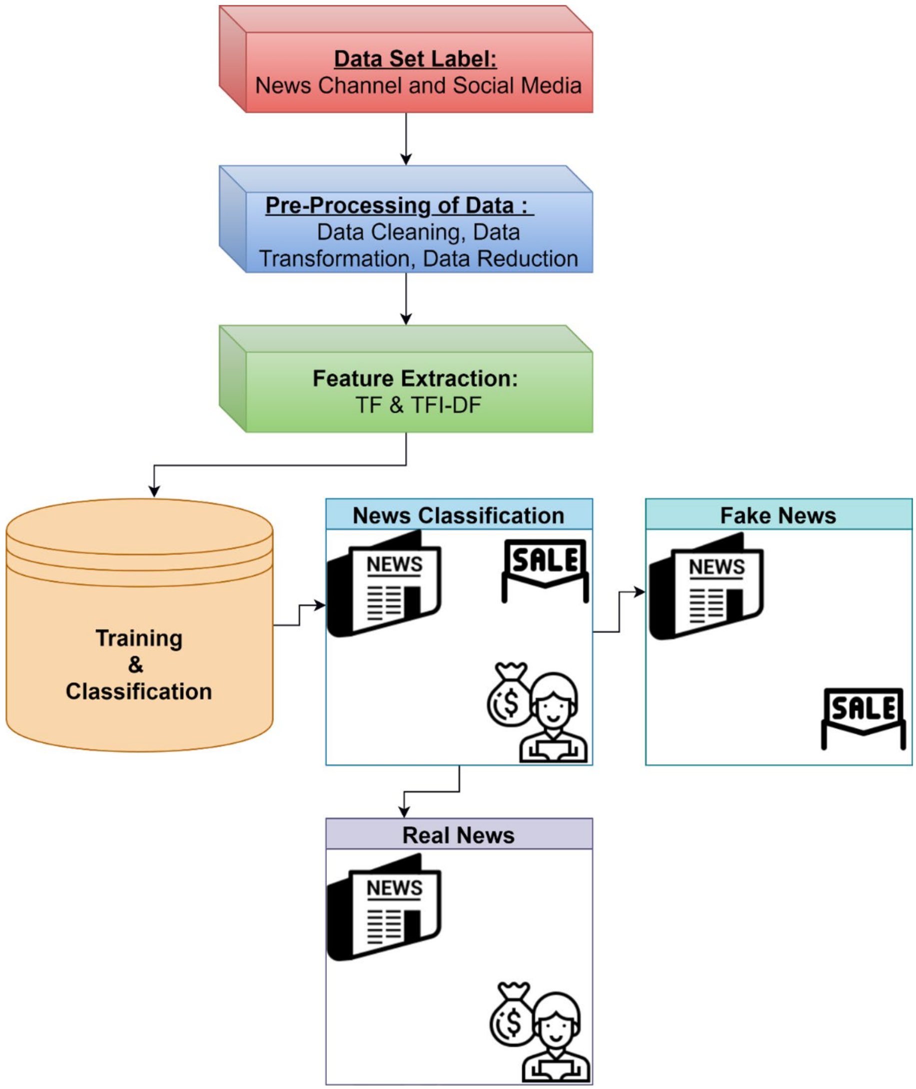

Fig. 3 Text Classification using ML

$$
\begin{aligned}
& FM_{ij} = Feature(TD_j) \\
& If \\
& TD_j \in TD_i \\
& Else \\
& TD_{ij} = 0 \\
& EndIf
\end{aligned}
$$

To say it in another way, (TF/TF-IDF) and TD_{ij} are parallel to each other for the term TD_j for a text document TD_i. The value of such a part is NULL (0) when there is no presence of the term in the document.

5 Experimental Setup

5.1 Date Set

A few public datasets are available in FND because it is comparatively a new area of research. We have used a unique dataset gathered by our team from a compiler of publicly available news articles. Our model also tested the dataset on Facebook, which has public accessibility. Real news articles were collected from the News website, and the reel news was gathered from a fake news website called www.ReelNews.com. The dataset collected from non-reliable websites has been working with Facebook to banish it. 25,000 fake news articles and 25,000 real articles were used for research. Political news articles were mainly focused on because, at present, spammers primarily target this sort of news. In particular, the news items from both reel and real stream took place in a similar timeline. The length of each article is better than 500 characters [36, 37].

The proposed result was loaded with all the modules like NumPy, Matplotlib, Pandas, Sky system learning libraries, and the NLTK libraries, abbreviated as the NLP tool. All the CSV format files were read; the ride was examined, checked, and corrected. A feature was chosen to predict class distribution, clean the missing interest, stem, and process info to build a unigram and bigram-like N-gram. Unigram will evaluate word form, and bigram will consider string and term length. Count vectorization, TF-IDF Transformer and Vectorizer are used to extract function [38, 39].

5.2 Count Vectorization

The accumulation of text documents is transformed into a token count matrix. Suppose there is no APRIORI dictionary, then an analyzer that can do any form of feature collection is not used. In that case, the number of features will be the same as the vocabulary size detected by testing the results.

5.3 TFIDF Transformer

Transformation of count matrix into regular TF/TF-IDF demonstration signifies term frequency, while TF-IDF signifies the reciprocal document-frequency of term-frequency time. Retrieval of knowledge is the reason for the standard term weighting scheme, and there is often a vigorous usage of database classification because features occur within a restricted fraction. It aims to use TF-IDF instead of a token's new frequencies in a given database to reduce tokens, which happens rather commonly in specified amounts with a few experimental insights.

5.4 TFIDF Vectorizer

The transformation of a collection of raw documents into a TF-IDF matrix is equivalent to the import of Count Vectorizer and Word2Vec. The feature was extracted for processing. The TF-IDF function is applied to the training dataset, counting words in a statement and vocabulary. In speech value, Part-of-Speech [14] tagging is one of the main components, and each of them is almost analyzed by NLP. POS-Tagging means labeling words in addition to their actual Part-Of-Speech Text classification with Word2Vec to differentiate our news as fake or non-fake using linear SVM Pipeline, which is used to merge feature selection and counter vector. The pipelines are created with the help of Functional Data Structures and Higher-Order Functions. Channels are a series of functions that reverses a value forever.

The formula to calculate TF-IDF for a term t of document D in a document set is

$$\mathrm{TF\text{-}IDF}\ (\mathrm{t},\ \mathrm{d}) = \mathrm{TF}(\mathrm{t}, \mathrm{d}) * \mathrm{IDF}(\mathrm{t})\ \text{and}$$

The IDF is calculated as

$$\mathrm{IDF}(\mathrm{t}) = \log\left[\mathrm{n}/\mathrm{DF}(\mathrm{t})\right]\ + 1(\text{If Smooth\_IDF} = \text{Fake}),$$

where n is the total number of documents in the document set, and DF(t) is the document frequency of t. The document frequency is the number of documents in the document set that has the term t. The impact of adding “1” to the IDF in the above equation is that its terms with zero IDF, i.e., terms that appear in all documents in a training set, will not be neglected. Training statement and label are adhered to with SVM and forecast the statement with taken m mean to calculate the arithmetic mean and the specified axis.

For instance, the flattened number is surpassed by average, apart from the defined axis. Float64 intermediate and return values are used for integer inputs and prediction of N-gram. A “Pickling” module pickle mechanism was imported, translating a hierarchy of Python objects into a byte stream, FND, by feeding it as a feature. Once the feature is called, it says, "Please insert the text of the news that you want to check:"

5.5 Experiments Training Data Set

The proposed ML algorithms are run on the dataset to predict the genuineness of the articles. By investigating the effect of the performance of size(n) of N-Grams, the experiments commenced. It began from unigram ($n=1$), then bigram ($n=2$), and slowly rose n by one till $n=4$ was reached. Besides that, each n value was evaluated with a combination of distinct features. The algorithms created the learning models, and these algorithms later predicted the labels allocated to the testing data. At the outset of our research, the model was applied to a fusion of news articles belonging to different periods with various social trends. Our model attained an accuracy of 98% on using this type of data. Hence, it was determined to collect our dataset that needs fake and original new items of the same year and month.

Moreover, to restrict the scope of the articles, news items concerning the 2021 Indian elections and the related articles were paid attention to. Altogether we picked 5000 articles, out of which the fake articles are 3000 are, and the real articles are 2000. This represents the dataset’s subset mentioned in the last section that concentrates, especially on political news.

5.6 TF-IDF Vectors with NLP

We investigated feature extraction methods of TF-IDF and TF. We also sorted the number of *p*'s features, in the range between 1000 and 50,000. Figure 4 displays the acquired results. This model adopts the headline-article pair of TF-IDF Vectors, their cosine resemblance (the measurement of the analogy between 2 non-zero vectors using standard metric) as input, and forecasts the output stance. The TF-IDF vectors were then passed to an NLP (Figs. a, b, 6a and b). NLP symbolized the words in a concealed and impeccable form for the use of other layers. The output probabilities of the stances were predicted in the final dense layer.

6 Results and Discussion

In this chapter, every experiment was done on the KAGGLE3.1 dataset of 40.5 k entries. The NN models also benefited other research, and there are many preprocessing approaches on this single dataset. The dataset was differentiated as train and test parts in a 0.43 ratio, and the train part offered 500 samples for validation. To accelerate the training speed, the input documents were curtailed to 500 encoded words. ML is a means to acknowledge data trends and use the same to envisage or take decisions automatically. Regression and classification are the two primary approaches used to concentrate on ML. The success of the algorithm to find fake/real news is estimated here. The hyper-plane is used there to divide the point with the maximum Margin of two data classes. Labelled training data are issued to other terms. The model creates a perfect hyperplane that classifies new illustrations. This hyperplane acts as a line that divides a plane into two sections, placing them on both sides of each unit (Figs. 5, 6).

Figure 7 shows the collection of entire news documents from the social forum. We have crawled a total of 2000 online social media news; among them, 1300 online social media news is true, and 700 online social media news is fake. In the fake categories, 24.12%

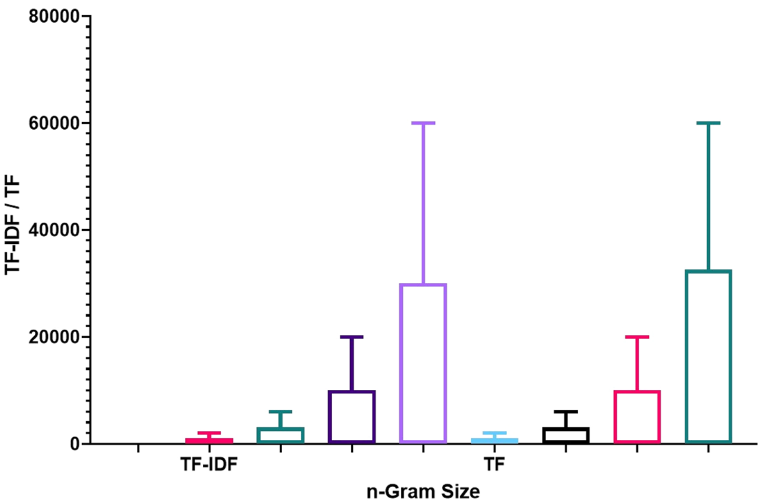

Fig. 4 Analysis of Hybrid SVM with TF-IDF and TF

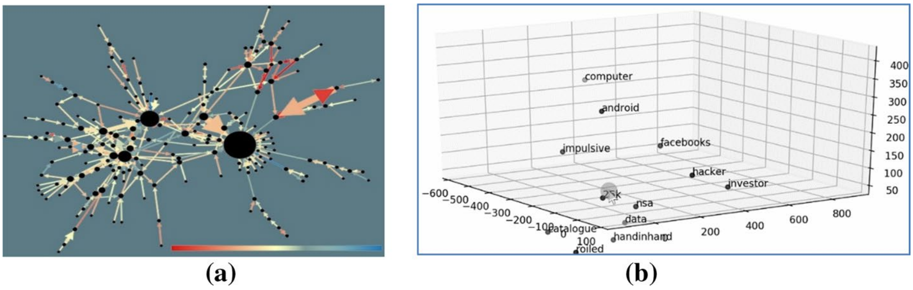

Fig. 5 **a** TF-IDF convert text sentences into numeric vectors; **b** Text Classification with NLP: TF-IDF versus Word2Vec versus BERT

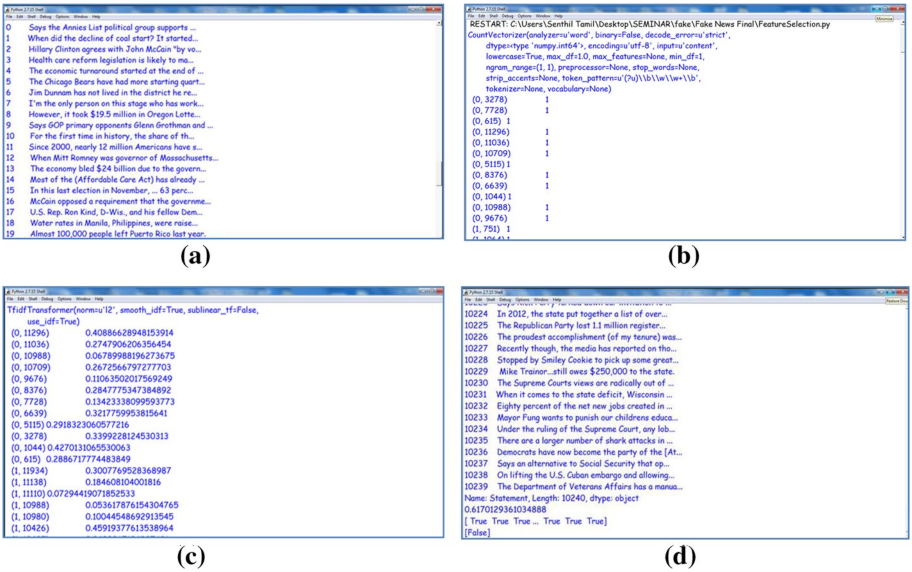

Fig. 6 **a** Loading the Online News, **b** Feature Extraction by Counter Vectorization. **c** Feature Extraction TFIDF Transform, **d** Classify News by Hybrid SVM is True/Fake

of source texts were collected from the online news channels, 18.91% collected from the Facebook User Comments, and 12.22% from the News Paper. Other sources such as Facebook User Posts contributed 10.12% of text documents.

The training dataset contains approximately 20,000 rows of data from various online articles. Researchers had to do many data pre-processing to train our approaches, as demonstrated by our source code. The mentioned characteristics tend in a proper training dataset: a unique ID for textual content, the title of a news article, the people of the news text, the text of the news article, the label of the textual content as untrusted or accurate Most of the data sets previously used to construct the ML classification model comprises the malware binaries. Feature extraction in the existing models is based on the manual selection of unique features identified by “looking” at the binary files. Most of the models have not employed a systematic ML- Hybrid SVM approach to extract the frequently occurring features. The few have intuitively applied the n-gram, with $n=1,2,3,4$. The selection of n

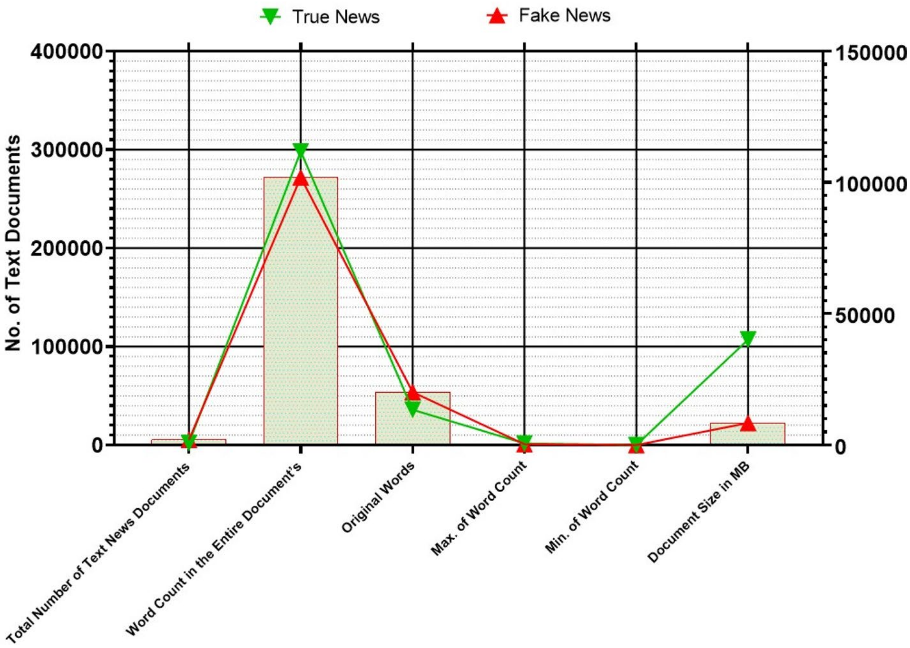

Fig. 7 Sources of News Document's Announcement in True versus Fake News

did not have any ML basis. We use a systematic Hybrid SVM approach, which helped us construct a solid framework to use various classification algorithms (Fig. 8).

Our modeling uncovered that all our features could be segregated as either sparse or dense features:

1. A large number of sparse features degraded the performance of most of the models since a lot of non-relevant features were being used for model construction.

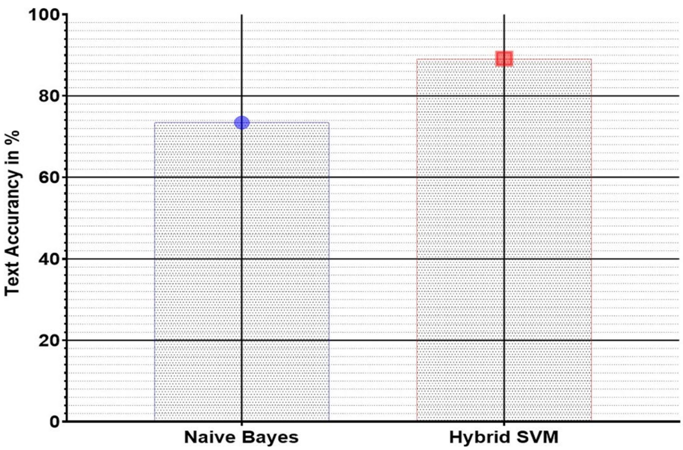

Fig. 8 Text Accuracy with ML Classifiers

2. The characteristic randomness of curating dense nodes from the subset of features recommended the SVM algorithm's robustness against over-fitting. It shows the best performance even before feature selection.
3. Apart from the proposed classifier, all the other classifiers did not perform well before feature selection.

6.1 TF-IDF Test Data

TF-IDF Vectors are represented in a complicated manner and are not dependent on word count alone. In the bi-parts of TF-IDF Vectors, TF represents Term Frequency, and IDF represents Inverse Document Frequency. This is described using the following Eqs. (1) and (2):

$$TF(t) = \frac{NumberOfWords' t' inTextDocument}{TotalWordCountInTextDocument} \tag{1}$$

$$ID(t) = \frac{TotalWordCountInTextDocument}{NumberOfWords' t' inTextDocument} \tag{2}$$

Different levels of matrix representation.

- Words Level -TF-IDF scores of terms
- N-gram Level—TF-IDF scores of N-Grams
- Character Level -TF-IDF scores of character-level n-Grams

N-gram's probability of the next term/character in the sequence is roughly calculated using the Markov model. The model was evaluated on test data. We propose to correctly measure the proximity of the envisaged stance to the original, and we chose *'Text Classification Accuracy'* as our evaluation parameter. Figure 9 is a description of the correct prediction for the models. NLP provides an accuracy of 95.16% when TF-IDF word vector representations are passed into it. The accuracy of our model was compared with the other

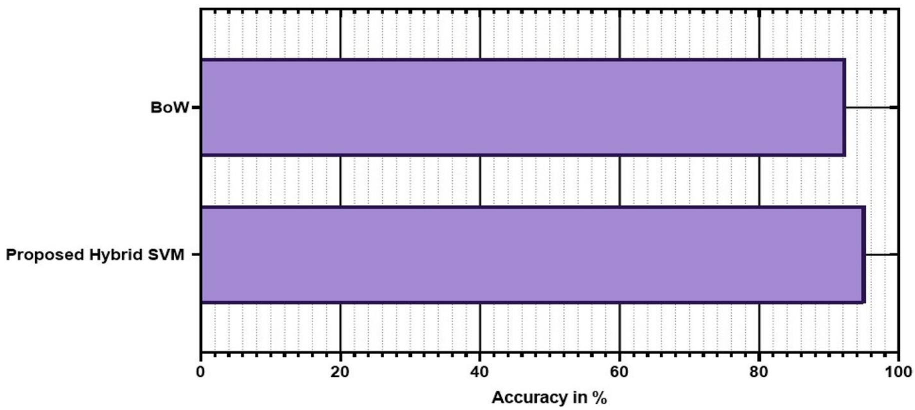

Fig. 9 Predicted Accuracy for the ML Models

models presented in the works of literature. The performance of our model is considerably better than the second-best performing model.

Figure 10 is an illustration of accuracy Vs. the number of training datasets. The study shows that the accuracy classification increases with the increased dataset, and with a text feature extraction, the TF-IDF predominates the BoW.

6.2 Performance of Text Accuracy using ML-Hybrid SVM Classifier

The 95.8% accuracy of the model testifies ML's prowess in handling fake news or true news classification centering on language patterns. The representation of this accuracy is possible by using the following confusion matrix that displays each group's prediction count. We collected fake news to get perceptive knowledge in finding the proper fake news classification's complicated and straightforward methods. Figure 11 depicts the fake news dataset that includes different categories and the misclassification rate of the following categories. The non-misclassification of one type was eliminated from this chart. The same method was used for identifying the real news sections that were mainly misclassified as fake news. The section of news was extracted from the URL. There are different sections, and this results in a few overlapping. Any completed sections are being expelled from these charts.

From all over the distinct body texts, the n-grams were captured. Each category of n-grams' weight activation was the maximum true and fake. In real news, "most real" is the truest weight activation, and "least real" is the fakest weight activation. The exact terms (i.e., "most fake" and "least fake") are used for fake news too. To sum up our findings, we paired the "most real" with the "least fake" *n*-grams and paired the "most fake" with the "least real" *n*-grams. We gathered from the *n*-grams captured by the model the 2000 most common words from both these groups. Next, the words shared in both categories were pulled out to obtain those rarely found as "fake/real" indicators. The accuracy rate of the brilliantly performed model on the test set is clearly shown in Fig. 12. The dataset includes only the articles that did not have the given word.

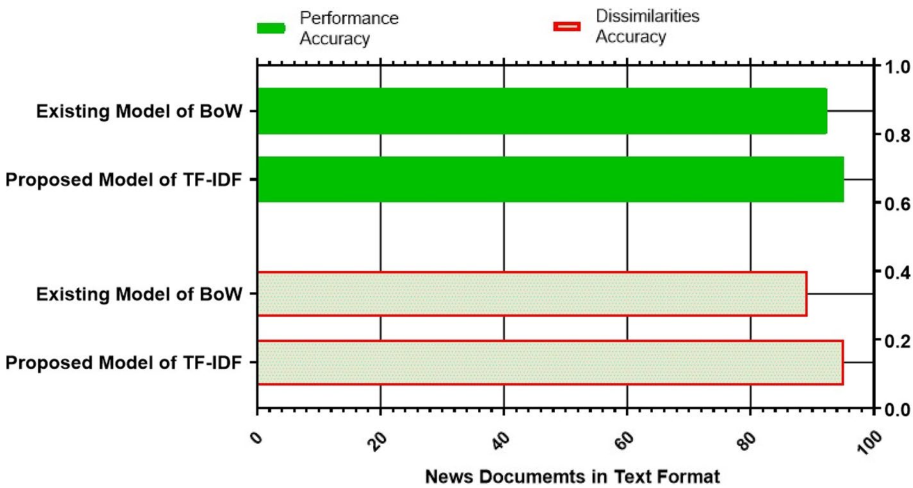

Fig. 10 Properties of Training Set of Text Accuracy

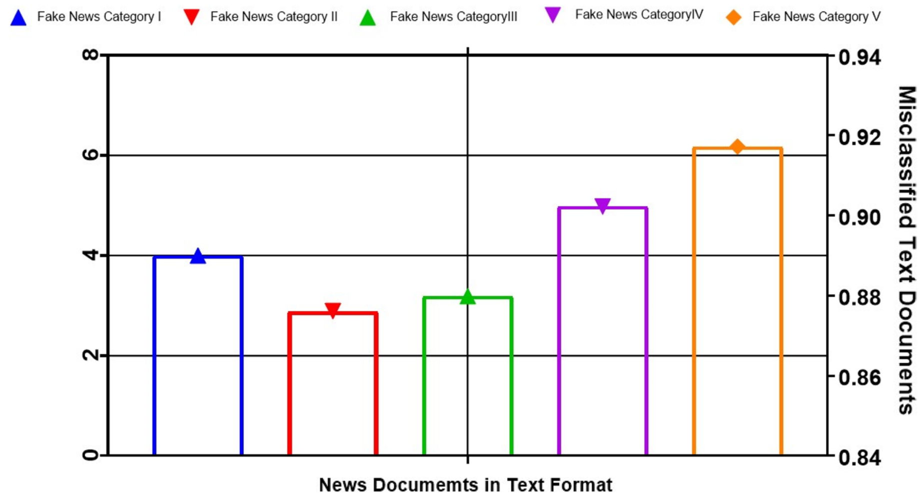

Fig. 11 Fake News Categories vs. Misclassification Rate

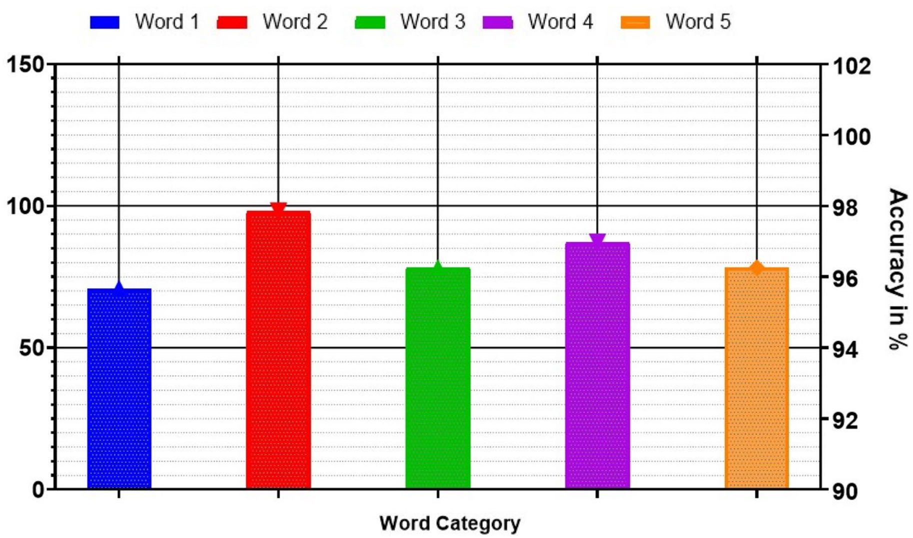

Fig. 12 Testing of accuracies using articles with each word

6.3 Cleaning of Text

Our data are pre-processed to shun away from any disturbing features that were recurrent, and it was divided into three key steps. Each incremental step has equivalent models trained and tested on the preprocessed data at the point represented by the step name. The fundamental steps are constructed on one another so that the second step has the pre-processing of the first step, and the third set is inclusive of the first two pre-processing methods. The

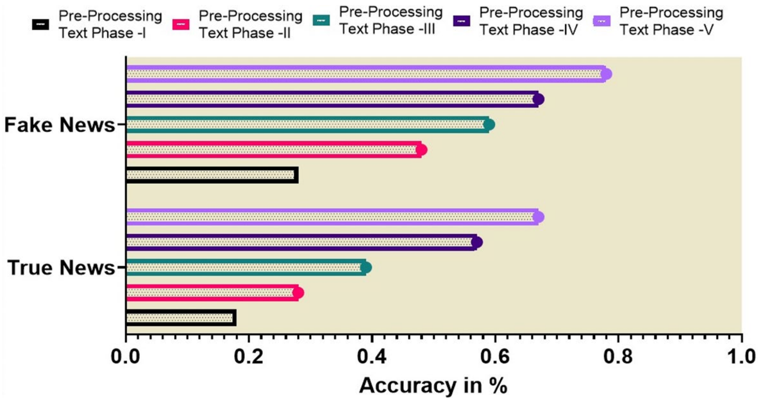

Fig. 13 Measures the Size of a Dataset Fake versus True

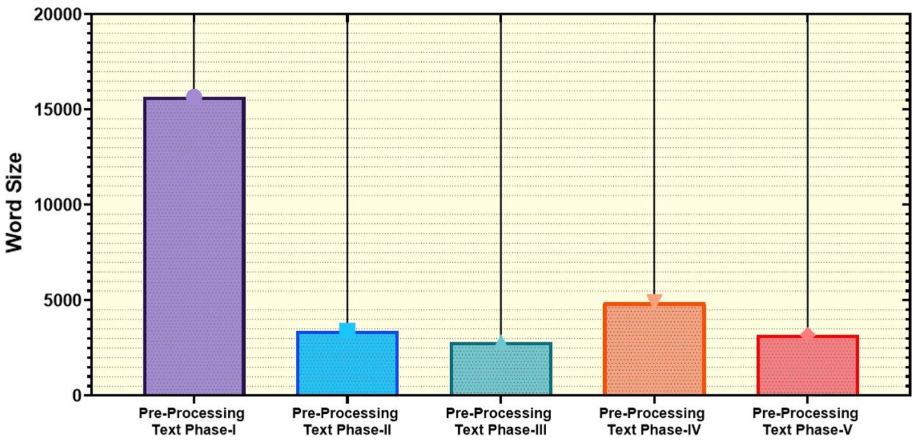

Fig. 14 Phases of Pre-Processing of Words Size

first step is mere preprocessing. The second step is removing any non-English words. The resultant action of changed weight distribution with more text cleaning is represented in Fig. 13. The relativity of this with the standard deviations and vocab size is described in Fig. 14. Figure 15 depicts the resultant change of accuracies with more cleaning.

7 Conclusion and Future Work

This article embodies the concept of fake news in the new world enriched with information. Artificial intelligence is an approach that is already being used for fully automated news classifiers into chapters, keyword selection, and text summarization in terms of

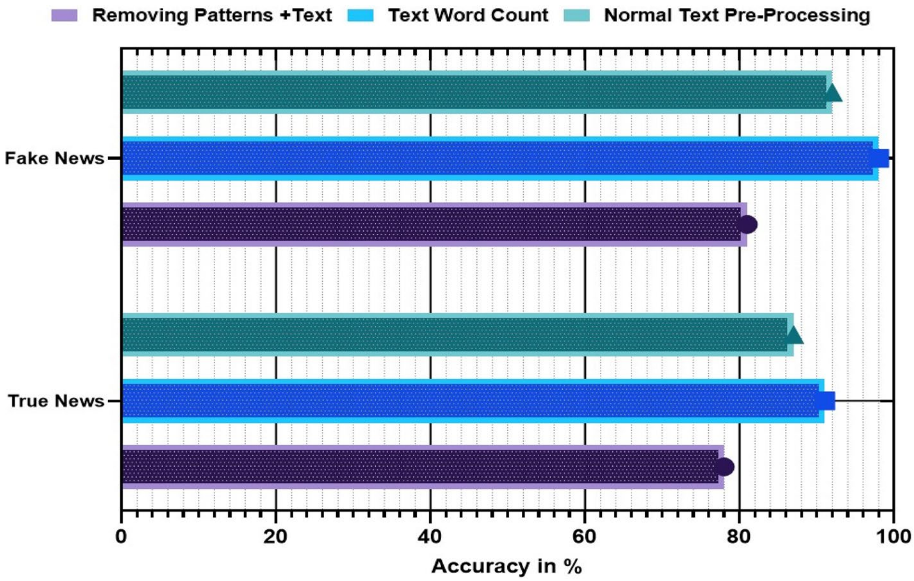

Fig. 15 Text Cleaning Accuracy of Ture versus Fake News

generating catchy news articles or synopses of first-hand data. The data authors used in our work were gathered from the Online Social Forum like Facebook and Twitter and included news reports from various domains to protect the large proportion of the news rather than particularly classifying world news. The finding of a perfect combination of pre-processing and CNN classification was achieved by conducting more than a hundred experiments, defining some specific limitations of the FND problem comparing with other text classification tasks. This leads to adopting Hybrid SVM, the best ML classification system for identifying optimistic fake news. Using a well-defined TF and TF-IDF in the NLP model, the existing model architectures can outperform by 2.5% and reach an accuracy level of 95.16% on test data. We could track social networking news like Facebook and Twitter. The proposed model used unigram features and a Hybrid SVM classifier and attained the highest accuracy of 91.23%.

Future works related to Fake News Detection are possible only on supervised models and texts that are not sufficient in all cases. To rectify this problem, most of the research spotlights additional information, like author information. I think the most promising approach would be the automatic fact-checking model that is expected to be knowledge-based. The outcome expected from the model would be to extract information for the text and check the database information, and it also warns the consumers that the news will be fake. Through this, they may get awareness towards these kinds of untrusted information. The drawback of this approach would be a periodical and manual updating of knowledge for sustainability.

References

1. Avinesh, P. V. S., Schiller, B., Caspelherr, F., Chaudhuri, D., Meyer, C. M., & Gurevych, I. (2018). A Retrospective Analysis of the Fake News Challenge Stance Detection Task. In Proceedings of the 27th International Conference on Computational Linguistics, pp 1859–1874, Santa Fe, New Mexico, USA.
2. Ahmed H., Traore I., Saad S. (2017) Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques. In: Traore I., Woungang I., Awad A. (eds) Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments. ISDDC 2017. Lecture Notes in Computer Science, vol 10618. Springer, Cham. https://doi.org/10.1007/978-3-319-69155-8_9
3. Akshay Jain & Amey Kasbe (2018) "Fake News Detection," IEEE International Students' Conference on Electrical, Electronics and Computer Sciences.
4. Allamanis, M., & Sutton, C. (2016) A "Convolutional Attention Network for Extreme Summarization of Source Code. In Int. Conf. on Machine Learning vol 48 (New York, NY) 2091–2100
5. Chen, Y.-C., Liu, Z.-Y., and Kao, H.-Y. (2017). IKM at SemEval-2017 Task 8: Convolutional Neural Networks for Stance Detection and Rumor Verification. In Proceedings of the 11th International Workshop on Semantic Evaluations (SemEval2017), pp 465–469, Vancouver, Canada.
6. Cody Buntain & Jennifer Golbeck (2017) Automatically Identifying Fake News in Popular Twitter Threads. IEEE International Conference on Smart Cloud (SmartCloud).
7. Conroy, N., Rubin, V., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology, 52*(1), 1–4.
8. Dai, J. J., Wang, Y., Qiu, X., Ding, D., Zhang, Y., Wang, Y., Jia, X., Zhang, L. C., Wan, Y., Li, Z., Wang, J., Huang, S., Wu, Z., Wang, Y., Yang, Y., She, B., Shi, D., Lu, Q., Huang, K., Song, G. (2019) Big DL: A distributed deep learning framework for big data. In Proceedings of the ACM symposium on cloud computing, Association for Computing Machinery. pp 50–60. https://arxiv.org/pdf/1804.05839.pdf, https://doi.org/10.1145/3357223.3362707
9. Fake news detection using naive Bayes classifier. (2017) IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), pp 900–903.
10. Felix Hamborg, Norman Meuschke, Corinna Breitinger, & Bela Gipp (2017) News-please: A generic news crawler and extractor. In Maria Gaede, Violeta Trkulja, and Vivien Petra, (eds.) Proceedings of the 15th International Symposium of Information Science, March 2017, pp. 218–223.
11. Gim´enez, M., Baviera, T., Llorca, G., G´amir, J., Calvo, D., Rosso, P., & Rangel, F. (2017). Overview of the 1st Classification of Spanish Election Tweets Task at IberEval 2017. In Second Workshop on Evaluation of Human Language Technologies for Iberian Languages (IberEval 2017).
12. Helmstetter S, Paulheim H (2018) Weakly supervised learning for fake news detection on Twitter. In: 2018 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM), IEEE, pp 274–277
13. Hnin Ei Wynne & Zar Zar Wint. 2019. Content-Based Fake News Detection Using N-Gram Models. In Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services (iiWAS2019). Association for Computing Machinery, New York, NY, USA, 669–673. DOI:https://doi.org/10.1145/3366030.3366116
14. Tagami, T., Ouchi, H., Asano, H., Hanawa, K., Uchiyama, K., Suzuki, K., Inui, K., Komiya, A., Fujimura, A., Yanai, H. et al. (2018) Suspicious News Detection Using Micro Blog Text. arXiv 2018, arXiv:1810.11663
15. Jiang, T., Jian Ping Li, Amin Ul Haq, Abdus Saboor & Amjad Ali (2021) A Novel Stacking Approach for Accurate Detection of Fake News." *IEEE Access, 9*, 22626–22639.
16. Rubin, V. L., Chen, Y., & Conroy, N. J. (2015). Deception detection for news: Three types of fakes. *Proceedings of the Association for Information Science and Technology, 52*(1), 1–4.
17. Xie, S., Wang, G., Lin, S., & Yu, P. S. (2012, April). Review spam detection via time-series pattern discovery. In Proceedings of the 21st International Conference on World Wide Web (pp. 635–636). ACM.
18. Yaqing Wang, Fenglong Ma, Zhiwei Jin, Ye Yuan, Guangxu Xun, Kishlay Jha, Lu Su, & Jing Gao. Eann: Event adversarial neural networks for multi-modal fake news detection. In Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining, pp 849–857. ACM, 2018.
19. Zubiaga, A. Aker, K. Bontcheva, M. Liakata, R. Procter (2018) Detection and resolution of rumors in social media: a survey. *ACM Computer Survey*, 51(2), 32:1–32:36.
20. Horne, B.D., Adali, S.: This just in: fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. In: the 2nd International Workshop on News and Public Opinion at ICWSM (2017)

21. Ahmad, I., Yousaf, M., Suhail Yousaf & Muhammad Ovais Ahmad (2020) Fake News Detection Using Machine Learning Ensemble Methods." *Hindawi-Complexity, 2020*, 1–11.
22. Islam, M. R., Liu, S., Wang, X., et al. (2020). Deep learning for misinformation detection on online social networks: A survey and new perspectives. *Social Network Analysis and Mining, 10*, 82. https://doi.org/10.1007/s13278-020-00696-x
23. Ma, J., Gao, W., Mitra, P., Kwon, S., Jansen, B. J., Wong, K.-F., & Cha, M. (2016). Detecting rumors from microblogs with recurrent neural networks. In Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI, address, pp. 3818,3824
24. Jan, B., Farman, H., Khan, M., Imran, M., Islam, I. U., Ahmad, A., Ali, S., & Jeon, G. (2017). Deep learning in big data Analytics: A comparative study. *Computers & Electrical Engineering, 1*, 1–13.
25. Jennifer Golbeck, Jennine B. Everett, Waleed Falak, Carl Gieringer, Jack Graney, Kelly M. Hoffman, Lindsay Huth, Zhenya Ma, Mayanka Jha, Misbah Khan, Varsha Kori, Matthew Mauriello, Elo Lewis, George Mirano, William T. Mohn IV, Sean Mussenden, Tammie M. Nelson, Sean Mcwillie, Akshat Pant, & Paul Cheakalos, Fake news vs satire: A dataset and analysis, 05 2018, pp. 17–21.
26. Kamran Kowsari, Mojtaba Heidarysafa, Donald E. Brown, Kiana Jafari Meimandi, & Laura E. Barnes. RMDL: Random Multimodel Deep Learning for Classification. Proceedings of the 2nd International Conference on Information System and Data Mining—ISDM '18, pages 19–28, 2018. arXiv:1805.01890.
27. Kim, J., Tabibian, B., Oh, A., Schölkopf, B., & Gomez-Rodriguez, M. (2018) Leveraging the crowd to detect and reduce the spread of fake news and misinformation. In Proceedings of the eleventh ACM international conference on web search and data mining, pp 324–332
28. Kuai, Xu., Wang, F., Wang, H., & Yang, Bo. (2019). Detecting fake news over online social media via domain reputations and content understanding. *Tsinghua Science and Technology, 25*(1), 20–27.
29. Lau, R. Y., Liao, S. Y., Kwok, R. C. W., Xu, K., Xia, Y., & Li, Y. (2011). Text mining and probabilistic language modeling for online review spam detecting. *ACM Transactions on Management Information Systems, 2*(4), 1–30.
30. Li, L., Cai, G., & Chen, N. (2018a) A rumor events detection method based on deep bidirectional GRU neural network. In 2018 IEEE 3rd international conference on image. Vision and computing (ICIVC), IEEE, pp 755–759
31. Markines, B., Cattuto, C., & Menczer, F. (2009). Social spam detection. In Proceedings of the 5th International Workshop on Adversarial Information Retrieval on the Web (pp. 41–48)
32. Umer, M., Imtiaz, Z., Ullah, S., & Mehmood, A. (2020). Gyu Sang Choi and Byung-Won On, "Fake News Stance Detection Using Deep Learning Architecture (CNN-LSTM)." *IEEE Access, 8*, 2169–3536.
33. Natali Ruchansky, Sungyong Seo, & Yan Liu. CSI: A hybrid deep model for fake news detection. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pp 797–806. ACM, 2017.
34. Tolmie, P., Procter, R., Randall, D. W., Rouncefield, M., Burger, C., Wong Sak Hoi, G., Zubiaga, A., Liakata, M. (2017) Supporting the use of user-generated content in journalistic practice. In Proceedings of the CHI Conference on Human Factors in Computing Systems. ACM, pp. 3632–3644
35. Qing Liao, Heyan Chai, Hao Han, Xiang Zhang, Xuan Wang, Wen Xia & Ye Ding (2021) An Integrated Multi-Task Model for Fake News Detection. IEEE Transactions on Knowledge and Data Engineering. https://doi.org/10.1109/TKDE.2021.3054993.
36. Rada Mihalcea, & Carlo Strapparava, The lie detector: explorations in the automatic recognition of deceptive language, Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, August 04–04, 2009, Suntec, Singapore
37. Rubin, V. L., Conroy, N. J., Chen, Y., & Cornwell, S. (2016). Fake News or Truth? Using Satirical Cues to Detect Potentially Misleading News. In Proceedings of NAACL-HLT (pp. 7–17).
38. Shivam B. Parikh & Pradeep K. Atrey (2018) Media-Rich Fake News Detection: A Survey. In IEEE Conference on Multimedia Information Processing and Retrieval.
39. Ahmad, T., Akhtar, H., Chopra, A., & Waris Akhtar, M. (2014) Satire detection from web documents using machine learning methods, pp. 102–105.
40. Bouazizi, M., & Ohtsuki, T. O. (2016). A pattern-based approach for Sarcasm detection on twitter. *IEEE Access, 4*, 5477–5488.

Roy Setiawan is a lecturer in the Department Management, Faculty Business, and Economic at Universitas Kristen Petra, Indonesia, with research interests in the field of Business and Management. He graduated from the informatics engineering and management science programs. He previously served as the Chairman of the Business Management Study Program and Head of the Leadership Laboratory at Universitas Kristen Petra. He is currently completing doctoral studies in Management at Universitas Airlangga Surabaya and is developing a new model in his research.

Dr Vidya Sagar Ponnam is an Indian academician who is serving as an Associate Professor in the Department of Computer Science& Engineering in KL University Vijayawada, Andhra Pradesh, India. He got the Ph.D. (Computer Science & Technology) from Sri Krishnadevaya University, Andhra Pradesh, India, in 2016. M.Tech. (Computer Science & Engineering) from Acharya Nagarjuna University, Andhra Pradesh, India, 2010. The major domain/specialization of doctorate Is Software Engineering application with Deep Learning, Image processing, Data Mining and Networking. I had around 10 yrs of IT industrial experience with major MNC's & I am currently acting as a reviewer/editorial member if international journals and organising members for international conferences.

Dr. Sudhakar Sengan is currently working as Professor and Director of International Relations in the Department of Computer Science and Engineering, PSN College of Engineering and Technology (Autonomous), Tirunelveli–627152, Tamil Nadu, India. He received PhD degree in Information and Communication Engineering from Anna University, Chennai, Tamil Nadu, and India. He has 20 years of Experience in Teaching / Research / Industry. He has published papers in 100 International Journals, 20 International Conferences, and 10 National Conferences. He is a Research Supervisor at Anna University in the Faculty of Information and Communication Engineering. His research interest includes Network Security, Information Security, and MANET, Cloud Computing, IoT. He received an award of Honorary Doctorate (Doctor of Letters-D.LITT.) from International Economics University, SAARC Countries in Education and Students Empowerment in April 2017. He has published more than 20 Indian Patents and 03 International Patents. He has published a 03 Textbook for Anna University syllabus. He is a member of various professional bodies like MISTE, MIEEE, MIAENG, MIACSIT, MICST, and MIEDRC.

Dr. Mamoona Anam is Visiting Assistant Professor in International Islamic University, Islamabad. She did her MS Computer sciences in "Copyright Protection of Plain Text using Zero watermarking approach". She is also the PhD scholar doing research on "Learning Based Bibliometric Research Performance Assessment of Degree Awarding Institutes using Stochastic Models".Her research interests are Cryptography and Networking, digital watermarking, text watermarking, artificial intelligence, Scientometrics, Bibliometrics, Information Retrieval, Data Mining, Social Network Analysis and Mining, Machine Learning, Data Grids.

Dr. Chidambaram Subbiah received B.Tech. (IT) degree from Anna University, Chennai in 2005, M.E,(CSE) from Anna University Chennai in 2009 and PhD from anna university Chennai in 2020. He is currently working as an Assistant Professor (Senior Grade), IT department, National Engineering College, Kovilpatti. His research work is published in many reputed conferences and journals. His current research includes Data Mining, Machine Learning and Networking.

Dr. Khongdet Phasinam is a lecturer from school of agricultural and food engineering, faculty of food and agricultural technology, Pibulsongkram Rajabhat University. He graduated in Doctor of Philosophy (Ph.D.) in Agricultural and Food Engineering, Institute of Engineering, Suranaree University of Technology, Thailand. He is interested in conducting research in agricultural engineering and agricultural machinery.

Dr. Manikandan Vairaven born in1966, completed his B.E degree in Mechanical Engineering at Maduraikamaraj university Madurai, M.E degree in Engineering Design at Bharathiyar University Coimbatore and PhD in Composites at Manonmanium Sundaranar university Thirunelveli. To his credit he has published 62 Journal publication in various reputed international and national journals and 25 conferences. Completed one DST sponsored project on Basalt Fiber Composites and received conference grant from DST. 6 scholars has completed PhD under his supervision and currently one is perusing in metal matrix composites. Has 31 years of experience in teaching and research currently holding the position as professor in mechanical engineering and Principal at PSN College of Engineering and Technology affiliated to Anna University, Chennai, Tamil Nadu, India.

Dr. Selvakumar Ponnusamy is currently Professor in the Department of Mechanical Engineering, PSN College of Engineering and Technology, Tirunelveli which is an autonomous institution affiliated to Anna University. He is also holding the post of Executive Director. He is the recipient of University ranks in his both UG and PG degrees. He has published more than 40 research papers in international journals and conferences. His research interest includes Nanofluids, Heat Transfer, Laser Texturing & Solar Energy. He has authored one book and three book Chapters.

Authors and Affiliations

Roy Setiawan[1] · Vidya Sagar Ponnam[2] · Sudhakar Sengan[3] · Mamoona Anam[4] · Chidambaram Subbiah[5] · Khongdet Phasinam[6] · Manikandan Vairaven[7] · Selvakumar Ponnusamy[7]

Roy Setiawan
roy@petra.ac.id

Vidya Sagar Ponnam
pvsagar20@gmail.com

Mamoona Anam
mamoona.anam.vt@iiu.edu.pk

Chidambaram Subbiah
chidambaramraj1@gmail.com

Khongdet Phasinam
phasinam@psru.ac.th

Manikandan Vairaven
vaimanikandan@yahoo.com

Selvakumar Ponnusamy
rpselvakumar@gmail.com

[1] Department Management, Universitas Airlangga, Jawa Timur, Indonesia

[2] Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram 522502, Andhra Pradesh, India

[3] Department of Computer Science and Engineering, PSN College of Engineering and Technology, Tirunelveli 627152, Tamil Nadu, India

[4] Department of Computer Sciences and Software Engineering, Faculty of Basic and Applied Sciences, International Islamic University, Islamabad, Pakistan

[5] Department of Information Technology, National Engineering College, Kovilpatti 628503, Tamil Nadu, India

[6] School of Agricultural and Food Engineering, Faculty of Food and Agricultural Technology, Pibulsongkram Rajabhat University, Phitsanulok, Thailand

[7] Department of Mechanical Engineering, PSN College of Engineering and Technology, Tirunelveli 627152, Tamil Nadu, India