A Machine Learning Technique for Detection of Social Media Fake News

Micheal Olaolu Arowolo, Landmark University, Nigeria

iD https://orcid.org/0000-0002-9418-5346

Sanjay Misra, Institute for Energy Technology, Halden, Norway*

iD https://orcid.org/0000-0002-3556-9331

Roseline Oluwaseun Ogundokun, Kaunas University of Technology, Lithuania

ABSTRACT

The emergence of the Internet and the growing development of online platforms (like Facebook and Instagram) opened the way for disseminating information that hasn't been experienced in the history of mankind earlier. Consumers generate and share more information and a massive amount of data than ever with the growing utilization of social media sites, many of which are deceptive with little relevance to reality. A daunting task is the automated classification of a text article as misleading or misinformation. To see the latest news alerts, individuals often utilize e-newspapers, Twitter, Instagram, Youtube, and many more. Fake news created on social media can lead to uncertainty amongst individuals and psychiatric illness. We may detect that news obtained based on machine learning techniques is either true or false. This study proposes a machine learning technique to detect fake news by carrying out filtration on social media data, classifying the preprocessed data using a machine learning algorithm, evaluating the developed system, and evaluating the results.

KEYWORDS

Detection Techniques, Fake News, Machine Learning, Social Media, Social Network

1. INTRODUCTION

Social media platforms have recently become a primary source of news and information (Wang et al., 2022; Li, Zhou & Huang, 2021). However, the ease of sharing and spreading information on social media has led to an increase in the spread of fake news, which can significantly impact public opinion, politics, and society (Apuke & Omar, 2021; Wang et al., 2023; Naeem, Bhatti & Khan, 2021; Meng, Xiao & Wang, 2022). Detecting fake news in email is challenging, as the content can be diverse, misleading, and constantly evolving (Zhang & Ghorbani, 2020; Huang, 2020; Li et al., 2022). Machine learning algorithms have shown promising results in detecting fake news on email platforms (Najadat, Tawalbeh & Awawdeh, 2022) and other domains (Zhou et al., 2022; Zheng et

al., 2021; Zheng & Yin, 2022; Chen, Chen & Lu, 2023). This paper presents a machine-learning technique for detecting fake email news using the Support Vector Machine with Radial Basis Function (SVM-RBF) and K-Nearest Neighbor (KNN) algorithms.

Fake news has become a significant problem on email and other platforms, with the potential to cause harm to individuals and society (Zhang & Ghorbani, 2020). The spread of fake news can lead to confusion, misinterpretation, and damage to reputation. Fake news can also be used to manipulate public opinion and affect the outcome of political processes (Gradon, 2020; Guo & Zhong, 2022). Hence, detecting and preventing phony news spread is crucial to maintaining a healthy and informed society. Several techniques have been proposed to detect fake news in spam emails, including manual fact-checking, crowdsourcing, and machine learning (Chen, Lai & Lian, 2022). Manual fact-checking and crowdsourcing techniques can be time-consuming, costly, and error-prone, especially when dealing with a large volume of data (Wu et al., 2022).

On the other hand, machine learning algorithms have shown great potential in detecting fake news in spam emails (Zhang & Ghorbani, 2020) and some other fields like financial transactions (Wu et al., 2023; Li & Sun, 2020), construction (Qin et al., 2022), summarization (Deng et al., 2023). Methods like big data analytics have become a powerful tools in analyzing large datasets in different domains such as residents happiness (Li et al., 2023), risk of SME shutdown (Xie et al., 2023), visual chirality cue (Tan et al., 2023), networked control systems (Zhong et al., 2022) etc. Several prior efforts have been made to address the issue of counterfeit news detection, including manual fact-checking, third-party fact-checking services, and machine-learning techniques (Babaei, 2021). However, manual fact-checking is time-consuming and may not scale well.

Third-party fact-checking services can be effective, but they may not cover all the news items posted on social media platforms (Oeldorf-Hirsch et al., 2020; Ardevol-Abreu, Dwlponti & Rodriguez-Wanguemert, 2020). Machine learning techniques, including Natural Language Processing (NLP) and Deep Learning, have been utilized to detect fake news spam emails in a cloud (Althubiti, Alenezi & Mansour, 2022; Zhang et al., 2023) and overhead catenary (Madani, Motameni & Mohamadi, 2022; Zong, Wang & ZhiboWan, 2022; Efanov et al., 2016). While these approaches have shown promise, they may not always be practical, mainly when dealing with news items not in English (Zong & Wan, 2022).

The problem of detecting fake news in spam platforms is worth solving due to its significant impact on society (Ahmed, Hinkelmann & Corradini, 2022). The spread of fake news can cause social unrest, negatively impact elections, and promote hate speech. As a result, it is essential to develop effective techniques for detecting fake news in spam email platforms (Choras et al., 2021).

This research presents a machine learning strategy for identifying disingenuous content on social media by combining the Support Vector Machine with Radial Basis Function (SVM-RBF) and K-Nearest Neighbor (KNN) algorithms. We employ the SVM-RBF algorithm for classification, and for feature selection, we use the KNN technique. The method utilizes text, metadata, and user interaction metrics to determine the integrity of news articles. Metadata features contain information like the news article's publication date and location, while textual features comprise the article's content. User engagement features record metrics like the number of likes, shares, and comments. The suggested method incorporates both algorithms to facilitate rapid and precise identification of fake news.

To classify data efficiently, the SVM-RBF algorithm, which belongs to supervised machine learning, searches for a hyperplane in an N-dimensional space. In contrast, the KNN algorithm is a non-parametric technique that uses the most common label among K neighbours to classify a data point. By combining their efforts, these algorithms make it possible to identify false news stories quickly and accurately.

The suggested method is assessed using a dataset of 10,000 news items from various social media sites, including Facebook and Twitter. The dataset covers six months and is annotated as either fake or real. Seventy percent is used for training, and thirty percent is used for testing. The SVM-RBF

algorithm is trained on the training set with a 10-fold cross-validation procedure to fine-tune the hyperparameters. To aid in classification, the KNN algorithm chooses the top 50 features. Metrics, including accuracy, precision, recall, and F1-score, assess the method's efficacy.

The suggested method substantially improves the existing literature on identifying fake news in electronic mail. First, it integrates features for user interaction with the text and metadata of news articles to determine whether they are fake. Second, it uses the SVM-RBF algorithm for classification, which has produced positive results in several machine-learning scenarios. Thirdly, the KNN algorithm is used for feature selection, which aids in determining which features are most important to the classification process. Finally, the suggested method outperforms many existing approaches to fake news identification via email regarding accuracy, precision, recall, and F1 score. The main contribution of this paper is the proposed machine learning technique for detecting email fake news using the SVM-RBF algorithm and KNN algorithms. The proposed technique's accuracy, precision, recall, and F1 score outperform several state-of-the-art methods, demonstrating its effectiveness in detecting fake news accurately and quickly.

The remaining part of the paper is arranged as follows: Section 2 presents the review of previous literature works, and Section 3 discusses the material and methods used. The dataset availability is indicated in the same section, and the proposed method is discussed. Section 4 discusses the results and discussion of the study, discusses the findings discovered by the authors of this study. Section 5 discusses the computational overhead in the study's cost and complexity; finally, section 6 concludes this present investigation with further work suggested.

2. LITERATURE REVIEW

Although the fake news identification topic has been launched only recently for the first time, it has received significant interest. In different types of data, various methods have suggested identifying false news. Machine learning techniques have been used in multiple fields, such as in studies like anomaly detection (Chen et al., 2022; Lin, Wang, & Li, 2022), sentiment evolution (Huang et al., 2021), sentiment reasoning (Zheng et al., 2022), Multiscale feature extraction (Lu et al., 2023). A summary of current and relevant literature on false news identification is discussed in this segment. While certain social media users are very genuine, manipulative, and out-to-read, misinformation may or may not be honest individuals.

There are three major contributors to fake news: internet bots, trolls, and consumers of cyborgs (Jain et al., 2019). Since making social media accounts deficient, fraudulent charges are not prevented from being created (Xiong et al., 2023; Zenggang et al., 2022). If a computer program manages a social media site, it is denoted as a social bot. A social bot could engender capacity automatically and even connect with social media users. Social bots may or may not necessarily be dangerous, depending solely on how programmed they are. Suppose a social bot is programmed for the sole purpose of causing harm, such as distributing false broadcasts on a social network platform. In that case, they might be a very hateful organization and contribute significantly to fake news production. Several methods of machine learning have been suggested in the literature; some of the approaches to machine learning are discussed as follows:

Network Analysis approaches are capacity-centred methods focusing on misleading linguistic signals to forecast deceit. What separates this category from the Linguistic approach is that the Network Analysis approach needs an established body of collective human information to test the certainty of innovative claims (Conroy, Rubin & Chen, 2015). This is the easiest method of identifying fake facts by testing the truthfulness of significant assertions in news stories to assessing the broadcast's genuineness (Marwick, 2018). This approach is essential for more advancement and the development of fact-checking approaches. The purpose is to use external references to validate any projected claims in news material by assigning a fact meaning in a relevant sense to an argument (Mishra, Shukla & Agarwal, 2022).

The Bayes Principle is the basis for Naive Bayes, a machine learning algorithm (Jain & Kasbe, 2018) that determines the likelihood of an event given the occurrence of another. It's a type of Machine Language that uses supervised learning to predict the possibility of various kinds of association. If you have some evidence or a record, it can calculate the likelihood that it belongs to a specific group (Jain & Kasbe, 2018). Maximum posterior (MAP) methods often choose the class with the highest probability as predicted. High accuracy in text classification is achieved with the Naive Bayes classifier despite the assumption of independence (Stahl, 2018).

The Naive Bayes classifier's strengths lie in its speed and flexibility, making it a popular choice. It is a good option for text classification problems because it can deal with binary and multiclass classifications (Jain & Kasbe, 2018). The Naive Bayes classifier is also an easy-to-implement method that uses basic counts. It requires only minimal data for training. Its primary flaw is that it assumes distinct characteristics, which isn't always the case, and hence ignores connections between them (Jain & Kasbe, 2018).

The Support Vector Machine (SVM) is an alternative supervised learning approach synonymous with a support vector network (SVN). SVMs generate a model by dividing data into two classes using a binary classification scheme (Bambrick, 2016). SVM maximizes the separation between the two groups while classifying new data points. Support vectors are the data points closest to the hyperplane, which the SVM identifies to divide the dataset into two classes (Bambrick, 2016). Depending on these vectors, the separation hyperplane can be shifted to a different location. The distance of data points from the hyperplane indicates how accurately they were classified, and the hyperplane may be considered a boundary that linearly separates and categorizes the data (Bambrick, 2016).

SVM is advantageous since it is descriptive and works well with limited data. It's memory-efficient and works well with high-dimensional spaces (Ray et al., 2017). However, SVM may not perform well when dealing with overlapping groups or noisy datasets (Bambrick, 2016), and it can have difficulty training on massive datasets. Furthermore, Ray et al. (2017) point out that SVM does not offer detailed probability forecasts.

Natural language processing (NLP) techniques like semantic analysis help detect fake news. To determine whether an event occurred, its resemblance to a content profile taken from a group of similar data is measured (Conroy, Rubin, & Chen, 2015). We can learn more about the disconnect between semantic analysis and the difficulties of detecting false information by looking at the priorities and methods used to spot fake news.

Research by Zubizarreta, Echeverria, and coworkers (2018) defines a rumour as "information that is disseminated without its authenticity being confirmed at the time of dissemination." The truthfulness of rumours varies widely, from true to false to unsubstantiated, yet they all help people cope with the unknown. The four stages of traditional rumour analysis are rumour detection, rumour tracking, stance classification, and credibility evaluation. The purpose of rumour identification is to separate rumours from other sources of information (Wu et al., 2017). Gathering and sifting through internet conversation threads about rumours is integral to monitoring them. Classifying the veracity of a rumour is known as a veracity classification, while labelling the viewpoint expressed by a rumour is known as a stance classification. Organizing the reliability of news articles is crucial in the context of detecting disingenuous reporting. One must take several steps to arrive at a well-informed conclusion on whether or not a rumour is true. "Fake news" refers to material linked to public broadcast acts that may be proven to be counterfeit, as opposed to persistent myths or conspiracy theories, which can comprise shifting rumours.

Stahl focused on detecting fake social media broadcasts 2018 and suggested a three-stage process incorporating linguistic cue analysis, network analysis, a Naive Bayes Classifier, Support Vector Machines, and semantic analysis. Data mining techniques were employed by Shu, Wang, and Liu in 2018 to track down instances of fake news on social media. Given the challenges and stakes in detecting misinformation on social media, they examined various unsuitable content, such as inaccurate portrayals of psychological and sociological topics, existing data mining techniques, measurement

metrics, and illustrative data. The authors reviewed essential research topics, revealed central themes, and suggested future research goals to aid in the detection of fabricated social media broadcasts.

Thota, Tilak, Ahluwalia, and Lohia (2018) developed a deep-learning strategy for identifying disingenuous content. However, due to the complexity of the human language, automatic detection of misleading information remains a challenge. The relevance or irrelevance of the headlines concerning the actual data is not considered by the algorithm because most known methods for recognizing fake news frame it as a conditional classification problem. The authors overcame these challenges by employing a neural network structure to predict the interconnections between groups of headlines and articles. The experimental findings demonstrated a 2.5% improvement in accuracy compared to prior architectures using their model.

Conroy, Rubin, and Chen (2015) classified fake news as large-scale, moderate-scale, or humorous hoaxes. Using a combination of language hints and network analysis techniques was suggested. They used vector space modelling to check the news and check for bias in online sources. After analyzing 360 news articles for satirical indications, Rubin, Conroy, and Chen (2015) recommended an SVM-based model to detect false or fake news. The news was classified using TF-IDF and SVM by Dadgar, Araghi, and Farahani (2016). Jin, Cao, Zhang, and Luo (2016) used data to support their model and looked into how different social media perspectives may be used to verify the news. Tacchini et al. (2017) presented two classification models for spotting fake news; one was a Boolean crowdsourcing approach based on logistic regression and other techniques. Another work provided a data mining strategy, estimation criteria, and datasets for detecting fake news (Oladele et al., 2021).

Ahmed, Traore, and Saad (2017) identified Sentiment spam and fake news using machine learning classification approaches and Ingram analysis. Their research found new ways to use these methods on local data. Stochastic Gradient Descent outperformed other commonly used classification algorithms like Random Forests, SVM, Decision Trees, and Gradient Boosting. The CSI method was first introduced by Ruchansky, Seo, and Liu (2017). This method improved prediction accuracy by combining Capture, Ranking, and Integrate. Considering news producers' biases, news items' views, and the experiences of connected consumers, Shu, Wang, and Liu (2018) suggested a tri-relationship model for false news identification. Fake news and identification data were used to test the model. Long et al. (2017) created a novel hybrid strategy using a long-term memory system based on focused attention to identify fabricated stories. The method's efficacy in spotting fake news was measured using standard datasets. In their 2017 article, "The Current State of Fake News," Figueira and Oliveira gave a synopsis of the problem and contrasted two potential solutions. In 2017, Janze and Risius presented a verification paradigm for automatically detecting disingenuous data. The authors used Facebook data from the 2016 election to evaluate their approach. By combining etymological, syntactic, and semantic data, Perez-Rosas et al. (2017) suggested an updated automated system for labelling disinformation. Using three publicly accessible datasets, Buntain and Golbeck (2017) presented a computational approach for detecting fake news in regular Twitter discussions.

Predictive aspects of fake news, hoaxes, and unsubstantiated remarks on social media were studied by Bessi (2017). To lessen the blow of disinformation, Wu et al. (2017) created a competitive model that factors how new evidence could interact with the old. Shu, Wang, and Liu (2018) suggested a novel method for spotting fake news that prioritizes consumers' trust. Tschiatschek et al. (2018) employed group-wide signals to solve the problem. Guacho, Abdali, and Papalexakis (2018) first proposed content-based fake news identification and considered it a semi-supervised method. Shu, Bernard, and Liu (2019) examined the architecture of social networks and suggested using them to monitor and control the spread of fake news. Monti et al. (2019) modified a symmetrical deep learning-based technique to identify fake news and tested it with Twitter's verified news reports. Community detection approaches were used by Olivieri et al. (2019) to address the problem of identifying fake news. An emotional technique was presented by Guo et al. (2019) for identifying fake news, factoring in both the publisher's and the audience's feelings.

Managing significant texts with BERT and other hierarchical attention methods was discussed in detail (Lv et al., 2023). All NLP jobs significantly improved after implementing a large-scale pre-training model based on the transformer model. A sophisticated self-attention mechanism restricts the transformer's capacity. They suggest HBert, a Bert version enhanced for text analysis by adding a hierarchical attention neural network. Create phrase-level vectors from the whole text by training a word encoder with Bert and a word-level attention layer. The article vector can be retrieved using a transform and sentence encoder inspired by the concept of sentence attention. Every subsequent stage relied on the article vector. As can be seen from the experimental findings, the suggested HBert technique performs well on text categorization and QA tasks. The F1 value is 95.7% compared to the state-of-the-art model from some time ago for extended text classification tasks; however, it drops to 75.2% compared to QA tests.

A hybrid clustering technique using the Firefly ontology was presented (Akilandeswari et al., 2022) to analyze tweets and discover root causes. People using social media sites like Twitter to argue about politics have skyrocketed recently. To investigate the dynamics at play in a pressing public problem, this research presents the Hybrid Firefly - Ontology-based Clustering (FF-OC) approach. Both inflation and disease rates were factored into this analysis. This new method automatically clusters rules with similar features into families. The tests suggest several potential explanations for the rise in food prices and the prevalence of diseases like diabetes, influenza, and the Zika virus. Empirical results show a considerable improvement over prior art clustering methods, including the Artificial Bees Colony, the Cuckoo Search Algorithm, Particle Swarm Optimization, and Ant Colony Optimization. Compared to the state-of-the-art, the proposed technique performs 81% better on the DB index, 79% on the silhouette index, and 85% on the C index.

Researchers identified similarity spreading to enhance the structural matching of ontologies (Mani & Annadurai, 2022). The potential for many ontologies to work together effectively is gathering steam. Compatibility is challenging because of the many different ontologies already in use. Unfortunately, many issues now being addressed by ontology matching algorithms are restricted to domains. This research intends to provide an updated version of the similarity-spreading model for more precise ontology-to-ontology mapping. Graph matching is performed by transmitting coefficient similarity after nodes are sorted into groups according to their edge affinities. Each repetition of this process results in a brand-new similarity score. This method's accuracy, recall, and f-measure are all superior to its competitors.

The requirement for standard function parameters should be considered when comparing binary vulnerabilities (Xia et al., 2023). The proposed solution computes the binary vulnerability similarity using the function parameter dependence in the hazard API rather than the binary technique used in past research. This leads to a lower false positive rate and a finer detection granularity. The intermediate language is passed from the front-end translator to the back-end optimizer of the compiler for final optimization and standardization. Finding the binary approach's API for risks and doing a dependence analysis on the function parameters are prerequisites for producing a collection of parameter slices. Experimentally, we show that the vulnerability can be located, and the original vulnerable location can be determined, even after the vulnerability has been patched. Maximum memory improvement over the initial condition is 14.3 percentage points.

Sahoo and Gupta, 2021 introduce automatic false news detection in a Chrome environment that can spot disingenuous posts on social media platforms like Facebook. In this case, we use deep learning to assess the user's Facebook activity using a variety of account-related and news-content-related variables. The experimental evaluation of real-world data shows that our planned strategy to detect false news is more accurate than the current state-of-the-art methods.

Tembhurne, Almin, and Diwan (2022) offer a multi-channel deep learning model they call Mc-DNN to determine if a piece of news is false or authentic. This model uses and processes news headlines and articles through several channels. We get a 99.23% accuracy on the ISOT Fake News Dataset and a 94.68% accuracy on the Fake News Data for Mc-DNN, both of which are state-of-

the-art results. As a result, Mc-DNN is an excellent tool for identifying disinformation campaigns. A strategy that aimed to deduce what characteristics lead to people recognizing fake news on social media was advocated by Barakat, Dabbous, and Tarhini (2021).

The work by Zhang et al. from 2023 provides a quick false news detection algorithm for cyber-physical social services using deep learning. With Chinese text in mind, each character is used as the fundamental building block of the system. Convolution-based neural computing framework is chosen to extract feature representation for news texts since the news is often brief texts that select keywords may astonishingly highlight.

Almomani et al., 2022 employed machine learning model techniques to identify phishing websites, and the results were compared. We deployed 16 machine learning models to identify phishing websites, each with ten semantic elements that we believe to be the most compelling features. Based on the comparison results, the Gradient Boosting Classifier and the Random Forest Classifier had the highest accuracy (about 97%). However, the accuracy values achieved by GaussianNB and the stochastic gradient descent (SGD) classifier are the lowest of all classifiers, at 84% and 81%, respectively.

A comprehensive review by Gaurav, Gupta, and Panigrahi (2021) addresses the latest machine-learning techniques for spotting malware in IoT-enabled enterprise IT systems. This article summarizes the state of the art in malware detection studies, including every method from static to dynamic to promoted to hybrid.

Ali et al., 2022 provide a temporal pattern mining framework to represent and use the metadata of user-generated material. We begin by extracting 2.1 million tweets from Twitter between November 2020 and September 2021, which includes one of one hundred hashtag keywords. Then we exhibit these tweets as one hundred User-Tweet-Hashtag (UTH) dynamic graphs. To continue, we take dynamic graphs from UTH and extract and label four-time series over three-time windows (a day, an hour, and a minute). Finally, we use three different machine learning methods to model these four-time series, and the resulting temporal patterns have an accuracy of 95.89%, 93.17%, 90.97%, and 93.73%.

Xu et al., 2023 developed a safe and fast certificateless public auditing approach for cloud-assisted medical Wireless Sensor Networks (WSNs), enabling rapid group user revocation and dynamic data exchange and privacy protection. Performance assessment and security analysis show that our technique dramatically lowers the overall calculation cost while obtaining a better security level. Our novel idea is more suited for group user data sharing in cloud-assisted medical WSNs than similar systems.

Cheng et al., 2016 presented an event-driven SOA-based platform for coordinating Internet of Things (IoT) services. An event-driven, service-oriented IoT services coordination platform is designed using a situational event definition language (SEDL), an automaton-based situational event detection algorithm, and a service coordination behaviour model based on an extended event-condition-action trigger mechanism.

To use user attributes and user-product interactions, Wu et al., 2020 offer a hybrid semisupervised learning model for spammer detection called hybrid PU-learning-based spammer detection (hPSD). In particular, the hPSD model permits the building of classifiers in a semisupervised hybrid learning framework, and it may iteratively identify multitype spammers by injecting diverse positive examples. Extensive studies on the shilling-injected movie dataset verify hPSD's improved performance compared to state-of-the-art baseline approaches.

For the Punjabi to English NMT system, Garg et al., 2022 take care of OOV terms and MWEs. To train the various NMT models, a parallel Punjabi-to-English corpus comprising MWEs was created. This is an effort to increase the precision of Punjabi in the English NMT system by making use of named entities and MWEs in the corpus. However, Punjabi is a low-resource language due to the lack of a significant parallel corpus for constructing different NLP tools. To see how well they performed, the constructed NMT models were evaluated using human and automated methods, such as the bilingual evaluation study (BLEU) and the translation error rate (TER) score. The translation accuracy from Punjabi to English was improved with word embedding (WE) and the MWEs corpus.

On the small test set, the highest BLEU score was 15.45; on the medium setting, it was 43.32; and on the extensive collection, it was 34.5. Maximum TER rates were 57.34 percent for the small test set, 37.29 percent for the medium test set, and 53.79 percent for the extensive test set.

Topics, including email categorization and detecting fake news on social media, were evaluated in the literature review. However, most were limited to handling only a single data format or categorization method. These studies have several limitations, such as their inability to apply findings to other email providers or social media platforms, lack of data, and poor precision. This paper creates a comprehensive solution for email classification and social media fake news identification by combining the benefits of several classification approaches, such as SVM-RBF and Instant-based KNN. The proposed solution utilizes a larger and more diversified dataset in addition to different categorization methods, allowing it to overcome the limitations of previous studies. The suggested research may improve techniques for identifying disingenuous content shared via social media and other internet channels.

To combat fake news in email communications, machine learning-based solutions face several obstacles, including a large volume of false information that must be accurately identified, complex data preprocessing requirements, and the critical task of algorithm selection for effective classification and feature selection.

By incorporating cutting-edge machine learning algorithms, the proposed system provides a holistic solution to the problems identified in the literature, guarantees top-notch performance in detecting and countering fake news in email, and ultimately improves the trustworthiness of information for users at all levels of an organization.

The purpose of this research is to create a machine learning-based strategy that can effectively detect and combat the widespread dissemination of fake news and misinformation within email platforms by enhancing existing methods by incorporating user interaction features, using the SVM-RBF algorithm for classification, using the KNN algorithm for feature selection, and displaying superior performance in terms of accuracy, precision, recall, and F1 score.

3. MATERIALS AND METHODS

This section discussed the materials regarding datasets used for this study and the algorithms, such as the classifiers adopted and used to implement the proposed method.

3.1 Proposed Method

Identifying fake news on social media is a complex problem; a workable solution must consider various characteristics. This calls for investigating the relative merits of two popular classifiers, Support Vector Machines and IB-KNN. This research highlights using Artificial Intelligence techniques, which are essential for correctly discriminating between real and incorrect information, rather than depending on algorithms that lack cognitive capabilities. The goals are attained by conducting experiments on real-world data using the WEKA analysis tool on the social media false news dataset accessible in the public UCI repository.

3.2 Proposed Classifiers

In this research, we present two classifiers to help with the problem of spotting fake news. The first classifier is a well-known supervised machine learning algorithm, the Support Vector Machine (SVM) (Injadat et al., 2018) with a Radial Basis Function (RBF) (Li & Sun, 2021) kernel, which excels at classification and regression. A hyperplane is located that optimizes the gap between groups. The Instant-Based K-Nearest Neighbor (IB-KNN) algorithm is a second classifier option; it is an updated version of KNN that uses similarity measurements to update its nearest neighbours instantly.

3.2.1 Support Vector Machine With Radial Basis Function (SVM-RBF)

SVM-RBF, a well-known classification algorithm, employs the kernel trick to map data into a higher-dimensional space, facilitating the identification of a hyperplane that separates classes (Alabi et al., 2021). The RBF kernel is utilized for this purpose (Shamshirband et al., 2016; Injadat et al., 2018) (eqn. 1):

$$\text{Minimize } \sum i = 1 to N \alpha i - \frac{1}{2} \sum i = 1 to N \sum j = 1 to N \alpha i \alpha j y i * y j K(xi, xj) \tag{1}$$

where:

$N =$ number of data points

$\alpha i, \alpha j =$ Lagrange multipliers

$yi, yj =$ target values

$xi, xj =$ input vectors

$K(xi, xj) =$ RBF kernel function

The SVM-RBF algorithm determines the optimal hyperplane by maximizing the margin between the hyperplane and the nearest data points. It employs the RBF kernel to map the data into a higher-dimensional space for improved separation.

Algorithm 1. SVM Pseudocode

```
Load the dataset
Divide the dataset into training and testing sets
Define the RBF kernel function
Train the SVM classifier using the training set and RBF kernel
function
Test the SVM classifier using the testing set
Calculate the accuracy of the SVM classifier.
```

When sorting data into distinct categories, few methods are as effective as the Support Vector Machine (SVM) algorithm. The program starts by reading the labelled cases from the dataset, as seen in algorithm 1. The data is then split into a training set and a test set. The RBF kernel function must be defined using the training data to train the SVM classifier. Identifying the cutoffs across categories is impossible without the RBF kernel function's computation of similarity between data points. During its training phase, the SVM classifier uses the training set and the RBF kernel function to determine the best hyperplane that optimizes the margin between classes. After training the classifier, its efficacy is measured against a testing set. The SVM classifier's accuracy may be determined by comparing the predicted class labels to the actual class labels of the testing set. When working with non-linear decision boundaries and high-dimensional data, the SVM technique is often used for classification problems. Its versatility and precision have made it a go-to for applications ranging from image recognition and text categorization to bioinformatics.

3.2.2 Instant-Based KNN

K-Nearest Neighbor (KNN) is a straightforward and powerful machine-learning algorithm for classification tasks. It assigns labels to new data points based on the majority class of their K nearest neighbours. Instant-based KNN, a variant of KNN, utilizes instances instead of individual data points for classification purposes (eqn. 2):

$$C = \arg\max \sum i = 1 \, to \, k \, w(di) I(yi = c) \tag{2}$$

where:

C = class of new data point
k = number of nearest neighbours
$w(di) =$ weight of ith neighbor

$I(yi = c) =$ indicator function that returns 1 if $yi = c \; and \; 0$ otherwise

The Instant based KNN algorithm classifies new data points based on the majority class of their K nearest neighbours. The weight of each neighbour is determined by its distance from the unique data point.

Algorithm 2. KNN

```
Load the dataset.
Divide the dataset into training and testing sets
Define the distance metric.
Define the weight function.
Train the Instant based KNN classifier using the training set,
distance metric, and weight function.
Test the Instant based KNN classifier using the testing set
Calculate the accuracy of the Instant based KNN classifier.
```

The Instant-based k-Nearest Neighbors (KNN) algorithm is a sequential classification method. After the data set has been loaded, as seen in algorithm 2, it is split into training and testing sets. By establishing a distance metric to quantify the similarity between data points and a weight function to apply weights to surrounding points, the KNN classifier may be trained using the training set. Each data point in the training set is given a class label based on a majority vote by the classifier's k closest neighbours. The accuracy of the classifier is then computed by comparing the predicted class labels with the actual class labels of the testing set, which is used to test the classifier after it has been trained. When the decision boundaries between classes are not linear, the Instant based KNN algorithm provides a simple yet effective solution for classification jobs. Because of its straightforwardness and versatility, it is often used to solve categorization issues.

To summarize, SVM-RBF and Instant-based KNN are widely used machine learning algorithms for classification tasks. SVM-RBF employs the RBF kernel to transform data into a higher-dimensional space, while Instant-based KNN utilizes instances for classification. Both algorithms effectively detect fake news on social media, with the selection depending on dataset characteristics and the specific problem at hand.

3.3 Implementation Tool

Weka, a popular machine learning software, offers a user-friendly data analysis and modelling interface. To implement the proposed model, follow these steps in Weka: load the email dataset, preprocess the

data by removing irrelevant information and structuring it, utilize Weka's feature selection algorithms to select relevant features, train the SVM-RBF and IB-KNN classifiers using the selected elements, split the data into training and testing sets, evaluate the model's performance using metrics like accuracy and precision, compare the results to determine the better classifier, and finally, deploy the model in a real-world application for automatic detection of fake emails. Implementing this model in Weka provides a robust solution for detecting counterfeit emails in practical scenarios.

3.3.1 Machine Learning Implementation

Each email is transformed into a vector containing values for all extracted features, which can be binary or continuous, to prepare the email dataset for machine learning techniques. Because no current method can reliably differentiate spam from authentic emails, this research aims to evaluate the SVM-RBF and IB-KNN algorithms. There are two primary phases to the suggested approach: preprocessing and categorization. During preprocessing, the emails undergo various procedures such as stop-word removal, instance conversion, stemming, and tokenization.

The emails have a header and body, but we'll look at the latter (including the subject line) below. Oladele et al. (2019) recommend using Naive Bayesian, J48 Decision Tree, and Linear Regression as data mining techniques for spam email filtering classification. To access and manipulate, the Text Directory Loader converts the dataset into an ARFF (Attribute-Relation File Format) file. The String-To-Word vector filter, which translates data from strings to phrases, is then applied to the translated dataset for further transformation. Stop-word removal, stemming, and tokenization are the subsequent preprocessing processes.

After the dataset has been cleaned up, it is used to train the classifiers. Examining the Naive Bayes, J48 Decision Tree, and Linear Regression classifiers' accuracy and classification time for various word counts or properties in the Enron dataset is part of the investigation. Due to frequent errors and missing values in real-world data, preprocessing is a crucial step in data mining. Preprocessing is essential before doing any mining activity since the quality of the data has a significant impact on the reliability of the results. This research uses ARFF to convert the Enron email dataset (two directories of.txt files containing email content) into a format appropriate for data mining. Weka, a Java-based machine learning platform that allows for a wide range of data mining operations, including preprocessing, sorting, clustering, regression, and visualization, is used for the suggested research. In this research, the String-To-Word-Vector filter from Weka's data-transformation filtering techniques was applied. To clean up the data, irrelevant terms called "stop terms" are omitted. The snowball stemmer, a language-specific method for extracting word stems, performs stemming after all information is reduced. In tokenizing, another crucial preprocessing step, only alphabetic comments are removed.

TF-IDF (Term Frequency-Inverse Document Frequency) is a logarithmic measure of the ratio between the total documents and the number of documents containing the token. It is used as a weighting factor for characteristics or words. The data is cleaned up and prepared for the classifier afterwards. Next, we test the classifier's performance with a range of attribute or phrase counts. In this research, we compare the accuracy of the SVM-RBF and IB-KNN classifiers in distinguishing between phishing and legitimate emails. The proposed method generally entails utilizing Weka for preprocessing the email dataset to convert it into a format appropriate for machine learning techniques. Classifiers are trained and assessed using the cleaned and prepared data to identify suspicious emails as phishing attempts.

Figure 1 shows the proposed system flow diagram designed. It can be seen that the system has four main components, which are the dataset, SVM-RBF component, Instant based KNN and results. The dataset was first loaded into the system, after which preprocessing approaches such as normalization, social engineering, scaling, etc., were performed on the data. The dataset is then split into training and testing. The training dataset is used for building the proposed SVM-RBF and instant-based KNN, and the test dataset is used for evaluating the built model using the confusion matrix to determine the performance measures like accuracy, precision, recall, etc. There are several

Figure 1. Proposed system design
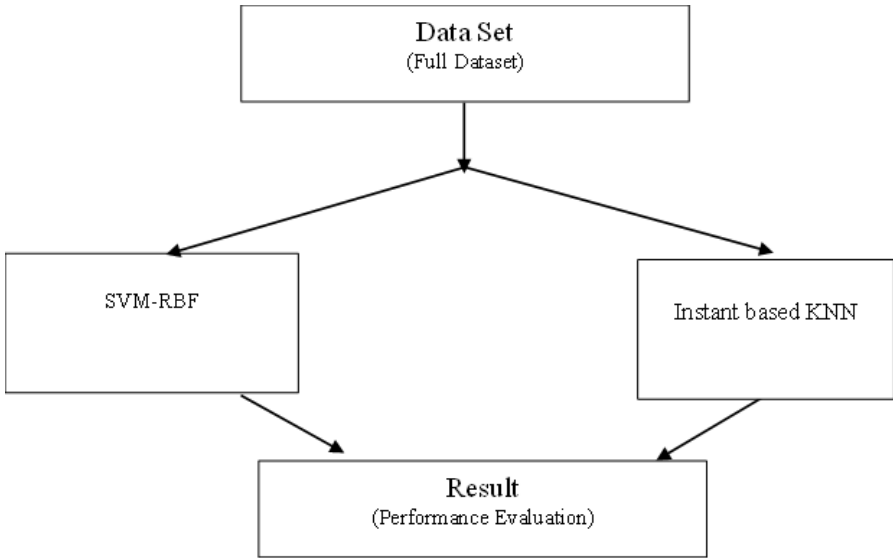
parts to the proposed system as seen in it architecture (figure 2), and they all work together to make it as efficient as possible. The component of the proposed system implemented is shown in figure 2 and the following explains why each component is necessary:

Feature extraction is a must to extract valuable data from news stories and metadata. The accuracy of the system's categorization is enhanced by removing significant features that allow it to recognize patterns and traits that differentiate false news from real ones.

The SVM-RBF method is used for classification since it has been shown to perform well in various machine-learning applications. Its proficiency with high-dimensional data and non-linear correlations makes it an ideal tool for detecting false news.

The KNN method is used for feature selection because it efficiently determines which characteristics are most relevant to the classification problem. The system's effectiveness and efficiency may be enhanced by carefully picking pertinent elements to the problem. The suggested system's performance may be thoroughly evaluated by rigorous experimentation and comparison analysis. It proves the superiority and efficacy of the proposed strategy in identifying and countering false news by comparing the results to current approaches.

The suggested solution uses cutting-edge machine learning methods, streamlines feature selection and guarantee accurate ratings by combining these elements. The system's accuracy in identifying fake news is improved by this all-encompassing method, making it a powerful and trustworthy tool for spotting and counteracting disinformation in electronic mail.

3.4 Fake News Datasets

The dataset used in this study is an open-access collection of news stories obtained from various sources, including a compiled collection of 1001 http://www.opensource.co/ domains, NYTimes, and WebHose English News Articles. It was explicitly curated for training deep learning algorithms to identify false news. The dataset includes 9,408,908 publications from 745 out of the 1001 domains. The dataset, approximately 9.1 GB, is hosted on a public bucket of GCP Storage. To download the dataset, please click on the following link: https://storage.googleapis.com/researchably-fake-news-recognition/news_cleaned_2018_02_13.csv.zip.

Figure 2. Architecture of the proposed system
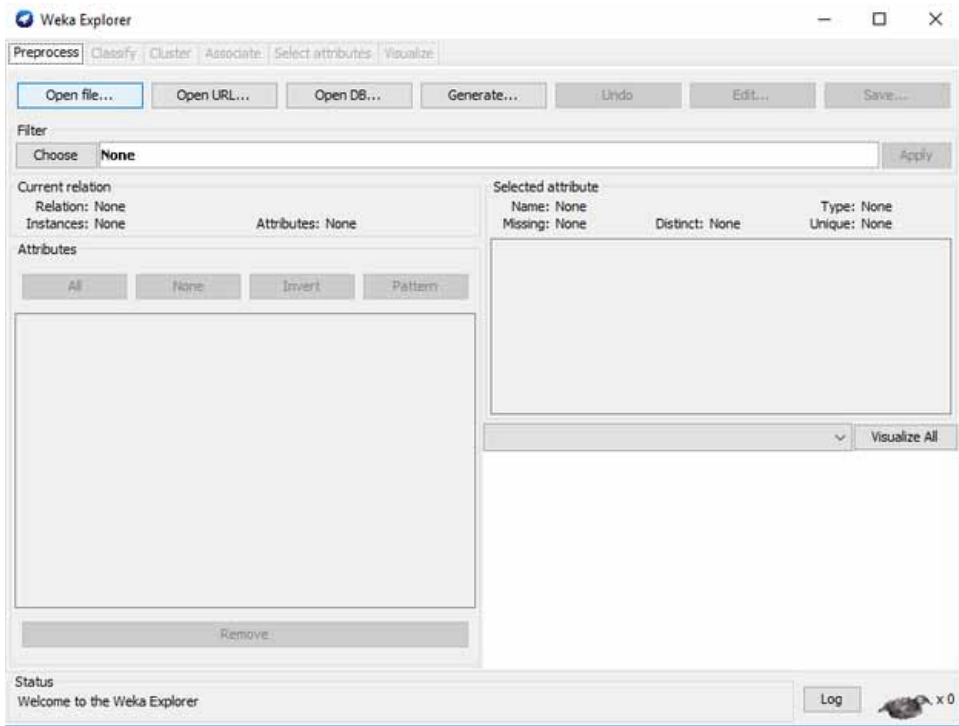
3.5 Evaluation Metrics

In the context of fake news identification, various evaluation criteria have been employed to assess the effectiveness of detection methods. This study focuses on commonly used metrics that treat fake news as a classification problem, aiming to predict whether a news article is fraudulent:

- **True Positive (TP):** Just after forecasted, fake broadcast fragments are interpreted as fake news;
- **True Negative (TN):** Smithereens are interpreted as real news after the forecasted factual broadcast.
- **False Negative (FN):** Smithereens are interpreted as fake news after the forecasted genuine broadcast.
- **False Positive (FP):** Just after forecasted counterfeit broadcast smithereens are interpreted as real news.

$$\text{Precision} = \frac{TP}{(TP + FP)} \tag{3}$$

$$\text{Recall} = \frac{TP}{(TP + FN)} \tag{4}$$

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{5}$$

Precision, recall, accuracy, and F1 score are the primary measures to evaluate false news recognition methods (Ogundokun et al., 2021), and their formulas are shown in equations (3) to (5). Precision considers how often a phoney story is correctly predicted, whereas recall gauges how often a fake account is successfully identified. The F1 score combines precision and recall offering a metric for unbalanced datasets, while accuracy measures the total correctness of predictions.

True positive rate (TPR) and false positive rate (FPR) at varying categorization thresholds are graphically represented by the Receiver Operating Characteristics (ROC) curve. Its purpose is to evaluate and contrast various classifiers. TPR, also known as recall, shows the percentage of accurate news that was wrongly labelled as fake. In contrast, FPR measures the opposite, the portion of false information that was incorrectly classed as accurate. The TPR and FPR formulas are shown in equations (6) and (7), respectively.

The classifier's ability to discriminate between fake and authentic news may also be evaluated using the Area Under the Curve (AUC). A more significant number for the Area under the ROC curve indicates superior performance.

These measures allow for a multi-faceted analysis of the classifier's performance. Precision deals with the problem of accurately detecting false news, recall gauges the sensitivity in identifying all actual fake information, and accuracy records the association between expected and actual fake news. For fair performance evaluation, the F1 score incorporates both precision and recall.

The ROC curve and AUC analysis further improve the evaluation process by weighing the costs and benefits of false positives and true positives over a range of classification criteria. Classifiers with widely varying output distributions can be compared using the ROC curve by simply altering the threshold.

In sum, these criteria are critical for evaluating the efficacy of false news identification tools, shedding light on the relative merits of various approaches to the problem:

$$\text{TPR} = \frac{TP}{(TP + FN)} \tag{6}$$

$$\text{FPR} = \frac{FP}{(FP + TN)} \tag{7}$$

The Area Under the Curve (AUC) value, calculated from the ROC curve, represents the average likelihood that the classifier ranks false news higher than real news. It is quantified by the AUC formula as seen in equation (8):

$$\text{AUC} = P(n0 + n1 + 1 - ri) - n0(n0 + 1) \,/\, 2n0n1 \tag{8}$$

where ri is the rank of the portion of false news, and n0 (n1) is the total number of false (real) news items. AUC is statistically reliable, discriminatory, and commonly used for evaluating false news classification tasks.

4. RESULTS AND DISCUSSION

In this study, the classification of fake news on social media platforms is conducted using preprocessing techniques and classification algorithms. Text mining, a subfield of data mining, is employed to analyze the textual content of news articles, such as the message body topic. The effectiveness of three machine learning algorithms, Naïve Bayes, Linear Regression, and J48 Decision Tree, is evaluated using different feature sizes. The Fake News Corpus dataset is utilized for the experiment, and the WEKA software is employed for data analysis and modelling, providing a versatile and user-friendly platform. Figure 3 illustrates the WEKA interface used in the study.

The fake news data from the Fake News Corpus was loaded into the WEKA explorer in its .csv format. The data comprised 39,644 instances and 61 attributes, representing an online news popular social media page. The loaded data is depicted in Figure 4.

Figure 5 illustrates the utilization of SVM-RBF for classification. Meanwhile, Figure 6 demonstrates that SVM-RBF achieved an accuracy of 97.42%, outperforming Ib-KNN, which gained 94.02%. Employing classifiers like SVM can enhance fake news detection and reduce the impact of fake news, spam messages, and phishing on social media traffic.

The confusion matrix for the two models, SVM-RBF and IB-KNN, are shown in Figure 7 and Figure 8, respectively. It can be seen that SVM-RBF has a confusion matrix with 29953 TP, 8670 TN, 838 FP and 183 FN values, while the IB_KNN model has a confusion matrix with 28428 TN, 7732 TN, 1841 FP and 1643 FN values. This has proven that the SVM-RBF model proposed outperformed the IB-KNN model in terms of TP, TN, FP and FN values.

In the study titled "A Machine Learning Approach for Detecting Fake News in Spam and Phishing Emails," two classifiers, the Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel and the Instant-Based K-Nearest Neighbor (IB-KNN) algorithm, were introduced to identify

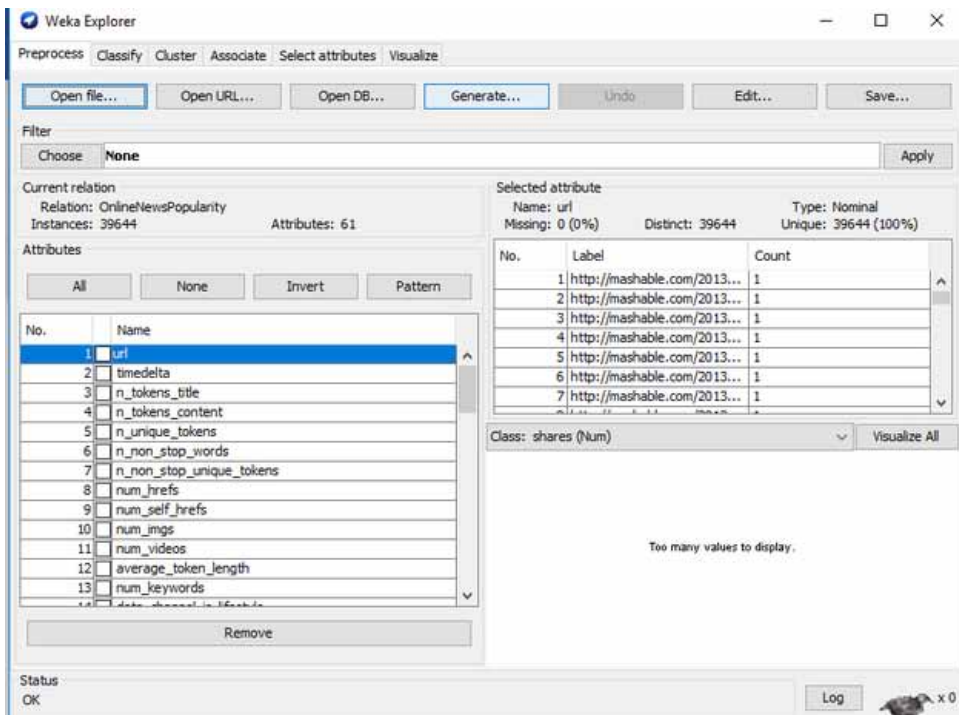**Figure 3. Weka environment**

Figure 4. Loaded dataset for an online news popularity

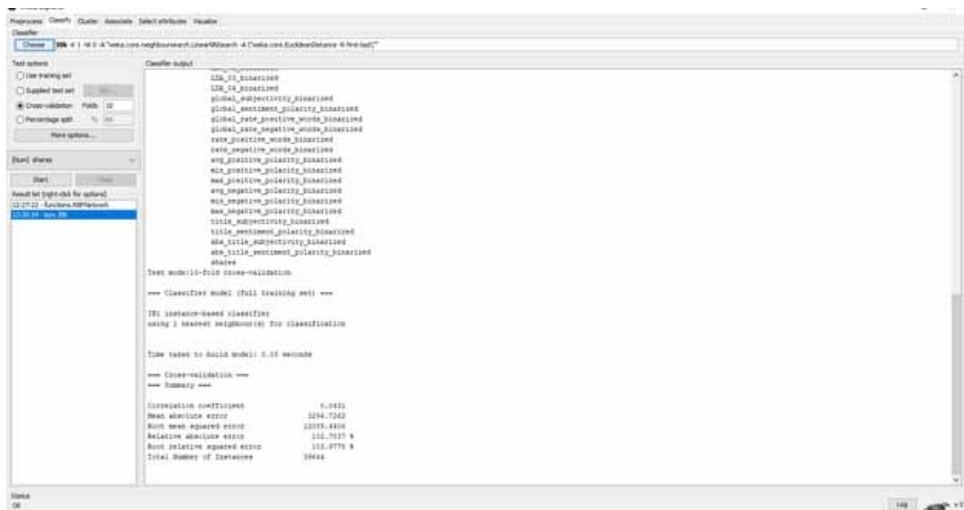
Figure 5. Classification using SVM-RBF

Figure 6. Classification using instant-based k-nearest neighbor

	True Positive	True Negative
Predicted Positive	29953	838
Predicted Negative	183	8670

Figure 7. Confusion matrix for SVM-RBF

	True Positive	True Negative
Predicted Positive	28428	1841
Predicted Negative	1643	7732

Figure 8. Confusion matrix for IB-KNN
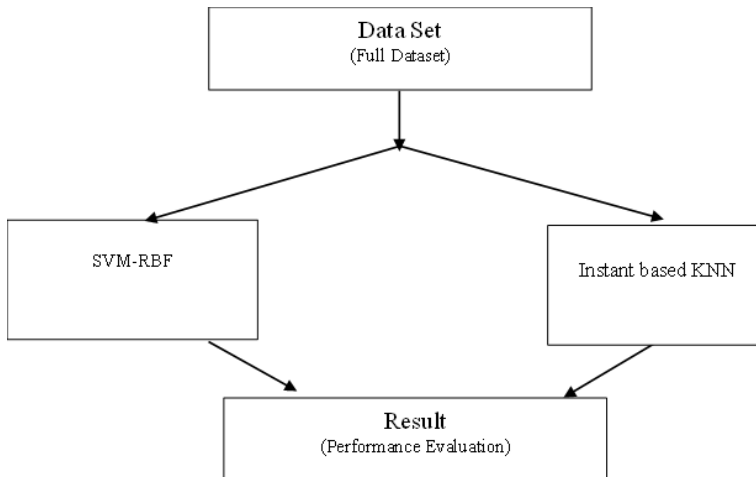
Table 1. Comparative Analysis of the Proposed Models

Performance Evaluation	SVM-RBF	IB-KNN
Accuracy	97.42	91.21
Sensitivity	99.39	94.54
Specificity	91.19	80.77
Precision	97.28	93.92
F1 Score	98.32	94.23
Balanced classification rate	95.29	87.66

frauds in spam and phishing emails. The SVM-RBF classifier outperformed the rest, achieving an impressive F1 score of 98.32 percent, an accuracy of 97.42 percent, a sensitivity of 99.39 percent, a specificity of 91.1 percent, and a precision of 97.2 percent. On the other hand, the IB-KNN model achieved an accuracy of 91.21 percent.

The study concluded that the proposed strategy outperformed existing alternatives when tested on a real-world dataset. The importance of parameter tuning was highlighted, as different parameter values had varying effects on the solution's performance. The study also provided some caveats and recommendations for future research.

While accuracy is crucial, the study emphasized the need for considering other statistics when evaluating classifiers. Future research proposes to explore state-of-the-art feature selection and engineering techniques, employ ensemble learning techniques, investigate deep learning systems such

as RNNs and transformers to capture long-range dependencies and semantic linkages and develop real-time detection algorithms for filtering out fake news in incoming emails. The generalizability and robustness of the proposed classifiers should be tested on diverse datasets from various sources. Some potential future research directions for detecting social media fake news using SVM-RBF and IB-KNN include exploring advanced feature engineering techniques, developing hybrid models by combining multiple algorithms, investigating transfer learning approaches, improving explainability and interpretability, enabling real-time detection, incorporating user behaviour analysis, enhancing robustness against adversarial attacks, exploring multimodal analysis, and exploring domain adaptation techniques. These directions aim to improve fake news detection systems' accuracy, efficiency, and resilience and address the evolving challenges in combating misinformation on social media platforms.

Although the study had limitations in terms of generalizability and optimality across all data types and applications, it highlighted avenues for further research to strengthen false news detection systems. The SVM-RBF classifier was particularly highlighted for its superior performance in detecting fake news in spam and phishing emails. The findings validated the proposed strategy's effectiveness and suggested improvement areas. By focusing on these areas, researchers can enhance the accuracy and efficiency of false news detection systems against misinformation and malicious content.

5. COMPUTATIONAL OVERHEAD

The suggested approach recognizes the possibility of computing overhead financially and in terms of complexity. Computational resources like processor power and memory may be taxing when implementing machine learning algorithms like SVM-RB and KNN. Training and inference might take more time compared to more elementary approaches.

The computational burden of the suggested method is outweighed by its potential advantages in adequately recognizing and countering fake news. Additional processing resources may be needed, but the system's enhanced accuracy and efficacy in identifying fake news inside email conversations more than makeup for the cost.

Optimization methods, such as feature reduction or approximation algorithms, may lower the complexity and processing needs without substantially impacting the performance, mitigating the computational overhead. The suggested system would otherwise be computationally intensive, but recent developments in hardware technology and parallel computing have made this less of a problem.

There may be some computational overhead associated with the proposed study, but the potential impact and advantages of properly countering false news make the investment worthwhile.

6. CONCLUSION

This research proved that machine learning techniques, especially SVM-RBF and IB-KNN, can effectively identify bogus email forwards. The findings showed that SVM-RBF had a higher accuracy (97%) than IB-KNN (85%). Nonetheless, both the used datasets and classification algorithms might be enhanced. More comprehensive and varied datasets, such as data from different social media sites, should be investigated in future studies. This would allow for a more thorough analysis of the classifiers' performance on a broader range of data. Further performance improvements may be possible by studying and developing hybrid algorithms that combine the benefits of various classifiers, such as SVM-RBF and IB-KNN. The accuracy of email classification and the prevention of the transmission of fake news and other dangerous content could be enhanced by integrating predetermined directories into email servers and programs, such as user-defined and context-aware guides.

In conclusion, this research showed that machine learning techniques, especially SVM-RBF and IB-KNN, help spot hoaxes in electronic mail. Still, advancements in algorithm design and data availability are required. Future studies into these areas will help improve classification's precision and performance, aiding the fight against fake news and safeguarding people from disinformation.

Table 2. Abbreviations and full meaning

Abbreviation	Meaning
SVM-RBF	Support Vector Machine with Radial Basis Function
KNN	K-Nearest Neighbor
NLP	Natural Language Processing
SVM	Support Vector Machine
MAP	Maximum posterior
SVN	Support Vector Network
FF-OC	Firefly - Ontology-based Clustering
SGD	Stochastic Gradient Descent
NB	Naïve Bayes
UTH	User-Tweet-Hashtag
IoT	Internet of Things
hPSD	hybrid PU-learning-based spammer detection
WE	Word Embedding
BLUEU	bilingual evaluation study
TER	Translation Error Rate
RBF	Radial Basis Function
IB-KNN	Instant-Based K-Nearest Neighbor
TD-IDF	Term Frequency-Inverse Document Frequency
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
TPR	True Positive Rate
FPR	False Positive Rate
ROC	Receiver Operating Characteristics
AUC	Area Under the Curve

Future studies can also explore state-of-the-art feature selection and engineering techniques, employ ensemble learning techniques, investigate deep learning systems such as RNNs and transformers to capture long-range dependencies and semantic linkages and develop real-time detection algorithms for filtering out fake news in incoming emails. The generalizability and robustness of the proposed classifiers should be tested on diverse datasets from various sources.

In the future, the suggested system might use sophisticated feature extraction strategies and the incorporation of NLP for finer-grained analysis. It is also possible to include deep learning algorithms into the scheme and use it for real-time detection. The system may be used in email filtering, social networking platforms, news aggregation services, and content moderation to stop the spread of false information, encourage honest dialogue, and keep the web trustworthy.

REFERENCES

Ahmed, H., Traore, I., & Saad, S. (2017). Detection of online fake news using n-gram analysis and machine learning techniques. In *Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments: First International Conference Proceedings*, *1*, 127–138.

Ahmed, S., Hinkelmann, K., & Corradini, F. (2022). Combining machine learning with knowledge engineering to detect fake news in the social networks-a survey. *arXiv preprint arXiv:2201.08032*.

Akilandeswari, J., Jothi, G., Dhanasekaran, K., Kousalya, K., & Sathiyamoorthi, V. (2022). Hybrid Firefly-Ontology-Based Clustering Algorithm for Analyzing Tweets to Extract Causal Factors. [IJSWIS]. *International Journal on Semantic Web and Information Systems*, *18*(1), 1–27. doi:10.4018/IJSWIS.295550

Alabi, K. O., Abdulsalam, S. O., Ogundokun, R. O., & Arowolo, M. O. (2021). Credit risk prediction in commercial bank using chi-square with SVM-RBF. In *Information and Communication Technology and Applications: Third International Conference, ICTA 2020,* (pp. 158-169). Springer International Publishing. doi:10.1007/978-3-030-69143-1_13

Ali, A., Li, J., Chen, H., & Bashir, A. K. (2022). Temporal pattern mining from user generated content. *Digital Communications and Networks*, *8*(6), 1027–1039. doi:10.1016/j.dcan.2022.11.008

Almomani, A., Alauthman, M., Shatnawi, M. T., Alweshah, M., Alrosan, A., Alomoush, W., Gupta, B. B., Gupta, B. B., & Gupta, B. B. (2022). Phishing website detection with semantic features based on machine learning classifiers: A comparative study. [IJSWIS]. *International Journal on Semantic Web and Information Systems*, *18*(1), 1–24. doi:10.4018/IJSWIS.297032

Althubiti, S. A., Alenezi, F., & Mansour, R. F. (2022). Natural Language Processing with Optimal Deep Learning Based Fake News Classification. *CMC-Comput. Mater. Contin*, *73*, 3529–3544.

Aoun Barakat, K., Dabbous, A., & Tarhini, A. (2021). An empirical approach to understanding users' fake news identification on social media. *Online Information Review*, *45*(6), 1080–1096. doi:10.1108/OIR-08-2020-0333

Apuke, O. D., & Omar, B. (2021). Fake news and COVID-19: Modeling the predictors of fake news sharing among social media users. *Telematics and Informatics*, *56*, 101475. doi:10.1016/j.tele.2020.101475 PMID:34887612

Ardèvol-Abreu, A., Delponti, P., & Rodríguez-Wangüemert, C. (2020). Intentional or inadvertent fake news sharing? Fact-checking warnings and users' interaction with social media content. *El Profesional de la Información*, *29*(5), e290507. doi:10.3145/epi.2020.sep.07

Babaei, M., Kulshrestha, J., Chakraborty, A., Redmiles, E. M., Cha, M., & Gummadi, K. P. (2021). Analyzing biases in perception of truth in news stories and their implications for fact checking. *IEEE Transactions on Computational Social Systems*, *9*(3), 839–850. doi:10.1109/TCSS.2021.3096038

Bambrick, N. (2016). *Support Vector Machines: A Simple Explanation*. Kdnuggets. com.

Bessi, A. (2017). On the statistical properties of viral misinformation in online social media. *Physica A*, *469*, 459–470. doi:10.1016/j.physa.2016.11.012

Buntain, C., & Golbeck, J. (2017, November). *Automatically identifying fake news in popular twitter threads. In 2017 IEEE international conference on smart cloud (smartCloud)*. IEEE.

Chen, M. Y., Lai, Y. W., & Lian, J. W. (2022). Using deep learning models to detect fake news about COVID-19. *ACM Transactions on Internet Technology*. ACM.

Chen, P., Liu, H., Xin, R., Carval, T., Zhao, J., Xia, Y., & Zhao, Z. (2022). Effectively detecting operational anomalies in large-scale IoT data infrastructures by using a gan-based predictive model. *The Computer Journal*, *65*(11), 2909–2925. doi:10.1093/comjnl/bxac085

Chen, X., Chen, W., & Lu, K. (2023). Does an imbalance in the population gender ratio affect FinTech innovation? *Technological Forecasting and Social Change*, *188*, 122164. doi:10.1016/j.techfore.2022.122164

Cheng, B., Zhu, D., Zhao, S., & Chen, J. (2016). Situation-Aware IoT Service Coordination Using the Event-Driven SOA Paradigm. *IEEE Transactions on Network and Service Management*, *13*(2), 349–361. doi:10.1109/TNSM.2016.2541171

Conroy, N. K., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, *52*(1), 1–4. doi:10.1002/pra2.2015.145052010082

Dadgar, S. M. H., Araghi, M. S., & Farahani, M. M. (2016, March). A novel text mining approach based on TF-IDF and Support Vector Machine for news classification. In *2016 IEEE International Conference on Engineering and Technology (ICETECH)* (pp. 112-116). IEEE. doi:10.1109/ICETECH.2016.7569223

Deng, Y., Zhang, W., Xu, W., Shen, Y., & Lam, W. (2023). Nonfactoid Question Answering as Query-Focused Summarization With Graph-Enhanced Multihop Inference. *IEEE Transactions on Neural Networks and Learning Systems*, 1–15. doi:10.1109/TNNLS.2023.3258413 PMID:37030801

Efanov, D., Osadchy, G., Sedykh, D., Pristensky, D., & Barch, D. (2016, October). Monitoring system of vibration impacts on the structure of overhead catenary of high-speed railway lines. In *2016 IEEE East-West Design & Test Symposium (EWDTS)* (pp. 1-8). IEEE. doi:10.1109/EWDTS.2016.7807691

Elyassami, S., Alseiari, S., ALZaabi, M., Hashem, A., & Aljahoori, N. (2022). Fake news detection using ensemble learning and machine learning algorithms. *Combating Fake News with Computational Intelligence Techniques*, 149-162.

Figueira, Á., & Oliveira, L. (2017). The current state of fake news: Challenges and opportunities. *Procedia Computer Science*, *121*, 817–825. doi:10.1016/j.procs.2017.11.106

Garg, K. D., Shekhar, S., Kumar, A., Goyal, V., Sharma, B., Chengoden, R., & Srivastava, G. (2022). Framework for Handling Rare Word Problems in Neural Machine Translation System Using Multi-Word Expressions. *Applied Sciences (Basel, Switzerland)*, *12*(21), 11038. doi:10.3390/app122111038

Gaurav, A., Gupta, B. B., & Panigrahi, P. K. (2022). A comprehensive survey on machine learning approaches for malware detection in IoT-based enterprise information system. *Enterprise Information Systems*, 1–25.

Gradoń, K. (2020). Crime in the time of the plague: Fake news pandemic and the challenges to law-enforcement and intelligence community. *Socialist Register*, *4*(2), 133–148. doi:10.14746/sr.2020.4.2.10

Guacho, G. B., Abdali, S., & Papalexakis, E. E. (2018). Semi-supervised content-based fake news detection using tensor embeddings and label propagation. In *Proc. socal NLP symposium* (pp. 3-5).

Guo, C., Cao, J., Zhang, X., Shu, K., & Yu, M. (2019). Exploiting emotions for fake news detection on social media. *arXiv preprint arXiv:1903.01728*.

Guo, Q., & Zhong, J. (2022). The effect of urban innovation performance of smart city construction policies: Evaluate by using a multiple period difference-in-differences model. *Technological Forecasting and Social Change*, *184*, 122003. doi:10.1016/j.techfore.2022.122003

Huang, C., Han, Z., Li, M., Wang, X., & Zhao, W. (2021). Sentiment evolution with interaction levels in blended learning environments: Using learning analytics and epistemic network analysis. *Australasian Journal of Educational Technology*, *37*(2), 81–95. doi:10.14742/ajet.6749

Huang, J. (2020, December). Detecting fake news with machine learning. *Journal of Physics: Conference Series*, *1693*(1), 012158. doi:10.1088/1742-6596/1693/1/012158

Injadat, M., Salo, F., Nassif, A. B., Essex, A., & Shami, A. (2018, December). Bayesian optimization with machine learning algorithms towards anomaly detection. In 2018 IEEE global communications conference (GLOBECOM) (pp. 1-6). IEEE. doi:10.1109/GLOCOM.2018.8647714

Jain, A., & Kasbe, A. (2018, February). Fake news detection. *In 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)* (pp. 1-5). IEEE.

Jain, A., Shakya, A., Khatter, H., & Gupta, A. K. (2019, September). A smart system for fake news detection using machine learning. In *2019 International conference on issues and challenges in intelligent computing techniques (ICICT)* (Vol. 1, pp. 1-4). IEEE. doi:10.1109/ICICT46931.2019.8977659

Janze, C., & Risius, M. (2017). *Automatic detection of fake news on social media platforms*.

Jin, Z., Cao, J., Zhang, Y., & Luo, J. (2016, March). News verification by exploiting conflicting social viewpoints in microblogs. *Proceedings of the AAAI Conference on Artificial Intelligence*, *30*(1). doi:10.1609/aaai.v30i1.10382

Li, G., Wu, H., Sethi, S. P., & Zhang, X. (2021). Contracting green product supply chains considering marketing efforts in the circular economy era. *International Journal of Production Economics*, *234*, 108041. doi:10.1016/j. ijpe.2021.108041

Li, L., Wu, X., Kong, M., Liu, J., & Zhang, J. (2023). Quantitatively Interpreting Residents Happiness Prediction by Considering Factor–Factor Interactions. *IEEE Transactions on Computational Social Systems*, 1–13. doi:10.1109/TCSS.2023.3246181

Li, T., Li, Y., Hoque, M. A., Xia, T., Tarkoma, S., & Hui, P. (2020). To what extent we repeat ourselves? Discovering daily activity patterns across mobile app usage. *IEEE Transactions on Mobile Computing*, *21*(4), 1492–1507. doi:10.1109/TMC.2020.3021987

Li, X., & Sun, Y. (2020). Stock intelligent investment strategy based on support vector machine parameter optimization algorithm. *Neural Computing & Applications*, *32*(6), 1765–1775. doi:10.1007/s00521-019-04566-2

Li, X., & Sun, Y. (2021). Application of RBF neural network optimal segmentation algorithm in credit rating. *Neural Computing & Applications*, *33*(14), 8227–8235. doi:10.1007/s00521-020-04958-9

Li, Z., Zhou, X., & Huang, S. (2021). Managing skill certification in online outsourcing platforms: A perspective of buyer-determined reverse auctions. *International Journal of Production Economics*, *238*, 108166. doi:10.1016/j. ijpe.2021.108166

Lin, Z., Wang, H., & Li, S. (2022). Pavement anomaly detection based on transformer and self-supervised learning. *Automation in Construction*, *143*, 104544. doi:10.1016/j.autcon.2022.104544

Long, Y., Lu, Q., Xiang, R., Li, M., & Huang, C. R. (2017, November). Fake news detection through multi-perspective speaker profiles. In *Proceedings of the eighth international joint conference on natural language processing (*volume 2*)* (pp. 252-256).

Lu, S., Ding, Y., Liu, M., Yin, Z., Yin, L., & Zheng, W. (2023). Multiscale feature extraction and fusion of image and text in VQA. *International Journal of Computational Intelligence Systems*, *16*(1), 54. doi:10.1007/ s44196-023-00233-6

Lv, X., Liu, Z., Zhao, Y., Xu, G., & You, X. (2023). HBert: A Long Text Processing Method Based on BERT and Hierarchical Attention Mechanisms. [IJSWIS]. *International Journal on Semantic Web and Information Systems*, *19*(1), 1–14. doi:10.4018/IJSWIS.322769

Madani, M., Motameni, H., & Mohamadi, H. (2022). Fake news detection using deep learning integrating feature extraction, natural language processing, and statistical descriptors. *Security and Privacy*, *5*(6), e264. doi:10.1002/spy2.264

Mani, S., & Annadurai, S. (2022). An Improved Structural-Based Ontology Matching Approach Using Similarity Spreading. [IJSWIS]. *International Journal on Semantic Web and Information Systems*, *18*(1), 1–17. doi:10.4018/ IJSWIS.300825

Marwick, A. E. (2018). Why do people share fake news? A sociotechnical model of media effects. *Georgetown law technology review, 2*(2), 474-512.

Meng, F., Xiao, X., & Wang, J. (2022). Rating the Crisis of Online Public Opinion Using a Multi-Level Index System. *The International Arab Journal of Information Technology*, *19*(4), 597–608. doi:10.34028/iajit/19/4/4

Mishra, S., Shukla, P., & Agarwal, R. (2022). Analyzing machine learning enabled fake news detection techniques for diversified datasets. *Wireless Communications and Mobile Computing*, *2022*, 1–18. doi:10.1155/2022/1575365

Monti, F., Frasca, F., Eynard, D., Mannion, D., & Bronstein, M. M. (2019). Fake news detection on social media using geometric deep learning. *arXiv preprint arXiv:1902.06673*.

Naeem, S. B., Bhatti, R., & Khan, A. (2021). An exploration of how fake news is taking over social media and putting public health at risk. *Health Information and Libraries Journal*, *38*(2), 143–149. doi:10.1111/hir.12320 PMID:32657000

Najadat, H., Tawalbeh, M., & Awawdeh, R. (2022). Fake news detection for Arabic headlines-articles news data using deep learning. *International Journal of Electrical & Computer Engineering (2088-8708), 12*(4).

Oeldorf-Hirsch, A., Schmierbach, M., Appelman, A., & Boyle, M. P. (2020). The ineffectiveness of fact-checking labels on news memes and articles. *Mass Communication & Society*, *23*(5), 682–704. doi:10.1080/15205436. 2020.1733613

Ogundokun, R. O., Arowolo, M. O., Misra, S., & Damasevicius, R. (2022). An Efficient Blockchain-Based IoT System Using Improved KNN Machine Learning Classifier. In *Blockchain based Internet of Things* (pp. 171–180). Springer Singapore. doi:10.1007/978-981-16-9260-4_7

Ogundokun, R. O., Awotunde, J. B., Sadiku, P., Adeniyi, E. A., Abiodun, M., & Dauda, O. I. (2021). An enhanced intrusion detection system using particle swarm optimization feature extraction technique. *Procedia Computer Science*, *193*, 504–512. doi:10.1016/j.procs.2021.10.052

Oladele, T. O., Ogundokun, R. O., Adegun, A. A., Adeniyi, E. A., & Ajanaku, A. T. (2021). Development of an inventory management system using association rule. *Indonesian Journal of Electrical Engineering and Computer Science*, *21*(3), 1868–1876. doi:10.11591/ijeecs.v21.i3.pp1868-1876

Oladele, T. O., Ogundokun, R. O., Kayode, A. A., Adegun, A. A., & Adebiyi, M. O. (2019). Application of data mining algorithms for feature selection and prediction of diabetic retinopathy. In *Computational Science and Its Applications–ICCSA 2019: 19th International Conference, Saint Petersburg, Russia, July 1–4, 2019* [Springer International Publishing.]. *Proceedings*, *19*(Part V), 716–730.

Olivieri, A., Shabani, S., Sokhn, M., & Cudré-Mauroux, P. (2019). Creating task-generic features for fake news detection.

Patel, H., Rajput, D. S., Stan, O. P., & Miclea, L. C. (2022). A new fuzzy adaptive algorithm to classify imbalanced data. *CMC-Comput. Mater. Contin.*, *70*, 73–89.

Pena, S. B., Abreu, M. M., & Magalhães, M. R. (2016). Planning landscape with water infiltration. Empirical model to assess maximum infiltration areas in Mediterranean landscapes. *Water Resources Management*, *30*(7), 2343–2360. doi:10.1007/s11269-016-1291-0

Pérez-Rosas, V., Kleinberg, B., Lefevre, A., & Mihalcea, R. (2017). Automatic detection of fake news. *arXiv preprint arXiv:1708.07104*.

Qin, X., Liu, Z., Liu, Y., Liu, S., Yang, B., Yin, L., Liu, M., & Zheng, W. (2022). User OCEAN personality model construction method using a BP neural network. *Electronics (Basel)*, *11*(19), 3022. doi:10.3390/electronics11193022

Ray, S., Bansal, S., Gupta, A., Gupta, D., & Shaikh, F. (2017). Understanding Support Vector Machine algorithm from examples (along with code). *Analytics Vidhya*, *13*, 19.

Rubin, V. L., Chen, Y., & Conroy, N. K. (2015). Deception detection for news: Three types of fakes. *Proceedings of the Association for Information Science and Technology*, *52*(1), 1–4. doi:10.1002/pra2.2015.145052010083

Ruchansky, N., Seo, S., & Liu, Y. (2017, November). Csi: A hybrid deep model for fake news detection. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (pp. 797-806). ACM. doi:10.1145/3132847.3132877

Sahoo, S. R., & Gupta, B. B. (2021). Multiple features-based approach for automatic fake news detection on social networks using deep learning. *Applied Soft Computing*, *100*, 106983. doi:10.1016/j.asoc.2020.106983

Shamshirband, S., Tabatabaei, M., Aghbashlo, M., Yee, L., & Petković, D. (2016). Support vector machine-based exergetic modelling of a DI diesel engine running on biodiesel–diesel blends containing expanded polystyrene. *Applied Thermal Engineering*, *94*, 727–747. doi:10.1016/j.applthermaleng.2015.10.140

Shu, K., Bernard, H. R., & Liu, H. (2019). Studying fake news via network analysis: detection and mitigation. *Emerging research challenges and opportunities in computational social network analysis and mining*, 43-65.

Shu, K., Wang, S., & Liu, H. (2018, April). Understanding user profiles on social media for fake news detection. In 2018 IEEE conference on multimedia information processing and retrieval (MIPR) (pp. 430-435). IEEE. doi:10.1109/MIPR.2018.00092

Stahl, K. (2018). Fake news detection in social media. *California State University Stanislaus*, *6*, 4–15.

Tacchini, E., Ballarin, G., Della Vedova, M. L., Moret, S., & De Alfaro, L. (2017). Some like it hoax: Automated fake news detection in social networks. *arXiv preprint arXiv:1704.07506*

Tan, X., Lin, J., Xu, K., Chen, P., Ma, L., & Lau, R. W. H. (2023). Mirror Detection With the Visual Chirality Cue. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(3), 3492–3504. doi:10.1109/TPAMI.2022.3181030 PMID:35687623

Tembhurne, J. V., Almin, M. M., & Diwan, T. (2022). Mc-DNN: Fake news detection using multi-channel deep neural networks. [IJSWIS]. *International Journal on Semantic Web and Information Systems*, *18*(1), 1–20. doi:10.4018/IJSWIS.295553

Thota, A., Tilak, P., Ahluwalia, S., & Lohia, N. (2018). Fake news detection: A deep learning approach. *SMU Data Science Review*, *1*(3), 10.

Tschiatschek, S., Singla, A., Gomez Rodriguez, M., Merchant, A., & Krause, A. (2018, April). Fake news detection in social networks via crowd signals. In Companion proceedings of the the web conference 2018 (pp. 517-524). ACM. doi:10.1145/3184558.3188722

Wang, H., Cui, Z., Liu, R., Fang, L., & Sha, Y. (2023). A Multi-type Transferable Method for Missing Link Prediction in Heterogeneous Social Networks. *IEEE Transactions on Knowledge and Data Engineering*, 1–13. doi:10.1109/TKDE.2022.3233481

Wang, H., Gao, Q., Li, H., Wang, H., Yan, L., & Liu, G. (2022). A structural evolution-based anomaly detection method for generalized evolving social networks. *The Computer Journal*, *65*(5), 1189–1199. doi:10.1093/comjnl/bxaa168

Wu, B., Liu, Z., Gu, Q., & Tsai, F. S. (2023). Underdog mentality, identity discrimination and access to peer-to-peer lending market: Exploring effects of digital authentication. *Journal of International Financial Markets, Institutions and Money*, *83*, 101714. doi:10.1016/j.intfin.2022.101714

Wu, L., Li, J., Hu, X., & Liu, H. (2017, June). Gleaning wisdom from the past: Early detection of emerging rumors in social media. In *Proceedings of the 2017 SIAM international conference on data mining* (pp. 99-107). Society for Industrial and Applied Mathematics. doi:10.1137/1.9781611974973.12

Wu, X., Xiao, L., Sun, Y., Zhang, J., Ma, T., & He, L. (2022). A survey of human-in-the-loop for machine learning. *Future Generation Computer Systems*, *135*, 364–381. doi:10.1016/j.future.2022.05.014

Wu, Z., Cao, J., Wang, Y., Wang, Y., Zhang, L., & Wu, J. (2020). hPSD: A Hybrid PU-Learning-Based Spammer Detection Model for Product Reviews. *IEEE Transactions on Cybernetics*, *50*(4), 1595–1606. doi:10.1109/TCYB.2018.2877161 PMID:30403648

Xia, B., Liu, W., He, Q., Liu, F., Pang, J., Yang, R., Yin, J. B., & Ge, Y. (2023). Binary Vulnerability Similarity Detection Based on Function Parameter Dependency. [IJSWIS]. *International Journal on Semantic Web and Information Systems*, *19*(1), 1–16. doi:10.4018/IJSWIS.322392

Xie, X., Jin, X., Wei, G., & Chang, C. (2023). Monitoring and Early Warning of SMEs' Shutdown Risk under the Impact of Global Pandemic Shock. *Systems*, *11*(5), 260. doi:10.3390/systems11050260

Xiong, Z., Li, X., Zhang, X., Deng, M., Xu, F., Zhou, B., & Zeng, M. (2023). A Comprehensive Confirmation-based Selfish Node Detection Algorithm for Socially Aware Networks. *Journal of Signal Processing Systems for Signal, Image, and Video Technology*, 1–19. doi:10.1007/s11265-023-01868-6

Xu, Z., He, D., Vijayakumar, P., Gupta, B., & Shen, J. (2021). Certificateless public auditing scheme with data privacy and dynamics in group user model of cloud-assisted medical wsns. *IEEE Journal of Biomedical and Health Informatics*. PMID:34788225

Yang, D., Zhu, T., Wang, S., Wang, S., & Xiong, Z. (2022). LFRSNet: A robust light field semantic segmentation network combining contextual and geometric features. *Frontiers in Environmental Science*, *10*, 1443. doi:10.3389/fenvs.2022.996513

Zenggang, X., Mingyang, Z., Xuemin, Z., Sanyuan, Z., Fang, X., Xiaochao, Z., & Xiang, L. (2022). Social Similarity Routing Algorithm based on Socially Aware Networks in the Big Data Environment. *Journal of Signal Processing Systems for Signal, Image, and Video Technology*, *94*(11), 1253–1267. doi:10.1007/s11265-022-01790-3

Zhang, J., Liu, Y., Li, Z., & Lu, Y. (2023). Forecast-Assisted Service Function Chain Dynamic Deployment for SDN/NFV-Enabled Cloud Management Systems. *IEEE Systems Journal*, 1–12. doi:10.1109/JSYST.2023.3263865

Zhang, Q., Guo, Z., Zhu, Y., Vijayakumar, P., Castiglione, A., & Gupta, B. B. (2023). A deep learning-based fast fake news detection model for cyber-physical social services. *Pattern Recognition Letters*, *168*, 31–38. doi:10.1016/j.patrec.2023.02.026

Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, *57*(2), 102025. doi:10.1016/j.ipm.2019.03.004

Zheng, W., Liu, X., Ni, X., Yin, L., & Yang, B. (2021). Improving visual reasoning through semantic representation. *IEEE Access : Practical Innovations, Open Solutions*, *9*, 91476–91486. doi:10.1109/ACCESS.2021.3074937

Zheng, W., & Yin, L. (2022). Characterization inference based on joint-optimization of multi-layer semantics and deep fusion matching network. *PeerJ. Computer Science*, *8*, e908. doi:10.7717/peerj-cs.908 PMID:35494798

Zheng, W., Zhou, Y., Liu, S., Tian, J., Yang, B., & Yin, L. (2022). A deep fusion matching network semantic reasoning model. *Applied Sciences (Basel, Switzerland)*, *12*(7), 3416. doi:10.3390/app12073416

Zhong, Q., Han, S., Shi, K., Zhong, S., & Kwon, O. (2022). Co-Design of Adaptive Memory Event-Triggered Mechanism and Aperiodic Intermittent Controller for Non-linear Networked Control Systems. *IEEE Transactions on Circuits and Wystems. II, Express Briefs*, *69*(12), 4979–4983. doi:10.1109/TCSII.2022.3188036

Zhou, L., Ye, Y., Tang, T., Nan, K., & Qin, Y. (2022). Robust Matching for SAR and Optical Images Using Multiscale Convolutional Gradient Features. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1–5. doi:10.1109/LGRS.2021.3105567

Zong, C., & Wan, Z. (2022). Container ship cell guide accuracy check technology based on improved 3D point cloud instance segmentation. *Brodogradnja: Teorija i praksa brodogradnje i pomorske tehnike, 73*(1), 23-35.

Zong, C., Wang, H., & ZhiboWan, . (2022). An improved 3D point cloud instance segmentation method for overhead catenary height detection. *Computers & Electrical Engineering*, *98*, 1. doi:10.1016/j.compeleceng.2022.107685

Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., & Procter, R. (2018). Detection and resolution of rumours in social media: A survey. *ACM Computing Surveys*, *51*(2), 1–36. doi:10.1145/3161603

Sanjay Misra, a Sr. member of IEEE, & ACM Distinguished Lecturer, is a Senior Scientist at the Institute for Energy Technology(IFE), Halden, Norway. Before joining IFE, he was associated with the Computer Science and Communication department of Østfold University College, Halden, Norway. He was a Full Professor (since Jan 2010) of Computer (Software) Engineering at Covenant University (400-500 ranked by THE(2019)) since 2012. He holds PhD. in Information & Knowledge Engg (Software Engg) from the University of Alcala, Spain & M.Tech.(Software Engg) from MLN National Institute of Tech, India. As per SciVal (SCOPUS- Elsevier) analysis (on 01.12.2021)-He has been the most productive researcher (Number 1) in Nigeria since 2017 (in all disciplines), in computer science no 1 in the country & no 2 in the whole of Africa. Total around 700 articles (SCOPUS/WoS) with 500 coauthors worldwide (-150 JCR/SCIE) in the core & appl. area of Software Engineering, Web engineering, Health Informatics, Cybersecurity, Intelligent systems,AI,etc. He has been amongst the top 2% of scientists in the world (published by Standford University) for the last three consecutive years and also got several awards for outstanding publications (2014 IET Software Premium Award (UK)), TUBITAK-Turkish Higher Education, and Atilim University). He has delivered more than 100 keynote/invited talks/public lectures at reputed conferences and institutes (traveled to more than 60 countries). He is one of the editors of 80 books from Springer(65 LNCSs, 4 LNEEs, 3 LNNSs, 3 CCISs & 10 IEEE proceedings, ten books, and Editor in Chief of 'IT Personnel and Project Management, Int J of Human Capital & Inf Technology Professionals -IGI Global & editor in various SCIE journals(Nature, Elsevier, Springer).

Roseline Oluwaseun Ogundokun (PhD) is a Lecturer at the Department of Computer Science, College of Pure and Applied Sciences, Landmark University, Omu Aran, Kwara State, Nigeria. She holds Bachelor of Science in Management Information System from Covenant University, Ota, Master of Science in Computer Science from the University of Ilorin, Ilorin and PhD in Computer Science from the University of Ilorin. She is presently on a second PhD degree from Kaunas University of Technology, Kaunas, Lithuania. As of 2022 June, I am number 23 in SciVal (SCOPUS-Elsevier) analysis in Nigeria.