Fake news detection in social media based on sentiment analysis using classifier techniques

Sarita V Balshetwar [1,2] · Abilash RS [3] · Dani Jermisha R [4]

Abstract
Fake news on social media, has spread for personal or societal gain. Detecting fake news is a multi-step procedure that entails analysing the content of the news to assess its trustworthiness. The article has proposed a new solution for fake news detection which incorporates sentiment as an important feature to improve the accuracy with two different data sets of ISOT and LIAR. The key feature words with content's propensity scores of the opinions are developed based on sentiment analysis using a lexicon-based scoring algorithm. Further, the study proposed a multiple imputation strategy which integrated Multiple Imputation Chain Equation (MICE) to handle multivariate missing variables in social media or news data from the collected dataset. Consequently, to extract the effective features from the text, Term Frequency and Inverse Document Frequency (TF-IDF) are introduced to determine the long-term features with the weighted matrix. The correlation of missing data variables and useful data features are classified based on Naïve Bayes, passive-aggressive and Deep Neural Network (DNN) classifiers. The findings of this research described that the overall calculation of the proposed method was obtained with an accuracy of 99.8% for the detection of fake news with the evaluation of various statements such as barely true, half true, true, mostly true and false from the dataset. Finally, the performance of the proposed method is compared with the existing methods in which the proposed method results in better efficiency.

✉ Sarita V Balshetwar
balshetwar.satara@gmail.com

Abilash RS
abilash.res2k21@yahoo.com

Dani Jermisha R
danijermisha01@yahoo.com

[1] YSPM'S, YTC, Faculty of Engineering, Satara, Maharashtra, India

[2] DBATU, lonere, Raigad, Maharashtra, India

[3] Bethlahem Institute of Engineering, Ulaganvillai, Tamil Nadu, India

[4] Arunachala College of Engineering For Women, Manavilai, Tamil Nadu, India

1 Introduction

People are spending more and more time interacting on social media, as the wide adoption of smartphones makes their access available almost anytime and anywhere, which is not the case with traditional media. In addition, they facilitate interaction with friends, families, and even strangers through the comment chains, be it through comments, discussions, or simply like and dislike buttons. This has made social media a main channel for the dissemination of news. However, new technologies and features can be used through social media platforms to spread fake news on a large scale. Such inaccurate information might result either from a deliberate attempt to deceive or mislead (disinformation) or from an honest mistake (misinformation). Rumours can fall into either of these two categories, depending on the intent of the source, given that rumours are not necessarily false but may turn out to be true. Unlike rumours, fake news is, by definition, always false and, thus, can be seen as a type of disinformation. Therefore, credible and reliable sources of information are needed so that the public does not fall prey to the intentions of those interested in manipulating reality.

Fake news can function as propaganda or misinformation, but it always appeals to the emotions of the public and the intent to cover rational responses, analysis, and comparison of information from several sources, encouraging inflammation and outrage and can easily lead to conspiracy theories and partisan biased content that negatively affects. The major source of news and data for the public is served by social media and online news articles because it is easily accessible, subsidized and readily available with one click. However, simultaneously, it also helps to spread false news that has significant negative effects on society, that is, messages that are deliberately misinformed. It has many similarities with spam messages since they share common features such as grammatical mistakes, false information, using a similarly limited set of words, and they contain emotionally coloured information that affects the reader's opinion [1]. To alleviate this problem, research on the identification of false news has gained more consideration recently. Despite the many computational solutions currently available to detect fake news, the lack of a comprehensive and community-based false news database has become one of the significant obstacles. Large-scale news passes over social media makes manual verification impossible, which promotes the design and implementation of automated systems for detecting fake news [10].

Detecting fake news is a layered process that involves analysis of the news contents to determine the truthfulness of the news. The news could contain information in various formats such as text, video, image, etc. Combinations of different types of data make the detection process difficult. In addition, raw data collected is always expected to be unstructured and contain missing values in the data. As fake news produces big, incomplete, unstructured, and noisy data [2], raw data pre-processing is extremely important to clean and structure the data before feeding it into detection models. Thereby, fake news creators use many new ideas to make their false creations successful, one of which is to stimulate the emotions of the beneficiaries. This leads to sentiment analysis, the portion of the analysis of the text is responsible for establishing the polarization and the emotional strength demonstrated in a text, which is used in false-news detection approaches as a system or complementary component [3]. It can be as easier as these binary positions such as positive and negative or sometimes the classification will be neutral. Sentiment analysis from text is beyond polarization and may include the determination of users' emotional conditions such as depression, anxiety,

excitement and anger [11]. Some sense dictionaries can help accomplish this task. Sentiment analysis from text like blogs, Twitter and news channels are fine-researched topic fields. However, this is the initial time research has been managed in the context of identifying false news on online social networks. For the motive of this current work, the perceptual analysis of the text is restricted from text messages to the negative and positive polarities of keywords.

While the above systems take various methods and perspectives, this paper is confined to dealing with fake news in both Machine Learning and NLP including missing data. In-depth, research concerning fake news requires additional statistical techniques to address their spontaneous and unstructured nature for data analysis. One challenge in analysing fake news is to properly handle missing data. As fake news is not created for research, most information within social media and news is not structured in a pre-organized manner. Accordingly, missing values are likely to occur, which leads to inconsistent or biased statistical results when applying regression-based analysis or classification [13, 20, 25]. These missing data problems are often encountered even when dealing with online data in a structured form. Two approaches are primarily used to handle incomplete data with missing values. A naïve approach is a deletion under which observations with missing values are excluded from the data analysis. The other approach is to implement an imputation that replaces missing values with plausible alternative values. Xin Liao studied the data hiding in encrypted images by using CS and discrete fourier transform. This show that this method can reduce the average extracted-bit error rate when the block size is appropriate. In practice, imputation methods are preferred because deletion is inefficient and may cause large biases in the model parameter estimates [19]. Among imputation approaches, multiple imputations, initially proposed by Enders et al. [9], are the most popular in both social science and biomedical science due to their statistical and practical advantages.

Multiple imputations yield precise estimates and accurate standard errors that can help obtain less biased results than when using single imputations [12]. These advantages make multiple imputations one of the best options for handling missing values [22]. As many statistical packages have already been developed for implementing multiple imputations on incomplete data, non-statisticians can easily handle missing values and then conduct statistical analyses on the imputed datasets. Given that online data are likely to have a considerable portion of missing values and be able to provide different forms of auxiliary information useful for creating alternative imputed values, multiple imputations are of great importance for fake news research.

It is of paramount importance that the researchers confront and try to eliminate the problem of fake news. There have been many methods that are proposed ranging from NLP analysis to clustering. Amer et al. [4] applied two machine learning supervised algorithms, i.e., Random forest and decision tree classifiers to detect Coronavirus covid-19 fake news, with this model, Count Vectorizer and Document Frequency Vectorizer as feature extraction after making a set of the initial set such as pre-process and normalization of the dataset. However, this work was presented during the labelling of the posts, demanding a lot of exhausting time in addition to labelling, and keeping up to date with what was happening during the pandemic. With that, the labels themselves were chosen, which may have occurred some mistakes during labelling.

Zhou et al. [28] proposed a theory-driven model for fake news detection. Fake news detection is then conducted within a supervised machine learning framework which enhances the interpretability of fake news feature engineering, and studies the relationships among fake news, deception/disinformation, and click baits. Experiments conducted on two real-world datasets indicate the proposed method can outperform the state-of-the-art and enable fake news early detection when there is limited content information. Datasets consisting of the ground truth of, e.g., both fake news and clickbait, are invaluable to understanding the relationships

among different types of unreliable information; however, such datasets are so far rarely available. Furthermore, it should be pointed out that effective utilization of rhetorical relationships and utilizing news images in an explainable way for fake news detection are still open issues. Kaliyar et al. [17] proposed coupled matrix–tensor factorization method to get a latent representation of both news content as well as social context. To classify news content and social context-based information individually as well as in combination, a deep neural network was employed with optimal hyper-parameters. For the task of fake news detection, a feature set can never be considered complete and sound. Jiang et al. [15] evaluated the performance of five machine learning models and three deep learning models on two fake and real news datasets of different sizes withholding out cross-validation. Moreover, the detection of fake news with sentiment analysis is required for different machine learning and deep learning models.

To fill this research gap and with the discussions, this study aims to introduce a new multiple imputation method for fake news detection research that incorporates social media and news content including both structured opinions shared via sub-ratings like user opinions in comments on specific content attributes and unstructured opinions in the form of text. This imputation model also adopts sentiment analysis. In terms of English word segmentation, since traditional machine learning methods cannot solve the long-distance dependencies of texts, it is difficult to analyse the information contained in the problem as a whole and grasp the user's true intention. Therefore, different machine learning-based models are implemented to detect and classify fake news. Each model's performance is measured to categorize various news items correctly, which revealed each model's ability to improve its accuracy in detecting fake news.

The rest of this paper is organised as follows, section 2 contributes a literature survey with different existing research. Section 3 studies the proposed methods carried out in this study. Section 4 analyse the implementation of the study. Finally, the study concludes with concepts in section 5.

2 Literature survey

Disseminated information and its dissemination process build a major problem in detecting these contents immediately, thus highlighting the importance of automatically identifying false news. To overcome this, Sahoo et al [23] proposed an automatic fake news identification technique for the environment chrome using this the detection of fake news on Facebook is possible. Specifically, this uses multiple features associated with a Facebook account in addition to some news content features to analyse the characteristics of the account across deep learning. Shu et al [24] present FakeNewsNet, a repository of fake news data, the news content includes two complete datasets with different features, spatiotemporal information, and social context to make facilitate fake news-related research. This comprehensive description of FakeNewsNet displays an analytical analysis of two datasets from various viewpoints and discussed the advantages of FakeNewsNet for potential applications in social media fake news research. SAF/S performs better in terms of accuracy and F1 score. SAF/A provides a similar result with 66.7% accuracy as SAF/S. This indicates that user engagements can help fake news detection in addition to news articles on the PolitiFact data set. Meanwhile, the selection strategy can be used for web search results to reduce noise in the data collection process.

The studies on user credibility in this context focus more on the frequency and timing of engaging in fake news propagation, rather than specification according to the content of users' tweets. Duan et al [8] approach this challenge by elaborating two features one is linguistic and another one is a sentiment feature from operators' tweet feed as well as retrieving the presence of hashtags, emojis, and political bias in their tweets. These features were later used to categorize operators as those who broadcast or did not broadcast fake news. 72% accuracy was obtained by this proposed approach, among the results in the first 4 positions acquired by systems for the task in the English language. Yet, in applications with diverse classification algorithms and the union of the different representations, not all combinations of representations increased the accuracy. NER in combination with other representations is not suitable for the use of SVMs or ANNs. Moreover, this limit had to be raised multiple times. The need for this is probably due to a large number of features (416,834).

A domain reputation analysis was proposed by Xu et al [26] that reveals the internet pages of real and fake news publishers revealing different registration behaviours, registration time, domain rankings and domain popularity. In addition, fake messages will disappear from the Internet after a certain time. This content on the false and original news corpus is unskilful in detecting false news, using time frequency-inverse document frequency (tf-IDF) and Latent Dichotomy Allocation (LDA) header modelling, while exploring document compatibility with word and word. Vectors are the most promising direction to predict original and false news. This shows the promising aspect of leveraging document similarity to distinguish fake and real news by measuring the document similarity of the news under tests with the known fake and real news corpus. On the other hand, the difference in the topics and word embeddings shows little or subtle difference between fake and real news.

Kumar et al [18] proposed a CNN + bidirectional LSTM ensembled network to gather fresh instances such as PolitiFact and build multiple information for the identification of original and false news and match multiple state-of-the-art approaches. Long Short-Term Memories (LSTMs), Convolutional neural networks (CNNs), attention mechanisms and ensemble methods are examples of multiple state-of-the-art approaches. This research collects 1356 news instances from various users via Twitter and media sources such as PolitiFact and creates several datasets for the real and the fake news stories. The study conclude that CNN + bidirectional LSTM ensembled network with attention mechanism achieved the highest accuracy of 88.78%, whereas Ko et al. tackled the fake news identification problem and achieved a detection rate of 85%. As the result, CNN + bidirectional LSTM ensembled network with focus mechanism obtained 88.78% of maximum accuracy. The results were satisfactory but not promising. The CNN architecture gave the lowest accuracy in comparison to the others that we studied. The LSTM architecture and bidirectional LSTM architecture performed significantly better in comparison to simple CNN architecture. We further increased our appetite for improved accuracy and incorporated more complex models as part of our methodology.

The hybrid deep learning design that merges recurrent neural and convolutional networks for false news identification was proposed by Nasir et al [21]. On two fake news datasets (ISO and FA-KES) this model was certified successfully, achieving the results of detection that are substantially better than other non-hybrid foundation techniques. A paired t-test was used to validate the statistical significance of the results; the experiments were repeated five times (using 5-fold cross-validation, i.e. 80%–20% split); and accuracy was reported at 95% confidence intervals. ISOT is chosen for training because it is much larger and has minimum space for improvement since many models perform above the 0.9 classification accuracy

threshold. Moreover, complex neural network architectures not be considered as part of the study.

A deep convolutional neural network (FNDNet) for false news detection was proposed by Kaliyar et al [16]. This prototype (FNDNet) is outlined instead of relying on hand-crafted features to learn automatically, about the one-sided features for false news identification build in the deep neural network across many hidden layers. As the result, each layer contains many features that will be extracted by a deep Convolutional Neural Network (CNN). Benchmarked datasets were used to train and test the model, and the proposed model achieved state-of-the-art results with an accuracy of 98.36% on the test data. Various performance evaluation parameters such as Wilcoxon, false positive, true negative, precision, recall, F1, accuracy, etc. were used to validate the results. Despite the high performance of our classifier, there is a scope for improvement. A multi-model approach (a combination of different learning techniques) is the main necessity for fake news detection for solving the multi-class fake news detection problem.

Choudhari et al [6] proposed a linguistic model to identify the properties of the content and language-driven features will also generate with the help of this. This linguistic prototype extracts particular news features such as syntactic, sentimental, grammatical, and readability. The language-driven model demands an approach to managing handcrafted feature problems and is time-consuming to maintain the trouble of dimensionality problems. Therefore, a continuous learning model based on neutrality is utilized to achieve the best results for detecting fake news. The results are drawn up to verify the importance of the extracted features of the linguistic model and finally, the integrated linguistic feature-driven model that can achieve an average of 86% accuracy in detecting and categorizing fake messages. However, extensive features/parameters for model performance are lacking. Examine the latent semantic feature-driven fake news detection model, and explore various variants of convolution neural networks for image-driven fake news detection.

The Structure-aware Multi-Head Attention Network (SMAN) was proposed by Yuan et al [27], which merges the content of news, issuing, and reposting connections of users and publishers, to collectively optimize the credibility prediction tasks and fake news detection. As a result, we can explicitly make use of publishers' and users' credibility to detect early fake news. The research conducted experiments on three real-world datasets, and the results show that SMAN can detect fake news in 4 hours with an accuracy of over 91%, which is much faster than the state-of-the-art models.

Dang et al [7] utilize Term Frequency-Inverse Document Frequency (TF-IDF) and word embedding has been implemented to the whole range of datasets to solve the problems in sentiment analysis, for example, sentiment polarity. As a result, comparative knowledge was carried out on the experimental outputs obtained for various designs and input features. The experiments also revealed that CNN outperforms other models, presenting a good balance between accuracy and CPU runtime. RNN reliability is slightly higher than CNN reliability with most datasets but its computational time is much longer. However, exploring hybrid approaches, where multiple models and techniques are combined to enhance the sentiment classification accuracy achieved by the individual models or techniques, as well as to reduce the computational cost is the reliable cost.

Abdullah et al [14] used the multimodal approach with Convolutional Neural Network (CNN) and Long Short-Term memory (LSTM) to classify the fake news articles that achieved significant performance. We worked on a database with 12 different categories of news articles and used linguistic cue approaches with machine learning. We classified news based on its

source and its previous history (such as domain name and/ or author name) with bimodal CNN and LSTM. Through reputable news sources, the model classifies reliable news articles with an accuracy of 99.7% on the training data and 97.5% on test data. However, as a piece of fake news can still be published on a reputable domain, we still had to consider other parameters such as news headlines.

Nida Aslam et al [5] proposed an ensemble-based deep learning model to classify news as fake or real using a LIAR dataset. Due to the nature of the dataset attributes, two deep-learning models were used. For the textual attribute "statement," Bi-LSTM-GRU-dense deep learning model was used, while for the remaining attributes, the dense deep learning model was used. Experimental results showed that the proposed study achieved an accuracy of 0.898, a recall of 0.916, a precision of 0.913, and an F-score of 0.914, respectively, using only statement attributes. Moreover, the outcome of the proposed models is remarkable when compared with that of the previous studies for fake news detection using the LIAR dataset. Despite the significant results achieved by the proposed study, there is still room for improvement. The model needs to be investigated using other fake news datasets.

From the aforementioned studies, the research gap present in each manuscript motivates me to study the hybrid methods for fake news detection. Among the various hybrid methods that exist in the literature, those that model the social graph that spreads the news, or the user and news source features (profile), cannot be applied when only the text of the news is available. From the hybrid methods that examine only the textual content of news, the combination of LSTM and CNN has shown promising results. However, so far, LSTMs have been used for providing word embeddings and CNN for doing the final classification. Accordingly, the research proposed Naïve Bayes, a passive-aggressive classifier and Deep Neural Network (DNN) to be implemented to detect fake news with multivariate missing values to tackle the issue [16]. The passive-aggressive classifier is used for training the ISOT and LIAR dataset and the Naïve Bayes was used to test the model for detecting fake news. Finally, DNN is used for validation purposes which efficiently classify fake and real news. A conclusion to the analysis of the related literature is that fake news has played a significant role in many real-time disasters. In order, manual interventions are of no use due to the multiple datasets which contain information sharing on the internet. Machine learning techniques have experimented on a range of datasets and deep learning techniques are still to be fully evaluated on fake news detection and related tasks. Table 1 illustrates the comparison of the state-of-the-art techniques.

3 Research proposed methodology

The usage of social media platforms has been increasing day by day. Due to the absence of regular supervision and oversight, and by lack of accountability, violators have been able to run uncontrolled and propagate false information. Therefore, the detection of fake news from social media is important for the current situation. Consequently, the article proposed novel fake news detection based on multivariate missing variables with useful features using classifier techniques.

The architecture of the proposed methodology is illustrated in Fig. 1. Intially, the latent variable formation is handled by key feature words with the content's propensity scores of the opinions on sentiment analysis using a lexicon-based scoring algorithm. Further, the paper, proposed a new multiple imputation strategy for handling multivariate missing variables in the ISOT of social media data. The researchers have used this imputation approach to bring in

Table 1 Comparison table for the state of art techniques

References	Objective	Pros	Cons
Sahoo et al [15]	Automatic fake news detection approach in chrome environment using machine learning and deep learning classifiers on which it can detect fake news on Facebook.	This analyzes both user profile and news content features	
Shu et al [23]	Fake news data repository FakeNewsNet contains two comprehensive data sets with diverse features in news content, social context, and spatiotemporal information.	FakeNewsNet would benefit the research community by studying various topics such as (early) fake news detection, fake news evolution, fake news mitigation, and malicious account detection.	This only shows the metadata of 5000 users in the provided link due to the space limitation.
Duan et al [24]	Multiple machine learning and deep learning algorithms to obtain the highest accuracy for detecting fake news patterns	Testing set on TIRA showed 70% accuracy on our highest-achieving model	Not all combinations of representations increased the accuracy. NER in combination with other representations is not suitable for the use of SVMs or ANNs.
Xu et al [8]	Term frequency-inverse document frequency (tf-IDF) and Latent Dirichlet Allocation (LDA) topic modelling is inefficient in detecting fake news,	Domain reputations and content characteristics of fake and real news will provide key insights for effectively detecting fake news on social media.	The difference in the topics and word embeddings shows little or subtle difference between fake and real news
Kumar et al [26]	Compare the CNN, LSTM, bidirectional LSTM model, CNN+LSTM ensembled network, and bidirectional LSTM+LSTM ensembles model to gather fresh instances such as PolitiFact and build multiple information for the identification of real and fake news	Use this research to combat fake news stories and neutralize the drastic effects of false information on a large scale	This research focussed mostly on the sentiments of news stories while not paying continuous attention To the credibility of the news sources themselves, due to resource limitations. Second, the classification models were unable to identify the semantic transition of real news to fake news
Nasir et al [18]	Proposes a novel hybrid deep learning model that combines convolutional and recurrent neural networks for fake news classification.	Approximately 100% accuracy on the ISOT dataset,	Over-fitted models expose high complexity and examine a lot more information than is probably needed to reach a decision.
Kaliyar et al. [21]	Deep convolutional neural network (FNDNet) for fake news detection.	This research will assist researchers in broadening the understanding of the applicability of CNN-based deep models	A hybrid approach can create more impact in the case of multi-label datasets.
Choudhari et al [16]	The linguistic model with neural-based sequential learning is proposed for fake news detection	Measure the importance of extracted feature sets as well and readability is considered the most rarely used feature out of all extracted feature	Extensive features/parameters for model performance are lacking

Table 1 (continued)

References	Objective	Pros	Cons
Yuan et al [6]	Structure-aware Multi-head Attention Network (SMAN) for fake news detection	SMAN can not only improve the detection performance but also significantly reduce the time required for the detection	–
Dang et al [27]	Deep learning models and related techniques applied to sentiment analysis for social network data	Good balance between accuracy and CPU runtime.	Address the problem of aspect sentiment analysis
Abdullah et al [7]	Multimodal approach with Convolutional Neural Network (CNN) and Long Short-Term memory (LSTM) to classify the fake news articles	The model classifies reliable news articles with an accuracy of 99.7% on the training data and 97.5% on test data	Fake news can still be published on a reputable domain, we still had to consider other parameters such as news headlines.
Nida Aslam et al [14]	Ensemble-based deep learning model to classify news as fake or real.	The outcome of the proposed models is remarkable when compared with that of the previous studies	Multiple dataset analysis is necessary.

missing values of sub-ratings on content posts in social media and the news. This suggested technique generates multiple imputed values for missing text content values based on sentiment analysis of text contents and verified user opinions. When compared to deletion and other imputation strategies, this imputation method is used to reduce inaccuracy and bias in the results of the study on online data. Subsequently, the useful features have to be extracted from the social media or news contents using Term Frequency And Inverse Document Frequency (TF-IDF). Finally, the classification of the detection of fake news has to be made, which has been done using multiple techniques, such as Naïve Bayes, passive-aggressive classifier, and Deep Neural Network (DNN). The description of the proposed method is detailed in further sections.

3.1 Data collection

For the identification of false news, numerous datasets have been provided. Having a big dataset for training the model is among the most important prerequisites for employing neural networks. Researchers employ two datasets for training deep models in this paper: ISOT false news and LIAR, both of which include a significant amount of documents. The size of ISOT's news statements is medium to long, whereas LIAR's is modest. There is both actual and fake news in the databases. The real news was gathered via reading articles from Reuters.com, while false news was gathered from untrustworthy sources highlighted by Politifact; within that case, the researchers additionally used data from FakeNewsNet.

3.2 Initial structure

Let X be the overall user opinions score, Y_k be the kth attribute user opinions score, $k = 1, …, K$ and W be the user opinions. Assume that X and W are observed over the entire data set but Y_k are subject to missingness. Now let R_k denote a set of indicators of the missingness Y_k. R_k takes
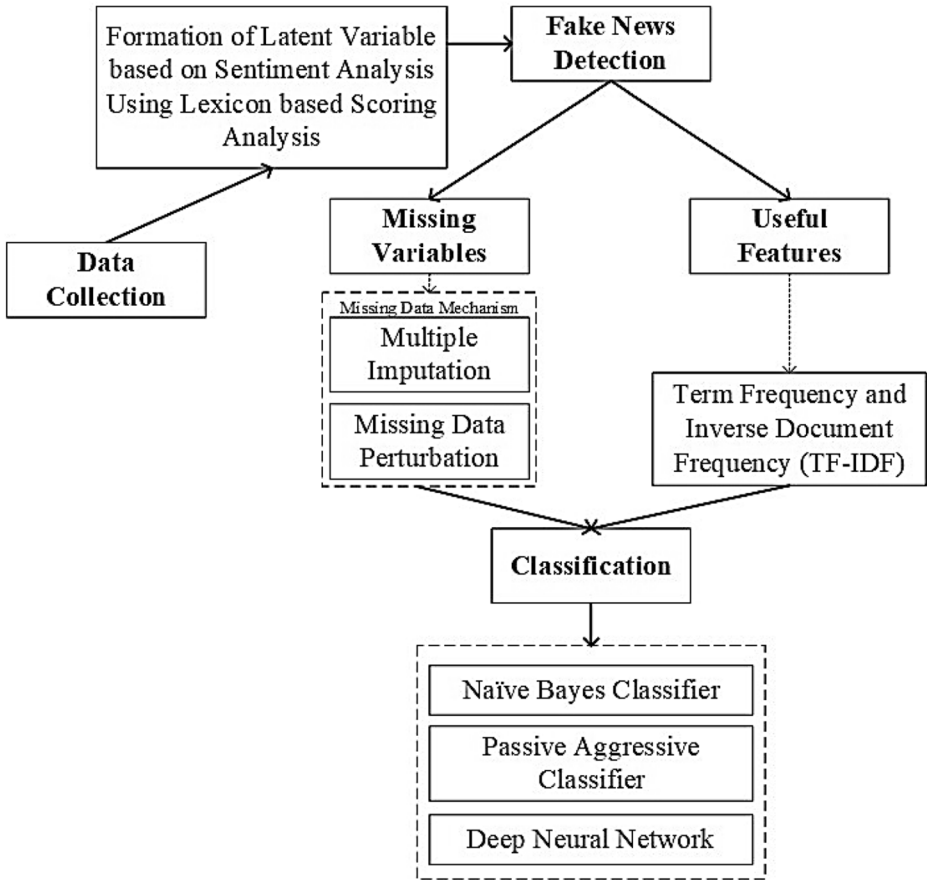
Fig. 1 Architecture of the proposed method

the value of 1 if Y_k is observed and 0 otherwise. Let Y_{obs} and Y_{mis} be the observed/ missing values of $Y = (Y_1 \dots, Y_k)$.

Assume MAR missing mechanism such that:

$$P(R|Y,X,W) = P(R|X,W) \tag{1}$$

Where $R = R_1, \dots R_k$. Thus, the missing values can be generated without specifying the response model.

3.2.1 Formation of a latent variable Z

Consider the predictive distribution of the Y conditional on X and W as the imputation model:

$$P(R|Y,X) \tag{2}$$

One simple method for generating the imputed values for Y is to use only X because W is unstructured, making it difficult to estimate the predictive distribution Eq. (2). However, this method may cause efficiency loss or biased estimation results when the user opinions are

strongly correlated with W given X. To avoid these problems, it needs to consider W, which makes fuller use of the observed text data.

For this purpose, this study needs to structure the unstructured W by creating a latent variable Z that can substitute W. This property of W is formulated as follows:

$$P(Y,X|W,Z) = P(Y,X|Z) \tag{3}$$

where $Z = g(W)$ is a numerical variable constructed from W through a converting mechanism $g(.)$. Text clustering, scoring analysis or similar statistical methods can be applied to obtain Z. Since Eq. (3) is not testable in practice, this study generally assumes that the newly created variable Z from the converting algorithm will be sufficient statistics of W, that is, Eq. (4) is satisfied.

The latent variable Z could be continuous or categorical, but we want this latent variable Z to satisfy the following desirable property:

$$P(Y,X,R|Z = z_1) \neq P(Y,X,R|Z = z_2) \text{ for } z_1 \neq z_2 \tag{4}$$

Equation (3) implies that the latent variable Z is informative in explaining the joint distribution of (Y, X, and R). Equation (4) can be decomposed into two properties as follows:

$$P(Y,X|Z = z_1) \neq P(Y,X|Z = z_2) \text{ for } z_1 \neq z_2 \tag{5}$$

$$P(R|X,Z = z_1) \neq P(R|X,Z = z_2) \text{ for } z_1 \neq z_2 \tag{6}$$

If at least one of two properties is satisfied, we can conclude that Z satisfies Eq. (4). If Z fails to meet Eq. (5) or Eq. (6), the imputation estimates may have larger standard errors than the estimates obtained using a simpler model $P(Y| X)$. On the other hand, if Z satisfies one of two conditions, adding an extraneous variable will not lead to additional biases, and the predictive distribution Eq. (3) will provide more robust results compared to when using a simpler model $P(Y| X)$, which is possibly exposed to the omitted variable problem. Thus, it is important to check if the created variable Z follows the required Eq. (5) or Eq. (6) to validate the new imputation method.

3.2.2 Lexicon-based scoring analysis

This study proposes a new method that does not require any training procedure or similarity score computation to construct a latent variable Z for our imputation. First, the proposed method chooses key feature words based on the parts of speech (POS) in the text content and then compute users' propensity scores of the opinions based on sentiment analysis. The proposed method essentially uses a lexicon-based scoring algorithm. As the false text content with users' subjective opinions is well represented through specific POS such as an adjective, adverb, and verb, this study extracts these POS and then makes scores on these selected features. The detailed algorithm is given below:

Step 1: Assign POS on words within each user's opinions and the main text contents using the natural language processing (NLP) toolkit.
Step 2: Select key feature words categorized as verbs, adverbs and adjectives.

Step 3:   Assign one of three values {positive; neutral; negative} to each keyword based on a dictionary of sentiment words, called a sentiment lexicon.

Step 4:   Compute a propensity score for user opinions i:

$$S_i = \left(C_{pos,i} + C_{neg,I}\right)/(\alpha + C_i) \tag{7}$$

In the realm of fake news detection, we regard each word feature as a treatment and each news sample as a subject. Then we formally define the propensity score as above. Where, $(\alpha > 0)$ is a tuning parameter, C_i is the number of key feature words and $C_{pos,i}$ and $C_{neg,i}$ is the number of positive/negative key feature words. The tuning parameter in step 4 adjusts the normalized degree. For short texts, the propensity scores will be more easily influenced by the sentiment of one key feature word than those of long texts. Thus, this tuning parameter is used to adjust the balance between short and long text data.

This procedure of constructing Z requires text-mining knowledge. There are many packages in Python that provides the proposed scoring algorithm. The NLP toolkit can be used to determine a POS for each word in text data.

Grammatical evidence To inspect real news and fake news, the grammatical feature is an important factor that is extracted through parts of speech (POS) tag evidence features. Out of all POS features, for the targeted problem noun, verbs, adjectives, and pronouns are viable features to define their authenticity. These features are designed to apprehend the deceiver cues in writing style to differentiate fake news. Details are shown in Table 2.

For a given news, defining the grammatical features references as noun count (x_{nou}), verb count (x_{verb}), adjective count (x_{adj}), adverb count (x_{adv}) and pronoun count (x_{pro}).

$$X_{gr} = \left\{x_{nou}, x_{verb}, x_{adj}, x_{adv}, x_{pro}\right\}$$

3.3 Missing variables

Social media has become a widespread element of people's everyday life, which is used to communicate and generate content. Among the several ways to express a reaction to social media content, "Likes" are critical. Indeed, they convey preferences, which drive existing markets or allow the creation of new ones. Nevertheless, the situation does not allow to give a dimension to the target universe of the respondents, leading to caution in the management of the missing values. Regarding the statistical analysis, the treatment of missing data represents a relevant problem.

Table 2 Grammatical-based features

Grammatical based Features	Example	Representation
Noun	Refugees can be singled out for a higher level of review	Refugees
Verb adjective	based on their age, nationality, or gender.	Higher level
adverb		gender
verb		Singled out

3.3.1 Missing data mechanism

This study develops a new multiple imputation method that can handle multivariate missing variables in social media or news data. Specifically, this imputation method is applied to missing values of user opinions on the posts or the news updates in social media texts. Based on the sentiment analysis of social media or news text data and observed user opinions of likes and comments, this method creates multiply imputed values for missing values of the post or news contents and the user opinions. It is important to first distinguish between missing data patterns, which describe observed and missing values, and missing data mechanisms, which relate to the probability of missingness. Common missing data patterns in surveys typically include unit nonresponse, where a subset of participants do not complete the survey, and item nonresponse, where missing values are concentrated on particular questions. In opinion polls, nonresponse may reflect either refusal to reveal a preference or lack of a preference.

Let $Y = y_{ij}$ be a ($n \times K$) dataset with each row $y_i = (y_{i1}, \ldots, y_{ik})$ the set of y_{ij} values of feature Y_j for example i. Let Y_{obs} define the observed values of Y and Y_{mis} define missing values. Let M define the missing data identity matrix $M = m_{ij}$, where $m_{ij} = 1$, if y_{ij} is missing and $m_{ij} = 0$, if y_{ij} is nonmissing. The missing data mechanism is missing completely at random (MCAR) if the probability of missingness is independent of the data, or

$$f(M|Y, \varphi) = f(M|\varphi) \ \forall Y, \varphi \tag{8}$$

where φ denotes unknown parameters. The MAR assumption is less restrictive than MCAR in that the probability of missingness depends only on the observed data, $f(M| Y, \varphi) = f(M| Y_{obs}, \varphi)$ for all Y_{mis}, φ. The missing not at random (MNAR) assumption is that the probability of missingness may also depend on the unobserved data, $f(M| Y, \varphi) = f(M| Y_{obs}, \varphi)$ for all Y_{mis}, φ. Researchers typically assume data are missing at random (MAR), which mitigates the identifiability problems of MNAR because the probability of missingness depends on data that are observed on all individuals. It is important to grasp these different types of missing data from a statistics point of view. The type of missing data will influence the thanks to accommodate filling within the missing values and detect missing values, and do some basic imputation and detailed statistical approach for handling missing data. Before, joint into code, it's important to grasp the sources of missing data. To investigate the multi-variate processes, data cleaning/preparation includes renaming the provided dataset and deleting columns, among other things. Based on Eq. (3), this study uses the widely used Multiple Imputation Chain Equation (MICE) for multiple imputations. The methods and techniques for cleaning data will differ depending on the dataset. The fundamental purpose of data cleaning is to find and eliminate mistakes and abnormalities to improve the quality of the dataset. The primary goal of data cleaning is to detect and remove errors and anomalies to increase the price of data in analytics and better noesis.

3.3.2 Multiple imputation

A modified imputation procedure is summarized below:

Step 1: Convert W to a latent variable Z using sentiment analysis on text data.
Step 2: Specify the imputation model for each missing variable with the fully observed data:

$$P\left(Y_k | Y_{-k}, X, Z, \delta = 1; \theta_k^{*(0)}\right), \ for \ k = 1, \ldots, K \tag{9}$$

where Y_k denotes the collection of Y except Y_k, $\delta = \Pi_{k-1}^K R_k$, and $\widehat{\theta}_k^{*(0)}$ denotes the parameter estimates of the imputation model for the kth missing variable on the complete cases.

Step 3: For given imputed values $Y^{*(t-1)} = \left(Y_{obs}, Y_{mis}^{*(t-1)}\right)$, the ith iteration of MICE is a Gibbs sampler that sequentially generates:

$$\widehat{\theta}_1^{(t)} \sim P\left(\theta_1 | Y^{*(t-1)}, X, Z\right) \tag{10}$$

$$Y_1^{*(t)} \sim P\left(Y_1 | Y_{-1}^{*(t-1)}, X, Z, \widehat{\theta}_1^{(t)}\right) \tag{11}$$

$$\widehat{\theta}_2^{(t)} \sim P\left(\theta_2 | Y_1^{*(t)}, Y_{-1}^{*(t-1)} X, Z\right) \tag{12}$$

$$Y_2^{*(t)} \sim P\left(Y_2 | Y_1^{*(t)}, Y_3^{*(t-1)}, \ldots, Y_k^{*(t-1)} X, Z, \widehat{\theta}_2^{(t)}\right) \tag{13}$$

$$\widehat{\theta}_K^{(t)} \sim P\left(\theta_K | Y_{-K}^{*(t)}, Y_K^{*(t-1)} X, Z\right) \tag{14}$$

$$Y_K^{*(t)} \sim P\left(Y_K | Y_{-K}^{*(t)}, X, Z, \widehat{\theta}_K^{(t)}\right) \tag{15}$$

Step 4 Iterate step 3 for large enough t until we have convergence.
Step 5 Independently repeat steps 3 and step 4 $M(>1)$ times so that this research creates M imputed data sets as the final imputation output.

Consider an imputation estimator constructed from the multiply competed for datasets. Let Q be an estimand defined with a known link function h, where $Q = h(Y, X)$. On the multiply completed data sets, the imputation estimator of Q is:

$$Q = M^{-1} \sum_{m-1}^{M} \widehat{Q}_m = M^{-1} \sum_{m-1}^{M} h\left(Y_m^*, X\right) \tag{16}$$

where Y_m^* denotes a set of imputed values of the mth completed data set, and $\widehat{Q}_m = h\left(Y_m^*, X\right)$ is the estimates obtained from the mth completed data set. The variance of the imputation estimator can be estimated using the variance formula, where:

$$\widehat{V}\left(\widehat{Q}\right) = \widehat{A} + \frac{M+1}{M}\widehat{B} \tag{17}$$

With the average of within-imputation variances:

$$\widehat{A} = M^{-1} \sum_{m-1}^{M} \widehat{V}\left(\widehat{Q}_m\right) \tag{18}$$

and the between-imputation variance:

$$\widehat{B} = (M-1)^{-1} \sum_{m-1}^{M} \left(\widehat{Q}_m - \widehat{Q}\right)\left(\widehat{Q}_m - \widehat{Q}\right)^T \tag{19}$$

3.3.3 Missing data perturbation

After randomly splitting each dataset, this study perturbs the training data so that the proportion of missing values in the set of categorical features Y_{cat} follows $\delta = \{0.1,\ 0.2,\ 0.3,\ 0.4\}$ according to the MCAR mechanism

$$\Pr(M = 1 | Y_{cat}.\varphi) = \delta \ \textit{for all} \ Y_{cat} \tag{20}$$

This research uses missing-data perturbation to study the impact of larger amounts of missing data; however, it is also a form of dropout noise that can be used to control overfitting during the training process and improve the generalizability of the model.

3.4 Useful feature extraction

Feature extraction solves the problem of finding the most compact and informative feature set. For classification and regression problems, defining feature vectors remains the most common and convenient method of data representation. The commonly used feature extraction method is driven by the size of the data table. With the increasing efficiency of data storage, the size of data tables is also increasing. Extracting effective features from text and avoiding useless data processing is the key to experimental research.

There are various "sizes" and "forms" of data in text information. An important point when extracting features in the text is structured data. Generally, the raw data that has not been processed in the text is converted into structured data. The process of tools to get valid information is called information extraction. In the same way, in this research, extracting useful features from the actual news content is a challenging task because fake news spreaders could make the content of the fake news look like real news.

3.4.1 Term frequency and inverse document frequency (TF-IDF)

In this work, the proposed method used term frequency and inverse document frequency (TF-IDF) to identify the useful features of news content. Existing work on fake news detection can be divided into two categories: unimodal and multimodal. To learn a more general joint

representation, a minimax game is set up between the event discriminator and feature extractors. Scaled-dot product attention has been applied to the fields of natural language processing (NLP) and computer vision (CV). In NLP and CV, the extraction of corresponding features, such as textual features and visual features, is a fundamental task, and it is also a key step in fake news detection. In this section, we review the related work on unimodal fake news detection, and hybrid fake news detection with a useful feature extraction technique is introduced. Term frequency works by looking at the frequency of a particular term you are concerned with relative to the document. Inverse document frequency looks at how common (or uncommon) a word is amongst the corpus. To summarize the key intuition motivating TF-IDF is the importance of a term is inversely related to its frequency across documents.TF gives information on how often a term appears in a document and IDF gives information about the relative rarity of a term in the collection of documents. TF-IDF technique is used to produce a composite weight for each term in the document which is called *tf-idf* weight. Calculating *tf-idf* weight has great importance in information retrieval and text mining tasks as it determines the significance of a term or word in a document as well as in a corpus.

$$tf-idf_{t,d} = tf_{t,d} \times idf_t \tag{21}$$

In Eq. 21, t means a term and d refers to a document. The term frequency $f_{t,d}$ means the measure of the frequency for a particular term t in a document, in other words, how many times term t appeared divided by the total number of terms in the document and inverse document frequency idf_t is the logarithm of a total number of documents in the corpus divided by the number of documents where term t appears. idf_t the measure helps in knowing the importance of the term t. From this method, this study observed that the derived weight matrix represents a large number of term features.

3.5 Classification

This study has thoroughly used ML to build the proposed models. The task of choosing the classifiers that emerged from the suitable properties of algorithms is, therefore, developed as a hybrid classifier model. As Naïve Bayes is popular for its multi-class prediction, it was picked up for its ease and robustness in predicting the class of the text. One of the problems with other methods is when new samples are collected, a model must be retrained to predict the output for new data. This is overcome by using a passive-aggressive classifier which trains the model incrementally, allowing modifications of the parameters only when needed, while discarding these updates when they don't alter the equilibrium. This study focused on the problem based on both conventional and deep learning architecture. A deep neural network was used to increase the efficiency in the identification of fake news. The following paragraphs dive deeper into each algorithm.

3.5.1 Naïve Bayes classifier

A naïve Bayes classifier assumes that features are statistically independent of one another. It explicitly models the features as conditionally independent given the class. Because of the independence assumption they are highly scalable and can quickly learn to use high-dimensional features with limited training data. Given a data point $\overrightarrow{x}$ of n features, Naïve Bayes predicts the class C_k for a data point, according to Bayes' theorem it can be factored as,

$$p\left(C_j|\overrightarrow{x}\right) = \frac{p\left(\overrightarrow{x}|C_j\right)p\left(C_j\right)}{p\left(\overrightarrow{x}\right)} = \frac{p\left(x_1,\ldots,x_n|C_j\right)p\left(C_j\right)}{p\left(x_1,\ldots,x_n\right)} \tag{22}$$

If this study uses the simplifying conditional independence assumption, that given a class (positive or negative), the words are conditionally independent of each other. Due to this simplifying assumption, the model is termed "naïve".

$$p\left(C_j|\overrightarrow{x}\right) = \frac{\left(\Pi p\left(x_i|C_j\right)\right)p\left(C_j\right)}{p\left(\overrightarrow{x}\right)} \tag{23}$$

Here the x_i represents the individual words of the document. The classifier outputs the class with the maximum posterior probability.

Laplacian smoothing If the classifier encounters a word that has not been seen in the training set, the probability of both the classes would become zero and there won't be anything to compare between. This problem can be solved by Laplacian smoothing.

$$p\left(x_i|C_j\right) = \frac{Count\left(x_i\right) + k}{(k+1)\left(No\ of words in class\ C_j\right)} \tag{24}$$

Usually, k is chosen as 1. This way, there is an equal probability for the new word to be in either class. Since Bernoulli's Naïve Bayes is used, the total number of words in a class is computed differently. For this calculation, each document is reduced to a set of unique words with no duplicates.

As considered, this classifier is best suited for small-size datasets, this study proposed a passive-aggressive classifier due to its specific properties.

3.5.2 Passive aggressive classifier

A passive-aggressive classifier is simple and its performance has been proven to be superior to many other alternatives. Let's suppose to have a dataset where $\overline{x}$ is the datapoint and y_i is the predicted output. Given a weight vector w, the prediction is simply obtained as: $\widetilde{y_t} = sign\left(\overline{w}^T \cdot \overline{x}_t\right)$, the algorithm works generically with this update rule:

$$\overline{w}_{t+1} = \overline{w}_t + \frac{\max(0, 1 - y_t(w^T \cdot x_t))}{\|x_t\|^2 + 1/2C} y_t \overline{x}_t \tag{25}$$

Hence this classifier trains the model incrementally. These are the conventional means of algorithms whose accuracy is limited when compared with deep learning architecture.

3.5.3 Deep neural network

The researchers chose to use tokenizers for feeding into Deep Neural Network models. Python offers a simple API, where one can vectorise text corpus by converting each word into vectors or sequences of integers. It splits the text into a list of tokens where the coefficient for each

token could be based on word count. The dictionary of the tokenizer has been prepared using the train set of each dataset. The tokenizer once constructed can be fit on the raw text data. Since each different article contained different number of words, this study used padding to keep the size uniform.

There are mainly 7 layers i.e. 1 input (takes prior model input), 1 output, and 5 hidden layers. Each attribute is individually tokenized and fed to the network. Each input layer takes the tokenized value from a single attribute of the dataset. Each Input layer is fed to the embedding layer which learns all the embeddings of the word. This layer allows us to take large inputs like sparse vectors representing words. The first DNN layer defines a filter of kernel size 3. Now there are 4 × 4 weights assigned to all neurons in the same block. The output of this layer is fed to a global max pool layer which is used to compute the single max value for each of its input channels. It is a great alternative to flattening. All three outputs from the max pool layers are concatenated to form a single layer. This layer is fed to a series of Dense layers which connects all the neurons to form a network. The final output is fed to the output layer, here CNN is connected to fully connected NN for decision making as either true or false. The fully connected layer with sigmoid activation is the final layer that will reduce the vector of height 8 to 1 for prediction ("fake", "real"). The Dense layers contained Relu activation functions. Activity, kernel, and bias regularizers along with Dropout were all used as seen fit.

3.5.4 Overfitting and cross-validation

Overfitting is one of the central problems in machine learning. It arises when the model performs poorly on unseen data while giving excellent results on training data. Cross-validation is a way to overcome such an issue; it aims to test the model's ability to correctly predict new data that was not used in its training. Cross-validation shows the model generalization error and performance on unseen data. K-fold cross-validation is one of the most popular versions. In this experiment, the researchers use k-fold cross-validation to ensure that the study avoids overfitting.

4 Result and discussion

This study employs ISOT and LIAR datasets to verify the proposed models. The dataset contains a total of 44,848 news articles, of which 21,417 are real and 23,481 are fake news. This dataset was applied for the key feature analysis using sentiment analysis and further. Three evaluation metrics, namely accuracy, F1-score and recall are used to evaluate the performance of these proposed models. One of the major challenges of performing classification on this dataset is handling missing values. Therefore, multiple classification techniques were performed to handle and get an accurate result, this classifies the result with the differences between before feature extraction and after feature extraction, and the dataset with imputation and dataset without imputation model is also evaluated.

The experiment of this paper was conducted on a system with a windows 10 operating system, a memory capacity of 6GB DDR3, and an Intel Core i3@ 3.5GHz would be the processor. With this system configuration, the proposed method is implemented using the software of Python.

4.1 ISOT dataset

The dataset contains two types of articles fake and real News. This dataset was collected from real-world sources; the truthful articles were obtained by crawling articles from Reuters.com (News website). As for the fake news articles, they were collected from different sources. The fake news articles were collected from unreliable websites that were flagged by Politifact (a fact-checking organization in the USA) and Wikipedia. The dataset contains different types of articles on different topics, however, the majority of articles focus on political and World news topics. The dataset consists of two CSV files. The first file named "True.csv" contains more than 12,600 articles from reuter.com. The second file named "Fake.csv" contains more than 12,600 articles from different fake news outlet resources. Each article contains the following information: article title, text, type and the date the article was published on. The data collected were cleaned and processed, however, the punctuations and mistakes that existed in the fake news were kept in the text.

Figure 2 depicts the ISOT dataset word cloud image of fake and real news. After preprocessing, the number of samples was reduced to 44, 143 (22, 727 fake and 21, 416 real). The similarity between real and fake news post word usage in the ISOT dataset suggests the real and fake news posts are not well separated by word frequencies alone. Most words highlighted in the fake news are prenatal, insurance, Shimkus and republicans and for true words, it seems to be refugees, trump, united, and administration.

4.2 LIAR dataset

LIAR is a publicly available dataset for fake news detection. A decade-long of 12.8 K manually labelled short statements were collected in various contexts from POLITIFACT.COM, which provides a detailed analysis report and links to source documents for each case. This dataset can be used for fact-checking research as well. Notably, this new dataset is an order of magnitude larger than previously largest public fake news datasets of similar type. The LIAR dataset4 includes 12.8 K human-labelled short statements from POLITIFACT.COM's API, and each statement is evaluated by a POLITIFACT.COM editor for its truthfulness.

To visualize the news content of the dataset, word cloud representations are used. Word cloud representations depict the frequency of the terms in a specific dataset of LIAR. This
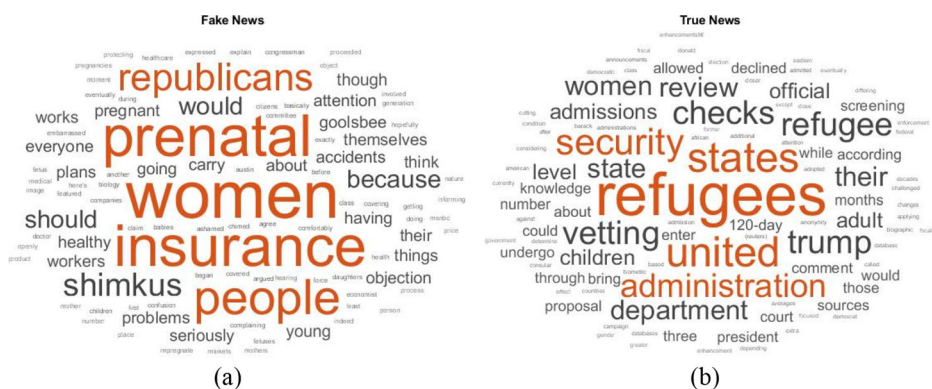

(a)   (b)

Fig. 2 ISOT dataset Wordcloud of fake and true words

system draws some exciting conclusions from the word cloud representation shown in Figs. 3a, b. Real news word cloud represents important entities that occurred in actual events like the Americans, texas, Obama, and state, while fake news word cloud highlights fake entities, such as Obama, state, president, tax and republican. It was interesting to observe that fake news generally uses appealing headlines and does not have specific content that denotes real news. It is observed that both fake and real news articles are generally related to a political campaign.

4.3 Spam detection with machine learning classifiers

The passive-aggressive classifier was trained on 70% of the ISOT dataset and the Naïve Bayes was used to test the model using the remaining 30% for detecting fake news. Similarly, for the LIAR dataset, 80% was used for training and 20% was used for testing. If the sentiment score is positive, then the sample is tagged as positive for the sentiment attribute, otherwise Negative is assigned. They have achieved good timeefciency performance as well as better classification performance. The model achieves an average accuracy of 90% over a five-fold cross-validation.

Table 3 illustrates the number of attributes in the dataset before and the number of attributes in the dataset after feature extraction for two datasets. The attributes used after feature extraction for the ISOT dataset are title, text subject, date, noun, verb, and adjective. Similarly, for the LIAR dataset, the attributes used after feature extraction are id, title, statement, subject, speaker, speaker job title, state info, party affiliation, noun count, verb count, adjective count and adverb count.

The confusion matrix of the DNN classifier is represented in Fig. 4. True-positive value and true-negative value of the random forest classifier are high means if the news is false, it hows false, and in the case of true news, it predicts the same. Here the output class predicts that fake words of 95% and real words of 67.3% respectively.

4.4 Error rate of dataset with MCAR perturbation

For varying amounts of MCAR-perturbed qualitative training data, error rates on the test set with and without missing data imputation (x-axis). The standard deviation of test set errors of ANNs trained using different convergences is used to calculate prediction intervals for ANNs.
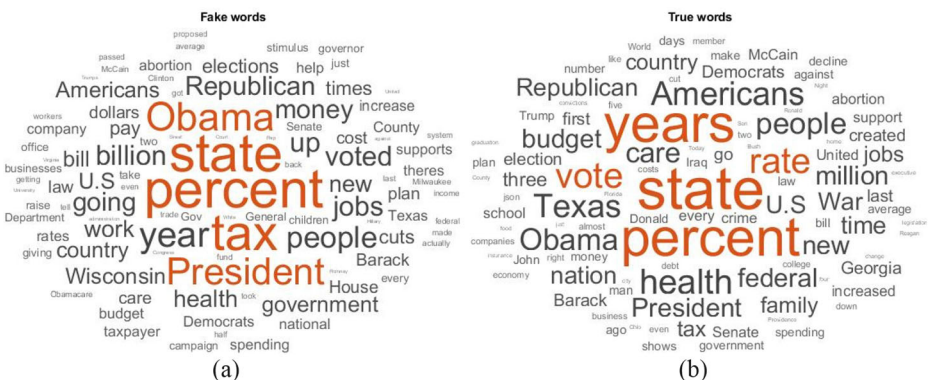

Fig. 3 LIDAR dataset Wordcloud of fake and true words

Table 3 Attributes in the dataset after feature extraction

ISOT		LIAR	
Attributes in the dataset	Attributes After Feature Extraction	Attributes in the dataset	Attributes after feature extraction
Title	Title	Id	Id
Text	Text	Title	Title
Subject	Subject	Statement	Statement
Date	Date	Subject	Subject
	Noun	Speaker	Speaker
	Verb	Speakers job title	Speakers job title
	Adjective	State info	State info
	Adverb	Party Affiliation	Party Affiliation
			Noun Count
			Verb Count
			Adjective Count
			Adverb Count

Prediction intervals for passive aggression follow the variance caused by increasing the maximum depth of the Naïve Bayes, and for deep neural networks, the result of several features evaluated are classified with fewer test set errors. The error rate of a dataset with missing data imputation was calculated with each classification technique proposed in this study, with results proportional to the test set error classified with MCAR perturbation as shown in Fig. 5(a). The error rate of a dataset without missing data imputation results in the non-compensative test set error with MCAR perturbation as illustrated in Fig. 5(b).

4.5 Evaluation metrics

The evaluation metric in this experiment is classification accuracy. Accuracy is the ratio of correct predictions to the total number of samples. Apart from accuracy, other performance
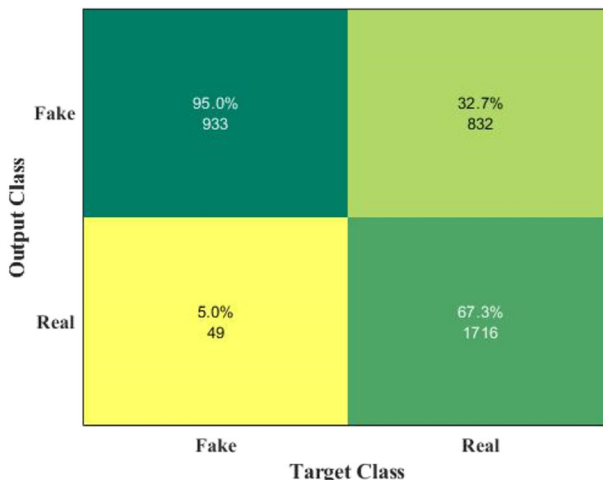
Fig. 4 Confusion matrix
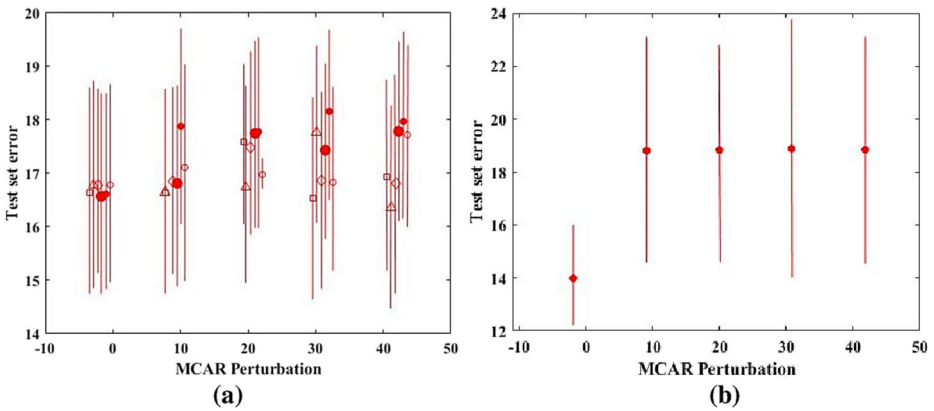
Fig. 5 Error rate of dataset (a) with and (b) without missing data imputation

measures, that is, True Positive Rate (TPR) also known as Recall, and F1 measures, are calculated based on equations below, respectively.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{26}$$

$$Recall = \frac{TP}{TP + FN} \tag{27}$$

$$F_1 = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \tag{28}$$

Where TP represents the number of True Positive results, FP represents the number of False Positive results, TN represents the number of True Negative results, and FN represents the number of False Negative results.

The performance of the models was evaluated based on the predicted outcome values using common statistical measures. In this study, the evaluation metrics of Mean Absolute Percentage Error (MAPE), Residual Sum Of Squares (RSS) and AUC are computed for the forecasting models which is shown in Eq. (29).

$$MAPE = \frac{1}{N} \Sigma_{i=1}^{N} \left(\frac{\left(\left| \hat{y}_i - y_i \right| \right)}{y_i} \right) *100\% \tag{29}$$

AUC is precisely the area under the ROC curve. An excellent system has an AUC close to 1 (it can perfectly distinguish between all fake and true news correctly), while a poor system has an AUC close to 0 (it would be considering all fake news as true and all true news as fake). AUC is more statistically consistent and more discriminating than accuracy, and it is usually applied in imbalanced classification problems, as is the case of fake news detection, where the number of ground truth fake news articles and true news articles have a very imbalanced distribution.
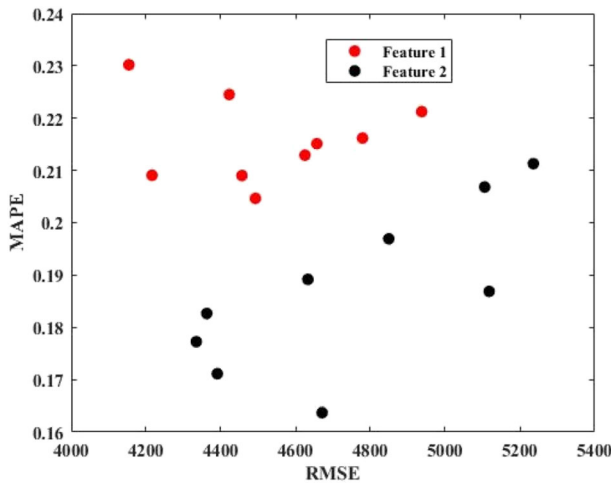
Fig. 6 MAPE performance analysis with RMSE

The MAPE performance result is demonstrated in Fig. 6. According to the result, as the number of features increases, the test error decreases. This indicates that the regression model gets more information from the additional features. However, after selecting the best features, adding more features decreases the model performance and increases the test error. This indicates that adding more features causes an overfitting problem due to the shared information within the features. Compared with selecting all 15 features, according to the result, selecting features reduces the test MAPE by about 5 cycles and test MAPE by about 0.23%. Therefore, in the following study, these selected features are used as the input features for the classification model.

Figure 7 depicts the performance analysis of the residual sum of square analysis. The residuals as an upright line connecting the actual values to the predicted value (red traces in the plot). The residual on that given datapoint is 0.5. However, if the scale meters, then that same datapoint has a residual of 7000. As calculated, the statistical analysis of the sum of all errors
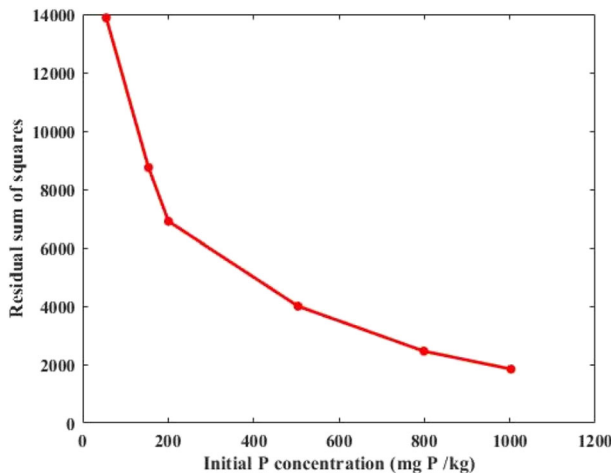
Fig. 7 Residual sum of square analysis
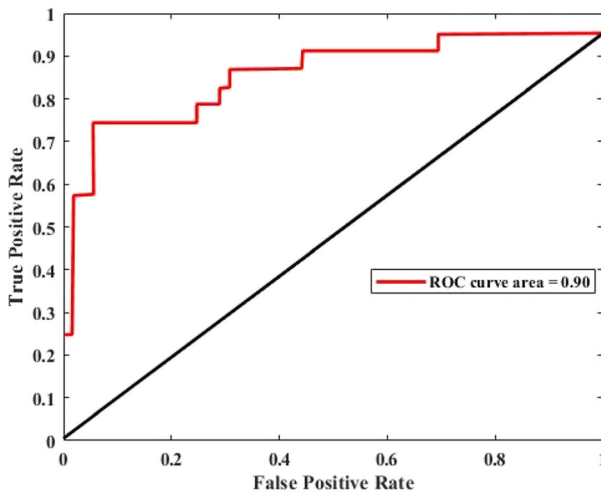
Fig. 8 ROC curve for fake news classifier

(the sum of residuals) is resulting in 0. This is because errors can positive or negative, as well the model underestimates and overestimates. By summing up the errors, all errors compensate for each other. This is a fundamental characteristic of the regression line and the method of least squares.

Figure 8 shows the Receiver Operational Characteristics (ROC) curve. The ROC curve focuses on the trade-off between the False Positive Rate (FPR) and the True Positive Rate (TPR) by adjusting the classification of each method. This proposed classifier got the highest accuracy score of 90%.

4.6 Outcome of classification

Table 4 represents the description of classification from the LIAR and ISOT dataset using the proposed classifier in view of the evaluation of accuracy, recall and F1 score. From this analysis, the dataset with barely true statements obtained a maximum of 26% as calculated by accuracy than other evaluation metrics, in before feature extraction. In after feature extraction, results in a maximum of 41% of recall as compared with others. In false statement analysis, the recall obtained a maximum of 49% in the evaluation before feature extraction and 50% in the evaluation after feature extraction. In a statement of half-true analysis, the accuracy results

Table 4 Classification report of LIAR dataset for proposed classifiers

Classifier	Label	Before feature extraction			After feature extraction		
		Accuracy (%)	Recall (%)	F1 Score (%)	Accuracy (%)	Recall (%)	F1 Score (%)
Multiple Classifier (Naïve Bayes, Passive Aggression,DNN)	Barely True	26	14	23	40	41	40
	False	24	49	30	43	50	47
	Half True	27	21	22	42	46	44
	Mostly True	27	28	30	39	51	43
	True	30	20	24	40	28	35

better with 27% of evaluation before feature extraction and in after feature extraction evaluation, recall obtains a maximum of 46%. In a mostly true statement, the F1 score obtains at 30% before feature extraction and 51% of the evaluation was obtained by recall in after feature extraction. Finally, the accuracy of 30% is achieved as a maximum before feature extraction for the true statement and 40% is the maximum accuracy obtained in after feature extraction. From this analysis, the analysis concludes that the calculation of multiple classifiers after feature extraction results better than before feature extraction.

The faster the training and prediction, the less time cost. The loss curves of the multiple classifications implemented in this study are shown in Fig. 9. It is obvious from the figure that the training loss and prediction loss were evaluated with similar training iterations of the data network. However, the loss of training and prediction of the classifiers have differed with fewer points of loss. Eventually, the prediction loss has been evaluated with minimum loss of curve as compared with the training loss.

To promote the work in this domain, more experiments were performed with the hybrid model trained on the ISOT dataset and tested on the LIAR dataset using the same configuration as before. ISOT is chosen for training because it is much larger and has minimum space for improvement since many models perform above the 0.9 classification accuracy threshold. Figure 10(a) and (b) show the ability of the ISOT-trained model to generalize on another dataset and plot the training and validation accuracy and loss values over the 10 epochs. Results show that while the training accuracy and loss are optimum after 6 epochs, the validation accuracy remains almost the same in all epochs and is lower than that achieved when training (and validating) on the LIAR dataset. The final performance after 10 epochs of training and testing on the whole LIAR dataset is shown in Fig. 10(b).

4.7 Comparison of proposed and existing methods

The proposed method of this study was compared with the existing methods such as CNN + LSTM [14] and Bi+LSTM+GRU [5]. Table 5 described the comparison of proposed and existing methods with various evaluation measures. From the table, the accuracy of the proposed and
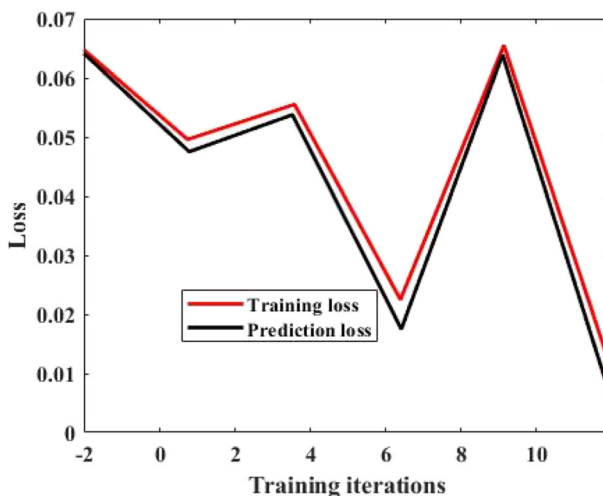

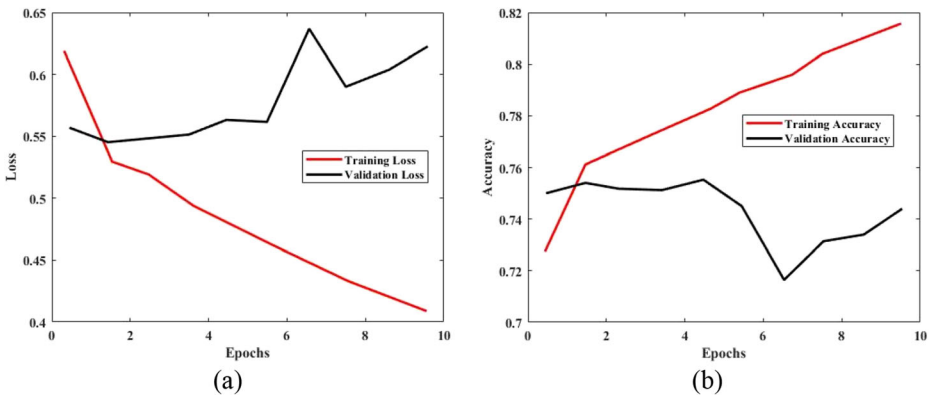**Fig. 9** Loss curve of multiple classifications

Fig. 10 Training and validation accuracy graph

existing methods is achieved to calculate maximum prediction analysis from the evaluation measures with percentage values of 99.8, 97.5, and 91.3. The recall and F1 score analysis are obtained better with the prediction of the proposed and the existing methods after the prediction analysis of accuracy. The average training time for the model without preprocessed input data was about 22 hours due to the more sequence length value. The preprocessed input data and the constraints helped reduce the sequence length to 120 from 200. The training time for the model was significantly reduced to 9 hours from 22 hours by a simple modification to the input samples and achieved the highest performance. Results showed the highest performance of 64 with 99.8% of accuracy and a corresponding F1-score of 95.6%. The results also indicate that input preprocessing helps increase performance with reduced training time. Accordingly, the time efficiency of the proposed framework is about 45 ms.

4.7.1 Ablation study of the proposed technique

To determine the relative importance of each module of the NPDNN, the article conducted a series of ablation studies on key parts of the model. The experimental comparison results are shown in Fig. 11.

This shows that these two attention mechanisms have a certain effect on our model performance. When the model used CNN-BiLSTM only, the performance of the model dropped by 2.6% to 3.2%, because the model loses very important propagation structure information. In addition, if only CNN-LSTM was used, the performance of the model dropped by about 8% on both data sets, because the model does not even read the text content of the news itself. However, the performance of Naïve Bayes on LIAR and ISOT data sets reached 8.83% and 8.92%, respectively, which proves that numerous clues can be used to detect fake news in the propagation structure.

Table 5 Comparison of proposed method and existing method

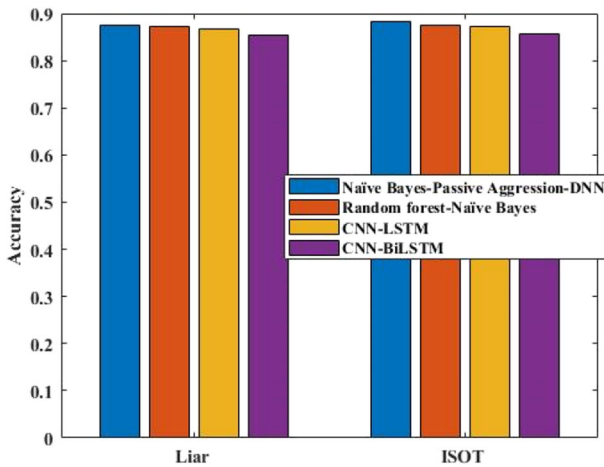| Evaluation measures	Proposed method	CNN+LSTM	Bi-LSTM-GRU
Accuracy	99.8%	99.7%	89.8%
Recall	98.9%	97.1%	91.6%
F1 Score	95.6%	93.8%	87.7%

Fig. 11 NPDNN ablation analysis in accuracy

5 Conclusion

The fundamental purpose of this study is to mitigate one of social media's drawbacks: the rapid distribution of fake news, which frequently misinforms people, produces false impressions and harms society. In recent years, an increasing number of methods for automatically detecting false news have been presented in the literature. The datasets and a set of explicit classes are two critical aspects that have a substantial influence on the accuracy of the existing models. Therefore, this study proposed several models for the detection of fake news which have been highlighted as follows.

The study of fake news detection was carried out using the ISOT and LIAR datasets with real and fake news contents from Reuters.com, Politifact and FakeNewsNet.

Initially, the proposed technique selects important feature terms relying on the parts of speech (POS) in the textual information, and then uses sentiment analysis to estimate users' control variables for opinions using lexicon-based scoring analysis.

For improving classification-based false news identification, a data imputation preparation approach is presented. This approach is based on the utilisation of data imputation techniques to handle missing values in a dataset.

Subsequently, the term frequency and inverse document frequency (TF-IDF) were used for the extraction of useful features from the datasets to help the detection accuracy.

Finally, the fake news was detected using multiple classification models. Initially, for the multiclass prediction and robustness of predicting the class of text, the Naïve Bayes model had been used. Secondly, the passive-aggressive classifier trains the model incrementally and eventually, the deep neural network was used to increase the efficiency to detect fake news.

The outcomes of this study revealed that the suggested method's overall evaluation achieved a 99.8% accuracy rate for detecting false news. The findings are compared with and without multiple imputation execution in the creation of multiple classifier calculations for test set errors. Using the developed technique, this study produced a higher prediction rate while

evaluating various statements from the dataset, such as barely true, half true, true, largely true, and untrue. Finally, the developed strategy's performance is compared to that of current methods, in which the proposed method proved to be more efficient. The proposed classification models paired with the suggested missing data variable models and feature extraction strategy outperforms baselines, according to experimental results. Certainly in future work, it would be fascinating to test this proposed strategy for false news identification on other data sets.

Appendix

```
clc;
clear;
close all;
warning off;
%% read data
global compoundScores sFeat yvalid xvalid
T=readtable('2020-04-21 Coronavirus Tweets.csv','TextType','string');
T=T(1:1000,1:22);
dat=(T(:,end));
Var=dat.Properties.VariableNames;
emb=readWordEmbedding('glove.6B.100d.txt');
[compoundScores,documents]=preprocessing(dat,Var,emb,T);
%% n gram model
bag = bagOfNgrams(documents);
vec2=tfidf(bag);
full_dat =full(vec2(:,:));
%% feature selection using hybrid TSA-HHO-SA
SearchAgents=5;
Max_iterations=10;
lowerbound   = 0;
upperbound   = 1;
dimension = size(full_dat,2);
fitness=@jFitnessFunction;
ho = 0.3;
% Hold-out method
HO = cvpartition(compoundScores,'HoldOut',ho);
[Best_score,Sf,concurve]=tsa_hho_sa(SearchAgents,Max_iterations,lowerbound,upperbound
,dimension,fitness,full_dat,compoundScores,HO);
figure;
plot(concurve,'-b','linewidth',2);
xlabel('Iterations');ylabel('Objective value');
sFeat = full_dat(:,Sf);
sFeat=sFeat(:,1:50);
xtrain = sFeat(HO.training == 1,:);
ytrain = compoundScores(HO.training == 1);
xvalid = sFeat(HO.test == 1,:);
yvalid = compoundScores(HO.test == 1);
```

```
%% hybrid BOA artificial neural network
N_iter=50;
Lb=-1;
Ub=1;
n=5;
fobj=@fitnANN;
accuracy=ANN_BOA(xtrain,ytrain,N_iter,Lb,Ub,n,fobj);
function [X]=initialization(N,dim,up,down)

if size(up,1)==1
   X=rand(N,dim).*(up-down)+down;
end
if size(up,1)>1
   for i=1:dim
      high=up(i);low=down(i);
      X(:,i)=rand(1,N).*(high-low)+low;
   end
end
end
global net yvalid xvalid compoundScores sFeat
net = feedforwardnet(10);
net.trainParam.epochs=100;
net = train(net, sFeat', compoundScores');
mdl=fitctree(xvalid,yvalid);
y = net(xtrain');
acc=length(find(round(y')==ytrain))/length(ytrain);
weights = getwb(net);
dim=size(weights,1);
[fmin,best_pos,Convergence_curve]=BOA(n,N_iter,Lb,Ub,dim,fobj);
net = setwb(net,best_pos);
y = net(xvalid');
y=round(y);
y=predict(mdl,xvalid);
accuracy=length(find(yvalid'==y))/length(y);
d1=find(y==0);
d2=find(y==1);
d3=find(y==2);
```

Snapshot of the Run Image

Declarations

Conflict of interest   The authors declare that they have no conflict of interest.

References

1. Ahmad I, Yousaf M, Yousaf S, Ahmad M (2020) Fake news detection using machine learning ensemble methods. Complexity 2020:1–11
2. Akinyemi B (2020) Department of computer science and engineering, Obafemi Awolowo University, Ile-Ife, Nigeria, Adewusi O, Oyebade a. an improved classification model for fake news detection in social media. Int J Inf Technol Comput Sci 12(1):34–43. https://doi.org/10.5815/ijitcs.2020.01.05
3. Alonso MA, Vilares D, Gómez-Rodríguez C, Vilares J (2021) Sentiment analysis for fake news detection. Electronics (Basel) 10(11):1348. https://doi.org/10.3390/electronics10111348
4. Amer AYA, Siddiqui T (2020) Detection of Covid-19 fake news text data using random Forest and decision tree classifiers. Int J Comput Sci Inform Secur (IJCSIS) 18(12):88–100
5. Aslam N, Ullah Khan I, Alotaibi FS, Aldaej LA, Aldubaikil AK (2021) Fake detect: a deep learning ensemble model for fake news detection. Complexity 2021:1–8. https://doi.org/10.1155/2021/5557784
6. Choudhary A, Arora A (2021) Linguistic feature based learning model for fake news detection and classification. Expert Syst Appl 169(114171):114171. https://doi.org/10.1016/j.eswa.2020.114171
7. Dang NC, Moreno-García MN, De la Prieta F (2020) Sentiment analysis based on deep learning: a comparative study. Electronics (Basel) 9(3):483. https://doi.org/10.3390/electronics9030483
8. Duan X, Naghizade E, Spina D, Zhang X (2020) RMIT at PAN-CLEF 2020: Profiling Fake News Spreaders on Twitter In: CLEF (Working Notes)
9. Enders CK, Baraldi AN (2018) Missing data handling methods. In: The Wiley Handbook of Psychometric Testing. Chichester, UK: John Wiley & Sons, Ltd; p. 139–85

10. Hannah Nithya S, Sahayadhas A (2022) Automated fake news detection by LSTM enabled with optimal feature selection. J Inf Knowl Manag 21(03). https://doi.org/10.1142/s0219649222500368
11. Harb JG, Ebeling R, Becker K. (2020) A framework to analyse the emotional reactions to mass violent events on Twitter and influential factors. Inform Process Manag 57(6)
12. Harb JGD, Ebeling R, Becker K (2020) A framework to analyze the emotional reactions to mass violent events on twitter and influential factors. Inf process Manag 57(6):102372. https://doi.org/10.1016/j.ipm.2020.102372
13. Harel O, Mitchell EM, Perkins NJ, Cole SR, Tchetgen Tchetgen EJ, Sun B, Schisterman EF (2018) Multiple imputations for incomplete data in epidemiologic studies. Am J Epidemiol 187(3):576–584
14. Javed Awan M, Shehzad F, Muhammad H, Ashraf M (2020) Fake news classification bimodal using convolutional neural network and long short-term memory. Int J Emerg Technol 11(5):197–204
15. Jiang T, Li JP, Haq AU, Saboor A, Ali A (2021) A novel stacking approach for accurate detection of fake news. IEEE Access 9:22626–22639. https://doi.org/10.1109/access.2021.3056079
16. Kaliyar RK, Goswami A, Narang P, Sinha S (2020) FNDNet-a deep convolutional neural network for fake news detection. Cogn Syst Res 61:32–44
17. Kaliyar RK, Goswami A, Narang P (2021) EchoFakeD: improving fake news detection in social media with an efficient deep neural network. Neural Comput Appl 33(14):8597–8613. https://doi.org/10.1007/s00521-020-05611-1
18. Kumar S, Asthana R, Upadhyay S, Upreti N, Akbar M (2020) Fake news detection using deep learning models: a novel approach. Trans Emerg Telecommun Technol 31(2) . https://doi.org/10.1002/ett.3767
19. Lang KM, Little TD (2018) Principled missing data treatments. Prev Sci 19(3):284–294. https://doi.org/10.1007/s11121-016-0644-5
20. Maniruzzaman M, Rahman MJ, Al-MehediHasan M, Suri HS, Abedin MM, El-Baz A et al (2018) Accurate diabetes risk stratification using machine learning: role of missing value and outliers. J Med Syst 42(5):92. https://doi.org/10.1007/s10916-018-0940-7
21. Nasir JA, Khan OS, Varlamis I (2021) Fake news detection: a hybrid CNN-RNN based deep learning approach. Int J Inform Manag Data Insights 1(1):100007. https://doi.org/10.1016/j.jjimei.2020.100007
22. Richardson HA, Simmering MJ (2020) Missing Data in Research. In: Missing data in research. Oxford Research Encyclopedia of Business and Management. Oxford University Press, In
23. Sahoo SR, Gupta BB (2021) Multiple features based approach for automatic fake news detection on social networks using deep learning. Appl Soft Comput 100(106983):106983. https://doi.org/10.1016/j.asoc.2020.106983
24. Shu K, Mahudeswaran D, Wang S, Lee D, Liu H (2020) FakeNewsNet: a data repository with news content, social context, and spatiotemporal information for studying fake news on social media. Big data 8(3):171–188. https://doi.org/10.1089/big.2020.0062
25. Sperrin M, Martin GP, Sisk R, Peek N (2020) Missing data should be handled differently for prediction than for description or causal explanation. J Clin Epidemiol 125:183–187. https://doi.org/10.1016/j.jclinepi.2020.03.028
26. Xu K, Wang F, Wang H, Yang B (2020) Detecting fake news over online social media via domain reputations and content understanding. Tsinghua Sci Technol 25(1):20–27. https://doi.org/10.26599/tst.2018.9010139
27. Yuan C, Ma Q, Zhou W, Han J, Hu S (2020) Early detection of fake news by utilizing the credibility of news, publishers, and users based on weakly-supervised learning
28. Zhou X, Jain A, Phoha VV, Zafarani R (2020) Fake news early detection: a theory-driven model. Digital Threats: Research and Practice 1(2):1–25