

TF-IDF, chi-square, and ANN machine learning techniques for fake news anomaly detection in online social networks

Venkata Ashok Kumar Boyina¹, N. Sabiyath Fatima², Omar Isam Al Mrayat³, Pallavi S. Bangare⁴, Kasula Venkata Durga Kiran⁵, D. Vetrithangam⁶, Sunil L. Bangare⁴, Mohan Awasthy⁷, Bhaskar Roy⁸, Batyrkhan Omarov⁹, Guna Sekhar Sajja¹⁰

¹Independent Researcher, Cumming, United States

²Department of Computer Science and Engineering, B.S.Abdur Rahman Crescent Institute of Science and Technology, Chennai, India

³Department of Software Engineering, Amman Arab University, Amman, Jordan

⁴Department of Information Technology, Sinhgad Academy of Engineering, Pune, India

⁵Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India

⁶Department of Computer Science and Engineering, University Institute of Engineering, Chandigarh University, Punjab, India

⁷Department of Engineering and Technology, Bharati Vidyapeeth (Deemed to be University), Navi Mumbai, India

⁸Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning), Asansol Engineering College, Asansol, India

⁹Al-Farabi Kazakh National University, Almaty, Kazakhstan

¹⁰Department of Information Technology, University of the Cumberlands, Williamsburg, United States

Article Info

Article history:

Received Apr 18, 2025

Revised Aug 12, 2026

Accepted Sep 5, 2026

Keywords:

Accuracy

Anomaly detection

Artificial neural network

Chi-square test

Fake news

Online social networks

Term frequency-inverse

document frequency

ABSTRACT

Platforms such as Facebook, WhatsApp, Twitter, and Telegram have a substantial influence on the dissemination of information in modern society. Numerous individuals rely on them without verifying the veracity of their information or the sources of their knowledge. False information is referred to as "Fake News" and is disseminated through both traditional and non-traditional media channels including social media. Contemporary media consumers uncritically embrace the news they read online, rather than taking the time to critically assess it. This method has led to a heightened trust in internet content, which has facilitated the spread of misinformation. This article describes a machine learning (ML)-based approach to classifying and detecting fake news anomalies in online social media data. This framework accepts social media data as input. Features are chosen using the term frequency identification term frequency-inverse document frequency (TF-IDF) and chi-square tests. The categorization model is then built utilizing approaches such as artificial neural network (ANN), support vector machine (SVM), and random forest (RF). The classification model is trained and tested on a preprocessed data set. Finally, fake news is discovered, and the suggested framework is evaluated using measures such as accuracy, precision, recall, and F1-score. ANN is performing better for fake news classification and detection.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Guna Sekhar Sajja

Department of Information Technology, University of the Cumberlands

Williamsburg, KY 40769, United States

Email: guna.sajja@gmail.com

1. INTRODUCTION

Facebook, WhatsApp, Twitter, and Telegram are platforms that influence the dissemination of information in modern society greatly. They make lots of people trust them without even having to check the validity or the sources of their knowledge. The fake information spreads through both the traditional and non-

traditional media, among them social media and is referred to as the fake news [1]. The consumers worldwide are social media fans because of its availability, low cost, and ease in the spread of information. Unfortunately, it resulted in spreading misinformation [2]. It is a commonly used term to define articles that are visibly and deliberately false as fake news [3]. In addition, it also includes those that appear to be a news but include factual errors that are intended to mislead readers into thinking that they are reading authentic information [4], [5]. Scholars are keen to establish the cases of unethical reporting even in the post-internet age. The spread of fake news is now seen as a serious threat to democracy in our country and the integrity of the news media. It has a tremendous impact on the three spheres: social, political, and economic. There is a plethora of fake and false news. Information plays a significant role in the way we form our worldview and how we make evidence-based judgements. The acquired facts are applied in developing the views towards groups or events. Nonetheless, it is not easy to make informed decisions when the evidence is biased, inaccurate, or otherwise mistaken. Studies have shown that about 93% of the American population rely on one or another types of online resources or technology to meet their informational needs. One of the main impacts of massive use of social media is the rapid dissemination of misinformation and lies.

Thereby, filtering false information has become a major issue. The automated fake news recognition is becoming essential in order to reduce negative effects of fake news. It is not easy to analyze news material and verify its credibility, which makes it a complex process. It is a daunting task to develop a solid and stable method to detect misinformation. Presently, social media has become part of the daily life of people. It is impossible to assume that all the news and information that is spread by it is accurate. It is common that news items and other internet information contains no references [6]. False news can also be issued as news online. When the information is incorrect or untrue, it results in fake news. This misinformation is in the form of unsubstantiated claims that are invented. The stories are presented in a way that gradually forms the minds of the reader and inculcates the misguided ideas. The similarity between authentic news and misinformation has been challenging to distinguish in the modern society as they are similar. The use of social media and its reliance has increased exponentially over the past few years, and the speed of communicating news via such media is unbelievable [7].

Contemporary media consumers blindly believe what they read on the internet rather than take time to critically evaluate the news. The given approach has led to the growing faith in the contents of the internet, which has added to the spread of misinformation. The catchy headlines, popular language, and other forms of misinformation quickly find their way to the media consumers because of the crucial role played by the social media in their spread. According to the researchers, the current media ecosystem is disorganized because people are not able to properly detect the misleading captions that they come across online [8].

Cases of fake news that were popular on the internet are few. There is a multiplicity of events that took place online and social media as a result of spreading misinformation on a virus called Zika that spread within the United States. The existing emphasis on social media throws the light on the health concerns and the rising number of sicknesses spreading per day. According to the authors of the study, consumers are more interested in misinformation that is being shared through social media compared to proper information provided by health groups. Algorithms of artificial intelligence can create fake profile images with ease, and posts and comments can trap users of social networks.

The identification of sources of disinformation is becoming difficult due to the spread of misinformation rapidly and the increasing use of social media. The utilization of deep learning and machine learning (ML) techniques will allow concluding about the authenticity of the article. To draw the line between fake news and real news, a person can use a number of textual traits. Despite the extensive research on ML to identify fake news, it is possible to determine the accuracy of fake news content accurately based on the key features of texts and employing the appropriate methodology [9]. ML approaches can prove useful in addressing the complexity by identifying trends that can distinguish between genuine news and fake news. Discovering fake news is a good fit with these algorithms because they are useful in processing serial data and context of knowledge. They also record contextual relationships and interconnections over time as well as highlighting important phrases in the text. ML methods make detection models more effective by extracting salient features due to understanding the text. Information on fake news requires the management of a significant amount of data, such as user data and information on a social network [10]. These algorithms are constantly learning on new data to fit the ever-evolving world of misinformation. In turn, neural network practices have demonstrated an efficient approach in the detection of fake news by optimizing the whole process of detection and efficiently managing various information.

This paper presents a ML-based approach to categorizing and detecting anomalies in social media fake news in online data. The information around the social media sites can be put into this framework. The data collected undergoes preprocessing in order to enhance the accuracy of classification and the removal of noise. Features are selected by the term frequency-inverse document frequency (TF-IDF) and chi-square tests. Then the model of categorization is developed based on such methods as the support vector machine

(SVM), artificial neural network (ANN), and random forest (RF). A training and evaluation of the classification model are conducted using a preprocessed data set. Finally, false news is detected and such metrics as accuracy, precision, recall, and F1-score are employed to evaluate the suggested method. The concept of ANNs is more efficient in detecting and classifying bogus news.

2. LITERATURE SURVEY

To address the limitations of ANN, a hybrid model that combines both ANN and autoregressive integrated moving average (ARIMA) is suggested and is said to have more comprehensive and accurate predictions as compared to the traditional hybrid ARIMA-ANN models. The research has also made use of neural networks, to forecast the underlying data with the help of the peculiarities of ARIMA models in the linear process of modelling on the preprocessed data. This allowed the researchers to identify and emphasize existing linear patterns in the data. Using the proposed methodology in the real-world data sets allowed the researchers to show that the traditional hybrid and single models could possess the improved forecasting accuracy [11].

They plan to create a prototype on the detection of false information with the help of polarity analysis [12]. The authors acquired the editorials of the credible sources of English-language news and commented on them. It was established after developing content-based features based on the use of n-grams and analysis using SVM, k-nearest neighbors (KNN), and the Naive Bayes that KNN classifiers performed better than SVM classifiers when used in polarity analysis to detect fake news. By using ML and natural language processing (NLP), authors created the methodology of aggregating news stories across multiple social media platforms, such as Facebook, Twitter, and WhatsApp. After this, they used SVM to determine reliability of the articles. The output of the model is assessed using modern models. Results of the model have an accurate characterization of 93.6% [13].

Adaptive boosts are necessary in the process of classifications made by gradient descents; therefore, the tree-based ensembles (gradient boosting) were used to identify fake news on the basis of content and context-based information. In order to maximize standard goal functions, the main elements and features of the methods were carefully chosen. The multiclass datasets of false news challenges (FNC) and some ML approaches were utilizing in this study. Their ensemble structures were far beyond most preceding norm, as the results of the study conducted show. The accuracy of the project in classifying fake news into four categories through the gradient boosting technique which is an ensemble ML framework has been realized to be 86% [14].

In order to distinguish between real and fake news, as well as to overcome the unbalanced data problem existing in the dataset, researchers applied an ensemble learning method on Indonesian fake news. RF ensembles performed better than multinomial Naive Bayes (MNB) and SVM (non-ensembles) with a F1-score of 0.98 and 0.43 and 0.74, respectively as per the study findings. This research revealed that the strategy yielded better conclusions than prior studies which used the same data [15].

Semantic features as well as other ML methods of identifying misinformation in digital media. In this study, there were five linguistic criteria to test the NB and RF classifiers. After evaluation based on both real and fake data of kaggle.com, the top model with bigram features and the RF classifier achieved an accuracy of 95.66%. Compared to unigrams, trigrams, and quadgrams, bigrams (word pairs) are a better indicator of news validity [16]. Hussna *et al.* [17] combined text data mining and the use of supervised ML algorithms to identify coronavirus disease 2019 (COVID-19) misinformation on social media. The text mining analysis and classification operations have been conducted on the text data. The experiment applies two algorithms, RF, and decision tree (DT), in order to identify incorrect information about COVID-19. According to the results, the measures of accuracy, precision, recall, F1-score, as well as support values have considerably improved when comparing the two methods. The results of the testing findings point to the fact that the proposed approach to detecting COVID-19 misinformation in social media had a trained and tested accuracy of 94.49% in RF. The level of performance reached by the DT was 92.07%. Following the negative publicity and false information, scholars revealed various facts about the disinformation of the online community. The findings clearly show the importance of accuracy, recall as well as the RF. The results of DT classifier are similar to the RF classifier.

Two different ML algorithms, RF, and DT (J48) were used in work [18] to detect fraudulent news. Out of the 20,761 samples, 4345 samples were to be tested. The first step was to remove the stop words in the data and then the removal of redundant special characters, numerical values, English alphabet and whitespace. The researchers prescribed two sets of features by extractions of TF-IDF features. According to the research findings, DT has scored high at 88.11 and RF at 84.97.

To detect fake news, the authors of proposed evaluate the completeness and validity of a report [19]. This algorithm built ensemble networks to process the news strip representations, producers, and titles in parallel because they offered trustworthiness scores of textual input components. Additional ML tools are used to achieve better accuracy such as SVM, KNN, and Naive Bayes. Measures that are utilized to

determine the effectiveness and performance of classifiers are accuracy, recall, and F1-score. The use of numerous algorithms emphasizes the performance effectiveness of the dataset.

Given the topicality of this issue, present an evaluation of the experimental study aimed at identifying fake news on the social media in the Bangali language. The features that were used to identify Bangla fake news were count vectorizer and TF-IDF vectorizer, and two supervised ML algorithms were used, namely MNB and SVM classifier. The article can be characterized by the framework that identifies the fake news [20]. This paper indicates that SVM using linear kernel is better than MNB and has a rate of 96.64 accuracy. Mahabub [21] suggested using ensemble voting-based intelligent detections in order to classify news as authentic or fraudulent. There were eleven different ML detection methods that were used in this research, namely: Naive Bayes, SVM, KNN, RF, logistic regression, gradient boosting, ANN, and AdaBoost. The ensemble voting was then used to select the three most common ML algorithms following the cross-validation. The experiment findings they had shown that their framework had the possibility of achieving an accuracy level of 94.5% or higher. Measures of receiver operating characteristic (ROC) improved during this experiment with memory, accuracy, and F1-scores. The technique can be applied to the classification of fraudulent personalities and messages, as well as the identification of news highlights worth referring to.

Weigh up detecting fake news within tweets by using context-based detection algorithms [22]. Credibility attributes of the user combined with the overall knowledge of social interactions, and the news content allow the automated detection of fake news on Twitter. The level of reliability is initially evaluated through examining user profile. The number of comments, likes, and retweets a post receives is an indicator of the view of the user and how much he or she is involved. The news has been analyzed using SVM model using radial basis function (RBF) kernel. An example of Twitter data was used to demonstrate the effectiveness of the strategy. Misleading and classified false news that misleads the viewers to believe false stories [23]. The study employed the use of language analysis to establish the content features, out of which the emotional, grammatical, syntactic, and readability news article dimensions were obtained. The models based on the language have a long time to process and the hand-crafted features are limited by the curse of dimensionality. Accordingly, a sequential learning algorithm based on neural networks was used to enhance the accuracy of detecting fake news. The model of the study based on the integrated linguistic features achieved an average detection and classification of false news of 86% and therefore confirmed the effectiveness of the retrieved linguistic features. Compared to the ML-based models and word embedding-based false news detection techniques, their sequential neural model showed better performance. Their comparison showed that feature based sequential models achieved the same assessment results at lower run times.

3. METHOD

This section presents a ML based methodology to classify and detect fake news anomalies in online social media data. This methodology is presented in Figure 1.

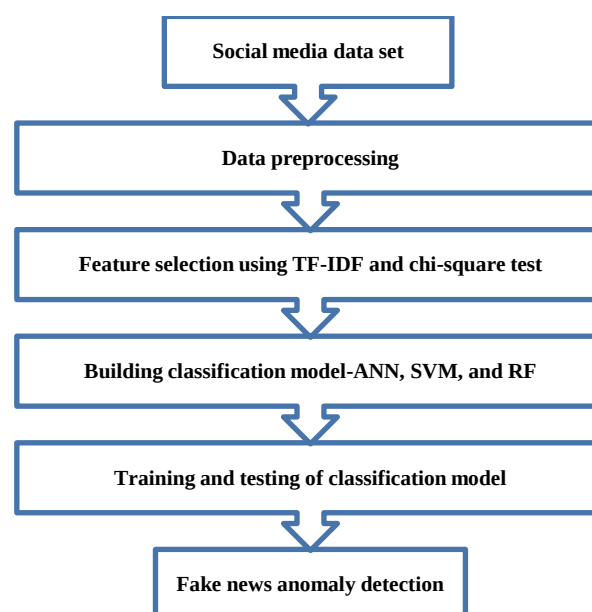


Figure 1. TF-IDF-chi-square and ANN ML techniques for fake news anomaly detection

In this framework, social media data is used as input. First of all, data set is preprocessed to eliminate any noise and improve classification accuracy. Features are selected using TF-IDF and chi-square test. Then classification model is developed using the ANN, SVM, and RF techniques. Training and testing of classification model is conducted using the preprocessed data set. Finally fake news are detected and proposed framework is analyzed on parameters like- accuracy, precision, recall, and F1-score. Table 1 presents the ablation study and architecture choice related details.

Table 1. Summary of architectural choices and ablation study validation

Architectural choice	Description	Validation (ablation study)	Impact on performance
Model type	Used a combination of ANN, SVM, and RF for classification.	Evaluated each model (ANN, SVM, and RF) separately, assessing their accuracy, precision, recall, and F1-score.	ANN performed the best with 98.45% F1-score, outperforming SVM (93.35%) and RF (89.35%).
Feature selection	TF-IDF and chi-square tests were used for feature extraction.	Tested performance with and without feature selection methods.	Feature selection improved model accuracy (ANN: 98.3%).
Preprocessing	Removed noise from data through tokenization, stopword removal, and lemmatization.	Compared models trained with raw data versus preprocessed data.	Preprocessing improved model precision (ANN: 98.3%) and recall (96.7%).
Training dataset	Used mediaeval (DB3) dataset containing 17,000 tweets (6,000 real, 9,000 fake).	Used smaller datasets to evaluate the impact of dataset size on model performance.	Larger dataset improved accuracy (ANN: 98.3%) and recall (96.7%).
Evaluation metrics	Performance evaluated using accuracy, precision, recall, and F1-score.	Evaluated models on multiple metrics to validate robustness.	ANN achieved the highest overall performance across all metrics (F1-score: 98.45%).

3.1. Data preprocessing

The initial important step in ML models of raw text data cleaning and preparation is preprocessing. It starts by changing the text to lower case so that such words as Fake and fake are similar and do not have any case sensitivity problems. To minimize noise, punctuations that do not usually contain any meaningful information used to detect fake news are removed. There is also the removal of stopwords including the, is, and. The stopwords are common in the text but do not add much value in terms of classification. The text is then tokenized to individual tokens (words) and it is processed more easily by models. Lastly, lemmatization eliminates words to their basic forms (e.g., running to run) to cluster related variations which enhances the efficiency of the model. Their combination of these preprocessing steps improves the data streams to concentrate on the most significant terms to avoid inaccurate analysis and classification.

3.2. Feature selection using term frequency-inverse document frequency and chi-square test

Features are chosen with the TF-IDF and chi-square methods. The text data collection is initially preprocessed by converting the text to lowercase. Punctuation is eliminated from the supplied dataset. Punctuation marks often lack semantic significance and may introduce noise during feature extraction. Subsequently, stopwords are eliminated, as they are ubiquitous yet uninformative for categorization purposes. The tokenization procedure divides the phrase into distinct words or tokens. Lemmatization transforms words into their basic or dictionary form ("better" → "good", "running" → "run"). The chi-square test assesses the independence of a term's occurrence from the class label. The more its dependence, the more revealing the trait becomes. Utilize it to identify the most pertinent terms for categorization [24], [25].

3.2.1. Term frequency-inverse document frequency

TF-IDF is a statistical value that can be used to determine the significance of a word to a document within a collection or corpus:

- Term frequency (TF) is defined as the frequency of the appearance of a word in a document.
- Inverse document frequency (IDF) is used to recompense the fact that some words are more common in all documents (such as common words), and then this decreases their significance.
- TF-IDF is based on the following concept: a word is considered frequent in a single document and is considered to be rare in all documents, so it is likely to have a significant amount of information. Although words such as breaking, exclusive or scam could be found more in fake news than regular news and thus could be more substantial when it comes to fake news detection.

3.2.2. Chi-square test for feature selection

Chi-square is a statistical test that is applied to understand whether a feature (term) is independent of the class label (fake or real news in this case):

- Chi-square test is used to test the independence of the appearance of a word (feature) and the target class (fake or real). Chi-square values are high, which means that a word has a strong relationship with the class label and can be considered relevant.
- Words such as scam, hoax, and unverified can be used more often on fake news stories. Such words can be detected by the chi-square test since these words are very strong indicators to the label of the fake class.
- Chi-square is to rank the features in terms of their relevance to classification. The characteristics with a more significant chi-square are said to be more telling and are relevant in terms of categorization.

3.2.3. Integration of term frequency–inverse document frequency and chi-square

TF-IDF which weighs terms by their frequency and inverse frequency across documents plus chi-square which tests the relationship between terms and the labels of the class form a very strong feature extraction and selection pipeline:

- TF-IDF is useful in finding significant terms through the frequency of occurrence of a term in a document or in the entire corpus.
- Chi-square is used to refine the features by testing the association between the terms and the class labels and allows the model to concentrate on the words which have the most predictive value.

3.3. Model building

An ANN has several processing components that utilize specified activation functions to take inputs and produce erroneous outputs. An ANN consists of input, hidden, and output layers. The uppermost layer collects input information, undergoes analysis, and generates 'n' outputs. All processes are conducted according to weights. Weights serve as data aids in addressing problems in neural networks [26]. The concealed information is recovered by the hidden layer as it progresses from the input layer to the output layer when substantial information is obtained. ANN are employed for the detection of counterfeit attributes. The information security online threats (ISOT) dataset is utilized for training in ANN, while the testing phase involves classifying state characteristics and identifying false information. Figure 2 illustrates the architecture of the ANN:

- Input layer: the input layer transports the data presented to the network. Initially, this information is little underdeveloped.
- Hidden layer: the hidden layer's primary function is to transform the raw data from the input layer into something that the output layer can use. There might be one or more hidden layers in an ANN.
- Output layer: to achieve the desired results, the output layer processes the data received from the hidden layer. Adjustable values along the connections between neurones in ANNs may be altered via a learning process that improves the model as it learns from observed data. A suitable cost function must be selected in addition to the learning process. ISOT used word embedding to detect false news by generating a set of prepared misleading phrases and assigning weights to gain fake characteristics. A linear function equation and a transfer function calculate the features from the input layer in the hidden layer to produce a hidden node's output feature.

$$N_i = \sum_{p=1}^P w_{i,p} I_p + B_i$$

where N_i is the value of the i th feature; $w_{i,p}$ and B_i are the weight of the p th input to the i th hidden node. Here is how the transfer function is defined,

$$y_i = f(N_i) = \frac{1}{1 + \exp(-N_i)}$$

where the sigmoid function used in this research serves as the transfer function and y_i is the output signal of the i th hidden node. Using the model's performance is evaluated MSE throughout the learning phase as,

$$MSE = \frac{1}{NN_{out}} \sum_{n=1}^N \sum_{o=1}^{N_{out}} (e_{n,o})^2$$

where N and N_{out} are the number of outputs and the number of occurrences, respectively; $e_{n,o}$ is the inaccuracy in training at the other output with n th instance; y the real result, and $\bar{y}$ is the result that ANN anticipated.

RF algorithms are a set of categorization approaches that aggregate decisions from many DT. Binary splits are frequently used to recursively partition the tree into homogeneous or very homogeneous terminal nodes when expanding the RF tree base learner. RF is frequently made up of hundreds to thousands of trees,

each built from a first data bootstrap sample. When a tree is created, a second layer of randomization is applied at the node level in addition to the randomization generated by creating the tree with a bootstrap sample of the starting data. RF selects a randomized sample of factors at each tree node, and only those variables are utilized as candidates to determine the node's optimal division. This two-step randomization minimizes the unpredictability of the forest ensembles, or bagging phenomenon, by de-correlating the trees.

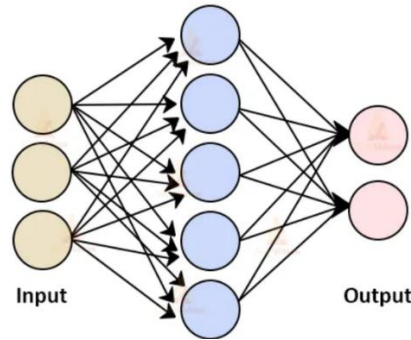


Figure 2. ANN architecture

The SVM is a data mining approach that uses supervised ML to analyze data and recognize patterns. Based on structural risk minimization analysis and statistical learning theory, the SVM divides the training set into two classes of positive and negative samples using a maximum margin hyperplane. The amount of characteristics that should be taught cannot be determined in advance. It is an effective ML method that seeks to identify a linear separation hyperplane with the highest margin for distinguishing data in a higher-dimensional environment. The conventional SVM method requires a significant training time [27].

In our paper, we adopted a mixture of ANN, SVM, and RF to classify fake news. TF-IDF and chi-square tests were used in the selection of the features, and the data was pre-processed by removing noise.

4. RESULTS ANALYSIS

4.1. Result and discussion

This section includes the results of the proposed fake news detector on various datasets, such as mediaeval (DB3) dataset, the LIAR and FakeNewsNet datasets. Another issue that we address is the effectiveness of different preprocessing procedures, such as tokenization, elimination of stopwords, and stemming/lemmatization, on model performance [28].

4.1.1. Performance across multiple datasets

We test the performance of our model on three different datasets to test how well it can generalize. The datasets used are:

- Mediaeval (DB3): this is a collection of 17,000 tweets (6,000 genuine and 9,000 fake) utilized in the training and testing.
- LIAR: this is a dataset containing 12,800 labeled short statements of real-world political conversations, where the labels are true, false, and mixed statements.
- FakeNewsNet: 20,000+ articles were gathered by online sources; the dataset was divided into fake and real articles.

Accuracy: accuracy is a measure that define the overall progress of the model. It measures the frequency at which the algorithm assigns the correct classification to a data point. Basically, it involves the proportion of data points that are correctly classified divided by the total number of classified data points, as given below.

$$\text{Accuracy} = \frac{(\text{TruePositive (TP)} + \text{TrueNegative (TN)})}{(\text{TruePositive (TP)} + \text{TrueNegative (TN)} + \text{FalsePositive (FP)})}$$

Precision: precision displays the model accuracy in terms of the proportion of true positives among all positive outcomes, including those that were incorrectly recognized.

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

where, TP represents true positive and FP represents false positive.

Recall: recall discusses the proportion of samples that are true positives out of all samples that should have been classified as positive.

$$Recall = TP / (TP + FN)$$

F-Measure: the F-measure combines both precision and recall presenting a weighted average known as the harmonic mean. The F-score can reach a maximum of 1.0 and a minimum of 0.

$$F1 = 2 \times ((precision \times recall) / (precision + recall))$$

Figure 3 and Table 2 shows that the ANN outperforms all measures, with the highest precision (98.3%) and recall (96.7%), yielding a good F1-score (98.45%). This model is extremely accurate and produces very few mistakes, demonstrating great sensitivity and specificity. SVM provides excellent performance across the board. In all measures, it outperforms RF. Solid F1-score (93.35%), indicating a favourable trade-off between recall and precision.

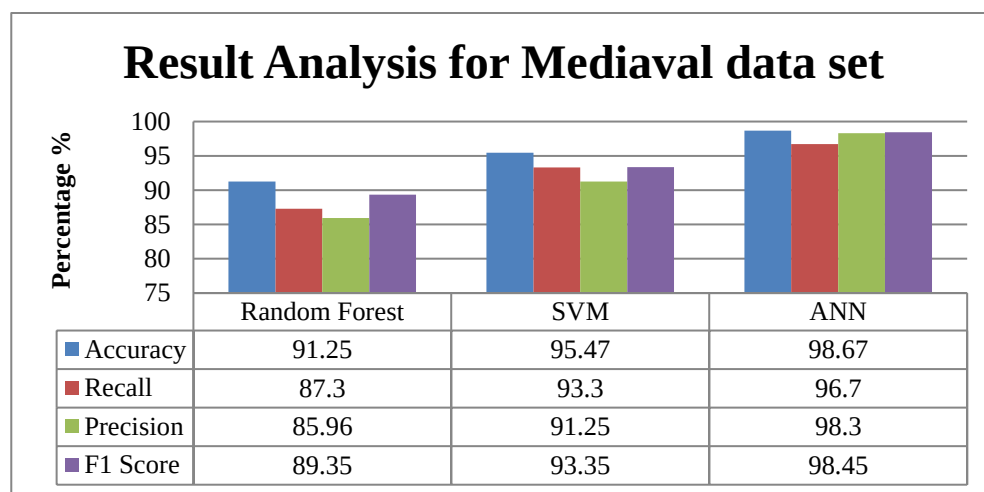


Figure 3. Accuracy, precision, recall, and F1-score of classifiers for fake news anomaly detection

Table 2. Performance comparison across datasets and impact of preprocessing steps

Dataset	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Mediaval (DB3)	ANN	98.67	98.3	96.7	98.45
	SVM	95.47	91.25	93.3	93.35
	RF	91.25	85.96	87.3	89.35
LIAR	ANN	96.3	94.1	92.5	93.3
	SVM	94.8	90.4	91.1	90.75
	RF	89.9	82.7	84.2	83.4
FakeNewsNet	ANN	97.4	95.6	94.1	94.8
	SVM	93.6	90.8	91.3	91.05
	RF	89.5	84.3	85.2	84.75

The preprocessing stages are important in enhancing the performance of the model. The datasets were subjected to the following steps:

- Tokenization: the text data was broken into single words or tokens to make the analysis easy.
- Stopword removal: insignificant words which occur regularly like the, is, in, etc. that do not help in the classification process were eliminated in the dataset.
- Stemming/lemmatization: words were made in their root form in order to standardize the data. As an example, running was changed to run and better too good.

These preprocessing steps were compared as to the effect they had on model training and testing. The Table 3 summarizes the findings of the various preprocessing settings.

As reported in Table 3, the use of all the preprocessing steps (tokenization, stopwords, and stemming/lemmatization) gives the best performance in all metrics. In particular, the accuracy improves by a wide margin in the preprocessing scenario, which proves the relevance of these procedures to improve the performance of the model.

Table 3. Impact of preprocessing steps on model performance

Preprocessing steps	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
None (raw data)	85.2	80.5	82.1	81.3
Tokenization only	90.1	85.3	86.5	85.8
Tokenization+stopword removal	93.2	88.2	90.1	89.1
Full preprocessing (tokenization+stopwordremoval+stemming/lemmatization)	98.7	98.3	96.7	98.45

4.2. Real-life scenario: social media moderation

The control of the content, such as the recognition of fake news, is always a problem in the case of social media, including Facebook, Twitter, or Instagram. The amount of data produced and the pace of information exchange with each other makes the world a place where false information is easily disseminated. The real-time identification and notification of such content before they become viral are a major issue in this case.

How the proposed model would adapt:

- Continuous learning and adaptation to new patterns: the latest model, which relied on TF-IDF, chi-square tests, and ANN, is useful to detect fake news trends in structured information. Nevertheless, concerning social media moderation, the content keeps changing. The model may be used with continuous learning in order to be adjusted. This will enable the model to keep on developing as new forms of fake news crop up so that it is never left behind in the current trends of deceit. The model may be provided with incremental training as an example; the model will be trained on every newly found piece of misinformation without going through retraining.
- Real-time processing: real-time analysis of content is needed on the social media platforms. In this case, the suggested ML model will be able to combine with the flows of social media data and implement the anomaly detection at the moment of the new data arrival. Through the implementation of a distributed solution (e.g., a cloud computing system), the detection of fake news can be conducted upon upload of content, which will allow responding much faster.
- Multi-modal data integration: the content of the social media is not only written; it can be in the form of images, videos and even memes and these are most frequently used to spread fake news. The model would be flexible to include multi-modal data processing. It might, as an example, analyze the visual contents of the posts with the help of computer vision models in addition to textual analysis. This would improve the capacity of the model to identify misinformation in more sophisticated forms of media.
- Contextual sensitivity: the model must be in a position to differentiate information that is contextually correct and misinformation. Sarcasm/satire- an example is that it may seem like misinformation, yet not always unhealthy. One way of solving this issue is by enhancing the contextual awareness of the model using NLP improvements, e.g., sentiment analysis or contextual word embeddings (e.g., BERT and GPT). With the use of such techniques, the model will be more accurate in detecting intent of content and cut off the false positives (i.e., identifying satirical content as fake news).
- User and source profiling: the other adaptation of the model is the use of user and source profiling. Users or groups of people in social media may be more inclined to posting fake news. The model can rank the flagging of the content in terms of origin and credibility of the source of the content by examining the historical behavior of the users and judging the reliability of the information sources of origin of such content.
- Scalability across platforms: since misinformation may be transmitted across various platforms, the model may be implemented in a cross-platform surveillance system. As an illustration, when fake news is identified on Twitter, the system is able to use the detected information to track such information in Facebook, WhatsApp, or any other platform. The scalability aspect will enable the model to keep on adapting and learning more social media platforms, enhancing its efficiency in identifying misinformation in the varied settings.

5. CONCLUSION

This paper presents a ML-based method for online social media data false news anomaly detection and classification. Social media data provides input for this approach. Chi-square testing and TF-IDF help to choose features. Then the classification model is developed using RF, SVM, and ANN. Preprocessed data sets are used both in training and testing the classification model. At last, bogus news is found, and the proposed structure is assessed in respect to accuracy, precision, recall, and F1-score. ANNs are doing better at detection and classification of bogus news. With the highest precision (98.3%) and recall (96.7%), ANN beats all metrics and produces a decent F1-score (98.45%). With considerable sensitivity and specificity, this model is quite accurate and generates rather few errors. All around performance is outstanding thanks to SVM.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Venkata Ashok Kumar Boyina	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
N. Sabiyath Fatima		✓				✓		✓	✓	✓	✓	✓		
Omar Isam Al Mrayat	✓	✓	✓	✓	✓	✓		✓		✓	✓		✓	
Pallavi S. Bangare	✓	✓		✓		✓		✓	✓	✓			✓	
Kasula Venkata Durga Kiran		✓				✓		✓	✓	✓	✓	✓		
D. Vetrithangam	✓	✓	✓	✓		✓	✓			✓	✓		✓	✓
Sunil L. Bangare	✓		✓	✓	✓	✓		✓	✓	✓			✓	
Mohan Awasthy	✓		✓	✓	✓	✓	✓		✓		✓		✓	
Bhaskar Roy	✓			✓		✓		✓	✓	✓			✓	
Batyrkhan Omarov		✓				✓		✓	✓	✓	✓	✓		
Guna Sekhar Sajja	✓		✓	✓	✓	✓		✓	✓	✓			✓	

- C : Conceptualization
M : Methodology
So : Software
Va : Validation
Fo : Formal analysis
- I : Investigation
R : Resources
D : Data Curation
O : Writing -Original Draft
E : Writing - Review &Editing
- Vi : Visualization
Su : Supervision
P : Project administration
Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

REFERENCES

[1] X. Li *et al.*, “A Survey of Multimodal Fake News Detection: A Cross-Modal Interaction Perspective,” *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 9, no. 4, pp. 2658-2675, Aug. 2025, doi: 10.1109/TETCI.2025.3543389.

[2] M. Azzeh, A. Qusef, and O. Alabboushi, “Arabic Fake News Detection in Social Media Context Using Word Embeddings and Pre-trained Transformers,” *Arabian Journal for Science and Engineering*, pp. 923–936, Apr. 2024, doi: 10.1007/s13369-024-08959-x.

[3] A. J. D. Forno, G. P. Richetti, and V. H. Knaesel, “Fake news detection algorithms – A systematic literature review,” *Data & Knowledge Engineering*, vol. 158, p. 102441, Jul. 2025, doi: 10.1016/j.datak.2025.102441.

[4] Y. Ding *et al.*, “EvolveDetector: Towards an evolving fake news detector for emerging events with continual knowledge accumulation and transfer,” *Information Processing & Management*, vol. 62, no. 1, p. 103878, Jan. 2025, doi: 10.1016/j.ipm.2024.103878.

[5] L. Zhang, X. Zhang, Z. Zhou, X. Zhang, P. S. Yu, and C. Li, “Knowledge-aware multimodal pre-training for fake news detection,” *Information Fusion*, pp. 102715–102715, Sep. 2024, doi: 10.1016/j.inffus.2024.102715.

[6] G. Pennycook and D. G. Rand, “The psychology of fake news,” *Trends in Cognitive Sciences*, vol. 25, no. 5, pp. 388–402, May 2021, doi: 10.1016/j.tics.2021.02.007.

[7] D. P. Calvillo, R. J. B. Garcia, K. Bertrand, and T. A. Mayers, “Personality factors and self-reported political news consumption predict susceptibility to political fake news,” *Personality and Individual Differences*, vol. 174, pp. 1–6, May 2021, doi: 10.1016/j.paid.2021.110666.

[8] C.-C. Wang, “Fake News and Related Concepts: Definitions and Recent Research Development,” *Contemporary Management Research*, vol. 16, no. 3, pp. 145–174, Sep. 2020, doi: 10.7903/cmr.20677.

[9] O. D. Apuke and B. Omar, “Fake News and COVID-19: Modelling the Predictors of Fake News Sharing among Social Media Users,” *Telematics and Informatics*, vol. 56, pp. 1–16, Jul. 2021, doi: 10.1016/j.tele.2020.101475.

[10] D. Carrion-Alvarez and P. X. Tijerina-Salina, “Fake news in COVID-19: A perspective,” *Health Promotion Perspectives*, vol. 10, no. 4, pp. 290–291, Nov. 2020, doi: 10.34172/hpp.2020.44.

- [11] M. Khashei and M. Bijari, "A novel hybridization of artificial neural networks and ARIMA models for time series forecasting," *Applied Soft Computing*, vol. 11, no. 2, pp. 2664–2675, Mar. 2011, doi: 10.1016/j.asoc.2010.10.015.
- [12] M. Jane, "Polarity Analysis of Editorial Articles towards Fake News Detection," in *Proceedings of the 2018 1st International Conference on Internet and e-Business*, Apr. 2018, pp. 108–112, doi: 10.1145/3230348.3230354.
- [13] A. Jain, A. Shakya, H. Khatter, and A. K. Gupta, "A smart System for Fake News Detection Using Machine Learning," in *2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*, Ghaziabad, India, 2019, pp. 1–4, doi: 10.1109/ICICT46931.2019.8977659.
- [14] R. K. Kaliyar, A. Goswami, and P. Narang, "Multiclass Fake News Detection using Ensemble Machine Learning," in *2019 IEEE 9th International Conference on Advanced Computing (IACC)*, Tiruchirappalli, India, 2019, pp. 103–107, doi: 10.1109/IACC48062.2019.8971579.
- [15] M. Fayaz, A. Khan, M. Bilal, and S. U. Khan, "Machine learning for fake news classification with optimal feature selection," *Soft Computing*, vol. 26, pp. 7763–7771, Jan. 2022, doi: 10.1007/s00500-022-06773-x.
- [16] P. Bharadwaj and Z. Shao, "Fake News Detection with Semantic Features and Text Mining," *International Journal on Natural Language Computing*, vol. 8, no. 3, pp. 17–22, Jun. 2019, doi: 10.5121/ijnlc.2019.8302.
- [17] A. U. Hussna, R. Alam, R. Islam, B. F. Alkhamees, M. M. Hassan, and Md Z. Uddin, "Dissecting the Infodemic: An In-Depth Analysis of COVID-19 Misinformation Detection on X (Formerly Twitter) Utilizing Machine Learning and Deep Learning Techniques," *Heliyon*, vol. 10, no. 18, pp. e37760–e37760, Sep. 2024, doi: 10.1016/j.heliyon.2024.e37760.
- [18] R. Jehad and S. A. Yousif, "Fake News Classification Using Random Forest and Decision Tree (J48)," *Al-Nahrain Journal of Science*, vol. 23, no. 4, pp. 49–55, Dec. 2020, doi: 10.22401/anjs.23.4.09.
- [19] A. Agarwal and A. Dixit, "Fake News Detection: An Ensemble Learning Approach," in *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2020, pp. 1178–1183, doi: 10.1109/ICICCS48265.2020.9121030.
- [20] M. G. Hussain, M. R. Hasan, M. Rahman, J. Protim, and S. Al Hasan, "Detection of Bangla Fake News using MNB and SVM Classifier," in *2020 International Conference on Computing, Electronics & Communications Engineering (iCCECE)*, Southend, UK, 2020, pp. 81–85, doi: 10.1109/iCCECE49321.2020.9231167.
- [21] A. Mahabub, "A robust technique of fake news detection using Ensemble Voting Classifier and comparison with other classifiers," *SN Applied Sciences*, vol. 2, no. 4, pp. 1–9, Mar. 2020, doi: 10.1007/s42452-020-2326-y.
- [22] V. C. Tran, V. D. Nguyen, and N. T. Nguyen, "Automatic Fake News Detection by Exploiting User's Assessments on Social Networks: A Case Study of Twitter," *Trends in Artificial Intelligence Theory and Applications. Artificial Intelligence Practices*, pp. 373–384, 2020, doi: 10.1007/978-3-030-55789-8_33.
- [23] A. Choudhary and A. Arora, "Linguistic Feature Based Learning Model for Fake News Detection and Classification," *Expert Systems with Applications*, vol. 169, p. 114171, Nov. 2020, doi: 10.1016/j.eswa.2020.114171.
- [24] D. E. Cahyani and I. Patasik, "Performance comparison of TF-IDF and Word2Vec models for emotion text classification," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 5, pp. 2780–2788, Oct. 2021, doi: 10.11591/eei.v10i5.3157.
- [25] Z. Zhan, "Comparative Analysis of TF-IDF and Word2Vec in Sentiment Analysis: A Case of Food Reviews," *ITM Web of Conferences*, vol. 70, pp. 1–7, 2025, doi: 10.1051/itmconf/20257002013.
- [26] S. Maham, A. Tariq, Muhammad, F. S. Alamri, A. Rehman, and T. Saba, "ANN: adversarial news net for robust fake news classification," *Scientific Reports*, vol. 14, no. 1, pp. 1–20, Apr. 2024, doi: 10.1038/s41598-024-56567-4.
- [27] P. Dedeepya, M. Yarrarapu, P. P. Kumar, S. K. Kaushik, P. N. Raghavendra, and P. Chandu, "Fake News Detection on Social Media Through a Hybrid SVM-KNN Approach Leveraging Social Capital Variables," in *2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, Salem, India, 2024, pp. 1168–1175, doi: 10.1109/ICAAIC60222.2024.10575681.
- [28] M. Lab, "LIAR-DATASET," Kaggle, [Online]. Available: <https://www.kaggle.com/datasets/doanquanvietnamca/liar-dataset>. (Accessed: Mar. 14, 2025).

BIOGRAPHIES OF AUTHORS



Venkata Ashok Kumar Boyina    is a senior researcher from Cumming, United States. His research areas are machine learning, deep learning, and generative AI. He can be contacted at email: Venkat65534@gmail.com.



N. Sabiyath Fatima    is associate professor in Department of Computer Science and Engineering, B.S.Abdur Rahman Crescent Institute of Science and Technology, Chennai. She is having vast experience in teaching, research and administration. She can be contacted at email: sabiyathfathima@crescent.education.



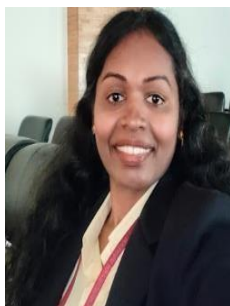
Omar Isam Al Mrayat    is a faculty in the Department of Software Engineering, Amman Arab University, Amman, Jordan. His research interests are machine learning and educational data mining. He can be contacted at email: o.mrayat@aau.edu.jo.



Pallavi S. Bangare    faculty in Department of IT, Sinhgad Academy of Engineering, Pune, India. She is having a vast experience in teaching, research and administration. Her research areas are machine learning, deep learning in educational and health care data analytics. She can be contacted at email: pallavi.bangare@gmail.com.



Kasula Venkata Durga Kiran    working as a professor and alternate HoD in the Department of Computer Science and Engineering at L University for 20 years teaching UG and PG engineering students. He completed his master of Technology in the Computer Science Engineering at KLCE Guntur District in 2007. He completed his Ph.D., in Computer Science and Engineering at Acharya Nagarjuna University, Andhra Pradesh, India. His main area of research is software engineering, network security, wireless sensor networks, data science, and machine learning. He has published 71 papers, all indexed in Scopus, with 14 Web of Science papers, and published 5 patents. Under his guidance, 8 Ph.D. scholars have been awarded Doctoral Ship till now and received 12 times the best teacher award for academics and also received the Drona award from IBM for doing a project. He can be contacted at email: kiran_cse@kluniversity.in.



D. Vetrihangam    is working as a professor in the Department of Computer Science and Engineering at Chandigarh University, Punjab, India. She has more than 17+ years of experience in teaching and research, and has published more than 90 research papers in various Scopus- and SCI-indexed journals and conferences and edited books, chapters. She is currently working in image processing, deep learning, machine learning, the internet of things, and remote sensing for various applications, including medical imaging, disease detection and prediction, cancer detection, soil moisture, and climate change. She has published books and book chapters, holds 12 published patents and 1 grant patent, and has more than 5 design granted patents. She also guides Ph.D. and M.E. research scholars working in machine learning, deep learning, explainable AI, and wireless sensor networks. She can be contacted at email: vetrigold@gmail.com.



Dr. Sunil L. Bangare    working as an associate professor in the Department of Information Technology at Sinhgad Academy of Engineering, Pune, India. He is in teaching profession for more than 17 years and have 1.6 years of industrial experience. He has published more than 100 papers in national and international journals and conference (Scopus and Web of Science indexing) and 8 book chapters. He can be contacted at email: sunil.bangare@gmail.com.



Mohan Awasthy    is principal and professor in Bharti Vidyapeeth, Navi Mumbai, India. He has keen interest in machine learning and deep learning based research areas. He can be contacted at email: awasthyamohan@gmail.com.



Dr. Bhaskar Roy, MIEI, LMISTE    is currently working in the Department of CSE (Artificial Intelligence and Machine Learning) at Asansol Engineering College, Asansol-713305, India. He has done his B.Tech. in Electronics and Communication Engineering from West Bengal University of Technology, Kolkata, India, M.Tech. in VLSI Design and Microelectronics Technology from Jadavpur University, Kolkata, India, and Ph.D. (Engg.) in Electronics and Communication Engineering from Brainware University, Kolkata, India in 2005, 2010, and 2023, respectively. He has a rich academic experience of more than twenty years. He has published sixteen utility and design Indian patents and authored eleven SCI/SCIE/Scopus-indexed international journals, seventeen national/international Scopus-indexed conference proceedings, and ten Scopus-indexed book chapters. He is the author of five textbooks. He is the recipient of the “Best Researcher Award 2022” by Novel Research Academy, India. His current research interests include sensors, optical detectors, artificial intelligence, cloud computing, soft computing, and IoT. He can be contacted at email: bhaskar.ec11@gmail.com.



Batyrkhan Omarov    received his bachelor's and master's degrees from Al-Farabi Kazakh National University, Almaty, Kazakhstan in 2008 and 2010, respectively. In 2019, he received his Ph.D. from Tenaga National University, Kuala Lumpur, Malaysia. His research interests include machine learning, deep learning, data science, image processing, medical image analysis, and artificial intelligence in medicine. He can be contacted at email: batyahan@gmail.com.



Guna Sekhar Sajja    works as an SAP consultant in the retail industry and is a research scholar at the University of the Cumberland's, USA. He has more than 10 years of work experience as an SAP consultant and covered in these areas as a supply chain management expert (manufacturing, medical, agriculture, and retail sector). His educational qualifications include a Ph.D. in Information Technology, complemented by master's and bachelor's degrees in Information Systems security and Mechanical Engineering. He published many research papers in reputed national and international journals and conference papers, and book chapters. He worked as a reviewer for Springer, Elsevier, InderScience, Web of Science, Hindawi (Wiley), PeerJ, and Taylor & Francis journals. He has a track record of publications that underscores my commitment to research and innovation, particularly in AI, machine learning, and their applications within supply chain management. An enterprising, dynamic go-getter with solid presentation, relationship-building, analytical and problem-solving skills, ensuring optimal utilization of resources, processes, and technology. He can be contacted at email: guna.sajja@gmail.com.