PMSST

Journal of Applied Science, Innovation & Technology (JASIT)

Journal homepage: https://prakritimitrango.com/documents-reports-and-publications/journal/

OPEN ACCESS

Review article

Unmasking social fake news: Machine learning approach

Ritika Sharma*, Gauri Kumari, Ashish Kumar Rauniyar, Gourab Das*, Rajnish Kr Mishra, Pawan K. Chaurasia

Department of Information Technology, Babasaheb Bhimrao Ambedkar University Central University, Lucknow – 226025, India

ARTICLE INFOR

Article history:
Received: 13 April 2024
Revised: 15 May 2024
Accepted: 23 May 2024
Published online: 30 June 2024

Keywords:
Fake News, Internet,
Social Media,
Articles,
Machine Learning Models.

ABSTRACT

The increasing use of social media, has led to inconsistencies in online news, causing confusion and uncertainty for consumers. The spread of the 'fake or false' news on social media platform is a matter of serious concern due to its destructive impact on social and national sector. There are a lot of on-going research works dedicated to fake or false news detection. Fake or false news and disinformation spread on social media platforms negatively affects stability and social harmony. This paper showcases 'fake news' detection models using machine learning algorithms. The paper categorizes and describes the best approaches in several landscape of 'fake news' (text) detection across different domains that include 'health, religion, crime, forged documents, jobs, and politics'. It explores into the problem's dimensions, existing methodologies, their comparative analysis, and proposes an innovative solution for the on-going battle against misinformation. In addition to creating a model with supervised ML algorithm that can classify the news as 'true or false' by using different tools. The model will undergo the feature selection methods, to experiment and chosen the best-fit features to obtain the accurate and best performance.

1. Introduction

Information that is inaccurate or misleading but is presented as factual or authentic is referred to as 'fake news'. It can take various forms, such as written articles, images, videos, or even social media posts. Fake news is typically created and spread with the intention of deceiving or manipulating the public, often for political, financial, or ideological reasons. False information frequently incites fear in individuals, interferes with society's regular operations, and puts the sustainable development of society. Characteristics of 'fake news' include: Inaccuracy, Deception, Manipulation, Sensationalism, Confirmation bias.

If tried to separate false news from true news; the most important factors to be considered such as: Accuracy and Truth, Credibility, Intent, Sources, Transparency, Consistency with Reality, Emotional Appeal. Fake news during natural and man-made disasters can spread panic, hinder response efforts, and compromise public safety. Misinformation can misguide decision-making and resource allocation, diverting attention from genuine relief initiatives and erode public trust in official communication channels. To counteract this threat, it is crucial to promote accurate, timely information from reliable sources, enhance media literacy, and use robust fact-checking measures. The impact of disaster-related fake news underscores the importance of maintaining an informed and vigilant public to navigate emergencies and foster the resilience in the face of adversity.

Concept of spreading fake or misleading information to influence and manipulate public opinion has not emerged recently; the "fake news" term gained more significance at present age. The origins of fake news can be tracked down to various historical and technological developments. Mitigation of 'fake news' is a blooming concern in today's information-driven world, as fake news can affect public opinion, undermine media credibility and disrupt social harmony. To combat this issue, various strategies and mitigation efforts have

***Corresponding Author:** *Email address: ritikasharma130201@gmail.com1 (R.Sharma); dasgaurav4401@gmail.com (G.Das)*

been implemented by individuals, organizations, and technology. These include fact-checking, AI algorithms, media literacy, and legal measures. The on-going battle to uphold the integrity of the information landscape is crucial for fostering a more informed and discerning society. It concise exploration of these strategies, shedding light on the multifaceted approaches employed by individuals, organizations, and technology to combat ‘fake news’.

As with increasing population their s also enhancement of technology to provide easer solution of the problems faced if a sector like social media which plays a great role in communication, transferring information via text/video based features. At present scenario this social media/information spreading sector being targeted by the scammers to spread “fake news” and trap people in various ways. Therefore it has become an important concern to detect and mitigate such spread of misleading information via social media/websites.

1.1 Existing text based fake news detection techniques/models

The existing fake news (text) identification and detection techniques and models that are used here include ‘logistic regression’ which is typically used to classify problems with binary solutions; decision tree which works best with small datasets and datasets with two possible outcomes; ‘random forest classifier’ which gives high accuracy robustness feature importance versatility and scalability also reduces over-fitting; and gradient boosting classifier. Here four types of models ‘logistic regression’, ‘decision tree classifier’, ‘random forest classifier’ and ‘gradient boosting classifier’ have been used to perform the detection for ‘fake news’.

The novelty of our approaches lies in the development of a machine learning framework specifically tailored to the challenges of social fake news detection. Drawing on a diverse range of features, including textual, temporal, and social network attributes, our algorithm aims to distinguish between authentic news and misinformation with greater accuracy and granularity than existing methods.

In this paper, we present our work on unmasking ‘fake news’ through a machine learning approach. We outline our objectives, describe our methodology, and discuss the implications of our findings for addressing the pervasive problem of misinformation in online environments. By shedding light on the mechanisms of fake news propagation and proposing effective detection strategies, we aim to contribute to the development of more ‘resilient’ and ‘trustworthy’ information ecosystems.

2. Literature Review

The scope of review is limited to research conducted within the past years, across various domains. This time frame ensures the focus on most recent advancements and methodologies in text-based method for identifying false news within these specific categories. According to the author (Verma et al., 2023) News is any information that informs the audience about events that are taking place and have the potential to have an impact on them individually or socially. ‘Social media’ sites developed into a popular venue for news dissemination for commercial, entertainment and political objectives in recent years. Social media is used by people to look for and consume news because it is convenient and spreads quickly. This product had both beneficial and detrimental effects. For pleasure and personal gain, people tamper with and disseminate true information as ‘false news’. As a result of a significant; volume of misleading material spreading on Facebook over the final three months of the campaign, ‘fake news’ had a crucial influence in the US presidential election at 2016 (Verma et al., 2021). Due to this tragedy, several academic and industrial organizations are now studying and attempting to control the phenomena of fake news proliferation.

Influence of fake news on the society led several scholars to utilized terminology like fake news, false news, rumour, misinformation and disinformation interchangeably. Although there is not a single, accepted definition of fake news, we may define it as any falsified or dishonest news material that leads readers to believe a falsehood. According to Klein and Wueller, “fake news” refers to the planned or intentional online propagation of incorrect information. Until a decade ago, only printed media could be used to distribute false information, but now, internet media has emerged as the simplest channel for doing so. The influence of fake news on politics, the economy, and public opinion may be detrimental (Ajao et al., 2018). Classification of news Real or false news was classified by some academics as a binary categorization, while rest viewed it as a multi-class classification, regression, or clustering problem.

According to the “author (Ahmad et al., 2020)” the inception of the ‘World Wide Web’ and the rapid advancement of social media platforms (e.g. Twitter and Facebook) have made it possible for knowledge to be disseminated in unprecedented ways. Besides other applications, news organizations benefited from the broad adoption of social media platforms by giving their subscribers access to news updated in almost real time. The news media transformed from print media to digital forms like real-time news ‘platforms’, ‘blogs’, ‘social media feeds’, and other digital media formats. ‘(Ahmad et al., 2020)’Consumers now have more access to most of the recent news at their fingertips. 70% of visitors to news websites come from “Facebook” recommendations. In their recent form, these social media platforms are very effective and helpful for enabling users to debate, share, and discuss topics like ‘democracy’, ‘education’, and ‘health’. ‘(Ahmad et al., 2020)’ However, some organizations also utilize these platforms negatively, frequently to obtain ‘financial advantage’, and occasionally to sway public opinion, influence people's attitudes, or propagate satire or ridiculousness. The phenomenon is sometimes referred to as false news.

Fortunately, a variety of ‘computational algorithms’ may be applied to identify some articles as false based just on their textual content (Liu and Wu, 2018). There are several repositories managed by researchers that flags ‘websites’ that are classified fraudulent. Most of these strategies involve fact-checking of websites like "Snopes" and "PolitiFact." The issue with these tools is that articles and websites that are fraudulent must be identified by human competence. Remarkably, the “fact-checking websites” only include stories from specific fields, like ‘politics’, and are not designed to detect false news from a variety of fields, including ‘sports’, ‘technology’ and ‘entertainment’.

Data on the 'internet' is available in various formats, including documents, movies, and audios. It might be challenging to find and classify content that has been published 'online' in an unstructured manner (such as 'news', 'articles', 'videos', and 'audios'), since this certainly needs human skill. (Gupta et al., 2021). However, anomalies that distinguish text articles that are misleading; in character from those that are based on facts can be found using computational approaches like "natural language processing" (NLP) (Gupta et al., 2021). Other methods analyse how false news is spread in contrary to legitimate news.

More precisely, the method examines the differences in how a factual article and a fraudulent one spread over a network. On a theoretical level, the response to an article may be divided into real and fake. The study investigates various characteristics in textual data that could be used to distinguish between authentic and false contents. Several machine learning algorithms are trained using a variety of 'ensemble methods' that are not fully explored in the existing literature by utilizing those properties. As the Machine learning models tend to reduce error rates by utilizing strategies like bagging and boosting, the 'ensemble learners' have demonstrated their results in a wide range of applications. These methods make it possible to train various 'machine learning algorithms' effectively and efficiently.

According to the (Palani et al., 2022), the rapid development of high-speed internet and information and communication technology (ICT), people are keen on reading engaging news based on current events on social media platforms e.g. Twitter1, Facebook2, and Weibo3. False and unconfirmed information was purposefully being spread by misinformation makers for a variety of political and economic goals. The two main categories of false news detection (FND) techniques are social context and content based techniques. While the latter, is linked to the news content of the item ('text', 'title', 'image', and 'video'), the former is more focused on user involvement statistics like, 'comments', 're-posts', and 'ratings'. The structure-based and post-based approaches are under the category of social context-based methods, respectively. (Wang et al., 2018). Whereas post-based approaches investigate the thoughts or feelings expressed by users in their 'postings', 'propagation', 'structure-based methods' focus on the patterns or trends of bogus 'news' on social media platforms. These two categories of online 'social-context' approaches encounter the following challenges as a result of the unstructured data: data analysis and collection, noisy data, and absent data. Therefore, the content-based strategy is the main emphasis of this research. The content-based approaches are easier to use and more practical for early-stage 'false news' detection.

The identification of 'fake news' is effectively improved in this work by the introduction of a new model called CB-Fake. The letters C and B in the term CB-False stand for the CapsNet and BERT models, respectively, while the word "Fake" stands for 'fake news detection'. For detecting 'fake news', an 'end-to-end framework' is created that integrates the BERT and "CapsNet" models. The "CB-Fake" model involves the mentioned steps: The pre-processing and 'vectorization' of the news pieces comes first.

Then, using BERT, the textual characteristics from the news material were collected. It makes use of transformer architecture's self-attention mechanism, to effectively extract the underlying semantic links between the words in a phrase. The use of the capsule 'neural network model', which attempts to extract educational visual elements from the images of news articles using the 'routing-by-agreement' method, is another significant addition of the proposed study. Finally, when compared to other cutting-edge algorithms in FND, a highly rich representation of data is achievable by merging high-level visual and textual information, which produced 'classification accuracy' of 93% for 'PolitiFact' and 92% for 'gossip Cop'. The ability to differentiate between bogus and true news has been implemented using the linked layer with SoftMax activation.

According to the Agarwal et al. (2023), in today's virtual world, 70% of the world has expressed their presence online. Social media platforms and applications enable and allow people to share their problems, raise their voices against the injustice, and gain public support (Ahmed et al., 2017). Governments' use social media to easily communicate their agendas and political parties use it to express their views in elections. However, there is also room for darkness and arrogance, as people can take advantage of this freedom to spread wrong feelings.

Anomaly detection in modern networks is complex due to the various types of networks with unique properties (Castillo et al., 2011). It is important for social media companies to take responsibility and prevent wrongdoing on their platforms. To do this, they need to design an equipped tool that can differentiate between true and fake news. This involves creating a dataset with both true and fake news, and training multiple learners to produce an optimal solution.

To detect 'fake news', 'social media' companies must prepare a dataset that includes both true and fake news. The dataset should contain both true and fake news (Shu et al., 2017), allowing the machine to differentiate both types of news. By doing so, social media companies can help prevent the spread of harmful news and ensure that no wrong person can harm others on their platform. Anomaly detection methods often use unsupervised approaches to detect anomalies in cases with limited labelled data and plenty of unlabelled data. To filter the dataset, an NLP algorithm can be used, ensuring a 60:40 ratio of true and false news. The classification algorithm extracts news features such as title, timing, source, location, and class. This classification helps correct the dataset and design a model based on features of correct and false news.

A trained and test dataset should be chosen in the right ratio, with an 8:2 ratio being the best for fake news detection. The model should then be prepared and the prediction algorithms, such as the MB algorithm or decision tree, should be chosen. This enables for the identification of the fake or false news and the steaming of incorrect intentions within time. The platform can recognize fake news even before it is posted, allowing necessary steps to be taken.

In high-dimensional, complex datasets, some methods rely on shallow practices that cannot keep up with the numerous interactions between structures and attributes. By preparing the model and choosing the best prediction algorithms, the platform can identify fake-ness of news and take necessary steps to prevent further damage.

In context of analysis of fake news, first step in any 'machine learning' project is data collection. This involves collecting a large dataset of news articles from various sources, such as the

LIAR dataset, which contains labelled examples of true and false statements made by politicians. The divide-count-sum mechanism was developed to address the issue of uneven crowd density distribution in photographs, making it nonsensical to count individuals while simultaneously observing the entire crowd.

Data ‘pre-processing’ is then performed to prepare the ‘data’ for ‘machine learning’, which may include tasks such as ‘data cleaning’, ‘stop words removal’, and ‘conversion of text’ into numerical vectors. Feature extraction is then done to identify ‘patterns’ in the text that may indicate the presence of fake news. Examples of features extracted include emotive language, logical fallacies, and exaggerated or sensationalist headlines. ML models like as “support vector machines”, “decision trees”, and “neural networks” can be used for fake news analysis. The model is trained on a subset of dataset and ‘validated’ on another subset to ensure accurate identification.

Model evaluation is then used to determine the model’s performance using metrics such as ‘precision’, ‘recall’, and ‘F1 score’ (Jwa et al., 2019). The trained model can then be deployed in a production environment to analyse new news articles and identify any that may be fake. According to the author ‘(Khanam et al., 2021)’ fake news is a form of ‘misleading information’ that can lead to unrest and misinformation. The organizations like the crosscheck and House of Commons project are working to combat this issue, but their scope is limited due to human manual detection. It proposes a system for ‘automated index scoring’ or ‘ratings’ for various publishers and news contexts, using a methodology to detect article authenticity based on ‘words’, ‘phrases’, ‘sources’, and ‘titles’.

The model employs “supervised Machine Learning” (ML) algorithms on an annotated ‘dataset’, which is manually classified and guaranteed. It then uses the feature selection methods to experiment and select the ‘best fit features’ for precision (Singh et al., 2022). The model tests unseen data, plots results, and detects and classifies fake articles, enabling future system integration.

According to the author (Kaliyar et al., 2021) ‘Social media platforms’ have replaced conventional print media as the main news source in the current technological era. Because social media platforms let us consume news much more quickly and allow for less controlled editing, false news is disseminated incredibly quickly and widely. Many practical techniques for detecting false news have been developed recently. These techniques combine a sequence of neural networks in order to encode ‘news content’ and ‘social context-level’ data, with ‘unidirectional text’ sequence analysis. Consequently, modelling the pertinent data of false news using a bidirectional training strategy is prioritized in order to enhance classification performance and capture long-distance and semantic connections in sentences.

3. Graphical Analysis

The graphical analysis of fake news of past years from the data of above review table is shown in Fig. 1.

Table 1 Various models and their accuracy

Objective	Models Used	Accuracy	References
To categorize genuine and fake news based on the message credibility.	MCred, CNN BERT	99.01 %	Agarwal et al. (2023)
Introducing ensemble techniques using diverse properties of the ‘linguistic feature sets’ to categorize news articles across ‘multiple domains’ as either false or true.	KNN, BAGGING, BOOSTING, LOGISTIC REGRESSION, CNN.	94%	Ahmad et al. (2020)
Detecting fake news at an early stage by analyzing both visual content and textual of the news article.	CAPSULE NEURAL NETWORK, BERT.	93%	Kaliyar et al. (2021)
Focusing on the detection of ‘fake news’ using ‘Machine Learning’ (ML) techniques.	LOGISTIC REGRESSION, NAÏVEBAYES, DECISION TREE.	97%	Khanam et al. (2021)
Detecting fake news involves a thorough review of it through two stages: characterization and disclosure.	NAÏVEBAYES, RANDOM FOREST, KNN, SVM, XGBOOST, DECISION TREE .	75%	Palani et al. (2022)
To showcase the efficiency of the model proposed (FAKEBERT) in detecting fake news.	FAKEBERT, KNN, NAÏVEBAYES, RANDOM FOREST, DECISION TREE.	98.90 %	Verma et al. (2021)

4. Domains of Fake News:

In this review the ‘fake news’ is classified into different categories.

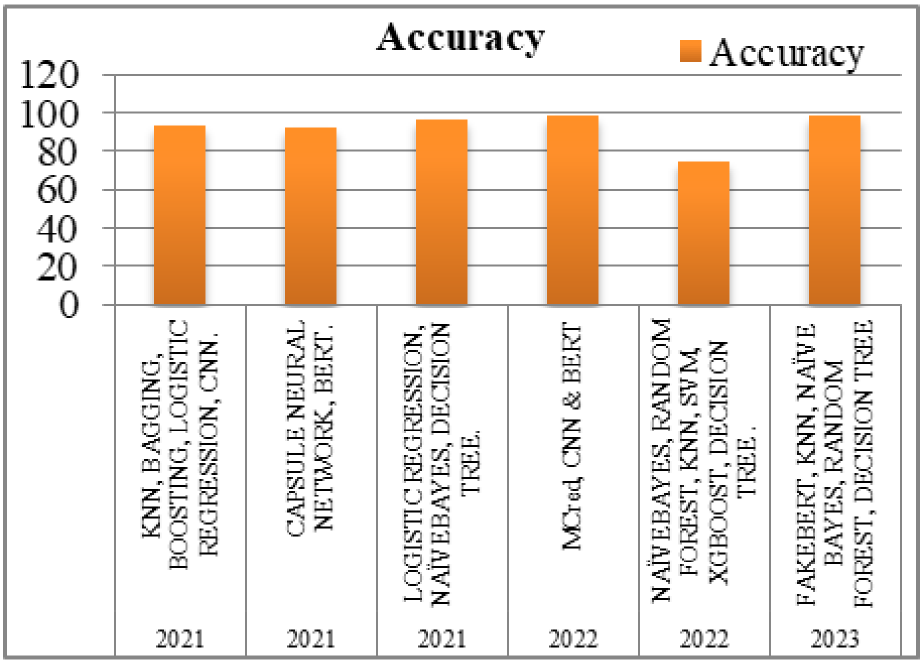

Fig. 1 Graphical presentation of accuracy of models

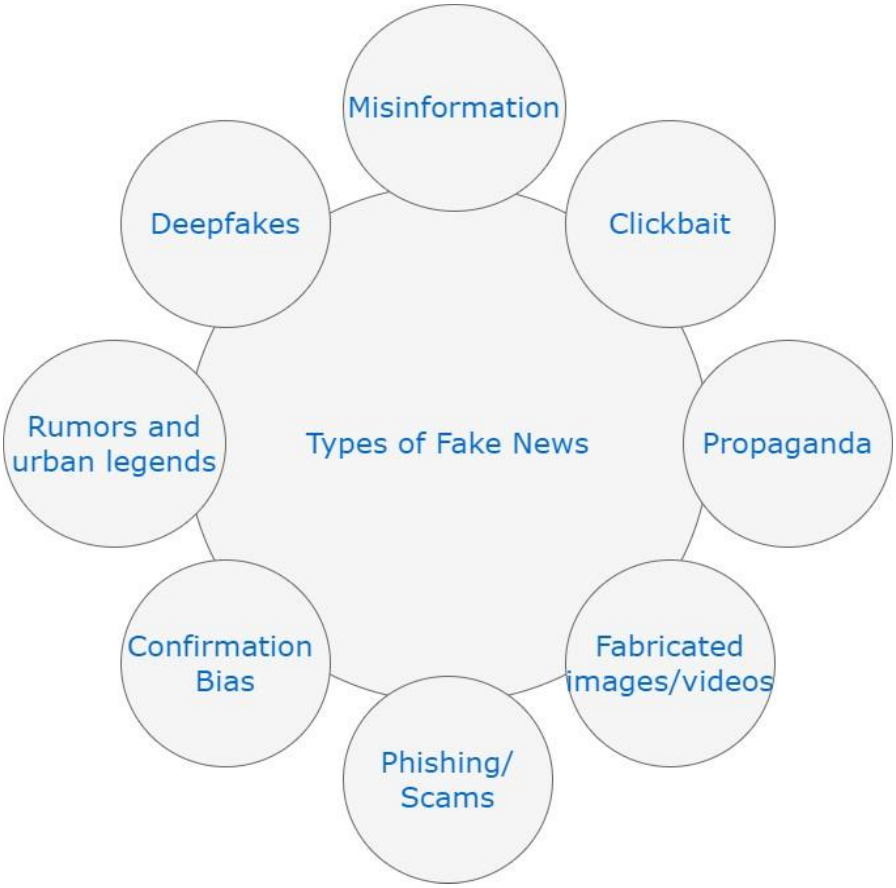

Fig. 2 Types of fake news

4.1. Health: Investigate the spread of fake health information, including false medical advice, miracle cures, and health-related hoaxes, which can have some consequences (Moreno-Castro et al., 2021).

4.2. Religion: Analyze how ‘fake news’ can affect religious beliefs and practices, exploring instances of religious hoaxes, false prophecies, and religiously motivated misinformation. (Therriault et al., 2022)

4.3. Crime: Explore the impact of fake news on criminal investigations and public perception, including cases of false accusations, crime-related conspiracy theories, and sensationalized crime stories.

4.4. Jobs: Discuss how misinformation can affect employment opportunities, job markets, and career decisions, including fake job advertisements, employment scams, and fraudulent career advice. (Dutta and Bandyopadhyay, 2020)

4.5. Politics: Examine the role of fake news in influencing political discourse, elections, and public policy. Explore cases of political misinformation, propaganda, and the spread of false political narratives. (Therriault et al., 2022)

5. Types of Different Fake News Existence

The 'fake or false' news is a type of false/misinformation can manifest in various forms and it's crucial to differentiate between different types of false information. They are categorised as shown in Fig. 2.

5.1 Misinformation: Inaccurate information spread without the intent to deceive. It often arises from genuine mistakes or misunderstandings. (Olan et al., 2022).

5.2 Clickbait: Sensationalized or misleading headlines designed to attract clicks and drive web traffic. The content may be exaggerated or distorted to generate interest. (Martínez-Sala et al., 2019)

5.3 Propaganda: Information spread by governments, organizations, or individuals to influence public opinion, often using biased or misleading narratives. (Aïmeur et al., 2023).

5.4 Deepfakes: AI-generated content, such as videos or audio recordings, that convincingly mimic real individuals, often used to create fabricated speeches or statements. (Westerlund, 2019).

5.5 Fabricated Images or Videos: Manipulated or entirely fabricated images or videos intended to deceive. Examples include photo-shopped images or digitally altered footage. (Nightingale et al., 2017).

5.6 Phishing and Scams: False information or deceptive tactics used to trap individuals into exposing personal information, passwords, or financial details. (Alkhalil et al., 2021).

5.7 Confirmation Bias: Presenting the 'information' in a way that reinforces 'pre-existing beliefs' and 'biases', without necessarily being false (Peters, 2022).
Rumours and Urban Legends: Unverified stories that circulate and are often based on fear, superstition, or exaggeration (DiFonzo and Bordia, 2007).

6. Methodology

The process of developing a 'Machine Learning' (ML) methodology for the detection of 'fake or false' news and mitigation involves the structured, step-by-step approach:

6.1. Existing Methodology for the fake news detection:

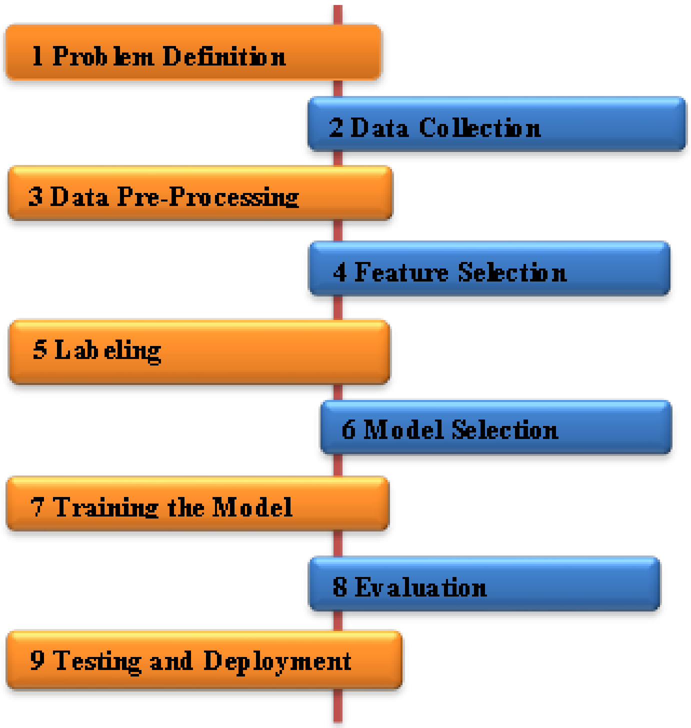

Fig. 3 A methodology using machine learning for detection of fake news

6.1.1 Problem Definition:

The problem of identification between real and fake news articles requires a clear definition and understanding of the specific context and goals of the fake news identification system (Fazil and Abulaish, 2018).

6.1.2 Data Collection:

It involves tasks like collecting a diverse datasets of news articles, which includes both fake and real news, to ensure a balanced and representative image of the desired news types (Sharma and Chaurasia, 2011).

6.1.3 Data Pre-processing:

Pre-process of data (text) and by removing 'stop words', 'special characters',' punctuation' and 'tokenize' it into words or sub-words, and convert it into numerical representations like TF-IDF vectors or word embedding.

6.1.4 Feature Selection:

Feature Selection involves extracting key features from text data, such as linguistics, textual patterns, and metadata, to differentiate between real and fake news (Sharma, 2021).

6.1.5 Labelling:

The task involves labelling the dataset to identify which articles are genuine and which are fake.

6.1.6 Model Selection:

Select the appropriate machine learning model or technique for classification, like as a CNN or RNN from deep learning, or the gradient boosting classifier, the 'decision tree classifier', 'logistic regression' or the 'random forest' from machine learning (Chaurasia, 2020). Here four models used to check for the fake news. They are as follows:-

Logistic Regression (LR):

A statistical model i.e. commonly used for 'binary classification' problems, like as in fake news detection. In logistic regression, the output is transformed using the logistic function (sigmoid) to ensure that the predicted values between 0 and 1 representing probabilities (Jee et al., 2022). A 'logistic regression' model for detection of fake news is being defined. There are set of features x_1, x_2,……, x_n that are utilized for predicting the part of news is either false or not. The logistic regression model is written as follows:

$$P(Fake\ News) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots \ldots\ldots\ldots + \beta_n X_n)}}$$

Here:

P (Fake News) is the probability of news being 'fake'.

e is the base of the natural logarithm.

β_0, β_1, …………, β_n are the coefficients of the model that are learned during the training process.

The function f(z) is defined as:

$$f(z) = \frac{1}{1+e^{-z}}$$

$$z=\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_n X_n$$

(Jee et al., 2022) During training, the logistic regression model is fitted to the training data by adjusting the coefficients β_0, β_1, … β_n to reduce the difference between the actual binary labels(fake or not fake) and predicted probabilities of the training examples.

After the model is trained, it can be used for predicting the probability of the new part of news being 'fake' by plugging in the values of its features into the trained model. If the predicted probability is greater than a certain threshold (commonly 0.5), news predicted is classified as false; otherwise, classified as it's not fake.

Decision Tree (DT) Classifier:

'Decision trees' are tree like structures that represents a series of decisions and their corresponding outcomes. They are easier and can handle both 'categorical' and 'numerical' data, making them suitable for fake news detection tasks.

Mechanism of Decision Tree Classifier:

Data Pre-processing: The first step; involves preparing the data for the 'decision tree' algorithm. This includes cleaning, removing irrelevant information, and converting text into numerical representations.

Feature Extraction: Applicable features are being extracted from all the 'pre-processed data'. Those features could include linguistic characteristics (e.g., word usage, sentiment), stylistic features (e.g., sentence structure, capitalization), and source credibility factors (Pérez-Rosas et al., 2018).

Construction of the Decision Tree: It is mainly constructed by recursively splitting the data into the subsets which are based mostly on the informative features. Each split creates a new node in the tree, and the branches represent the decision rules.

Training and Evaluation: The decision tree is trained on set of labelled dataset of the fake and real news posts or articles. The "trained models" are then evaluated on the separate set of the test dataset, to assess its performance.

Applications of Decision Tree Classifier in Fake News Detection:

Social Media Monitoring: Decision trees can be used to analyse posts of social media and detecting potential fake news articles in real-time.

News Feed Curation: Decision trees can be integrated into news aggregation platforms to filter out fake news and provide users with more reliable information.

Fact-Checking Tools: Decision trees can be used to develop fact-checking tools that assist journalists and researchers in verifying the authenticity of news articles.

Random Forest (RF) Classifier:

A 'random forest classifier' for detection of fake news contains the 'ensemble learning' method which is a combination of 'multiple decision trees' to the performance and generalization. A representation of the RF (Random Forest) classifier for detection of 'fake news':

Ensemble of Decision Trees: A Random Forest consists of 'N' number of 'decision trees' where each 'decision tree' is trained independently on a selection of the training sets of data with replacement. Each 'decision tree' is trained using a random selected set of features at the each split, adding an element of randomness.

Voting Mechanism: During classification, each decision tree in the 'random forest' independently classifies the input news article.

The result of final prediction or forecast is obtained by the most number of 'votes' among the trees.

Training: For each 'decision tree' in the forest: -A randomly picked subset of the training data is done with 'replacement'.

At each split, a randomly picked subset of features is considered. The decision tree is grown until a stopping criteria is achieved (example: maximum depth or minimum samples in a leaf node).

Final Prediction: For the given news article, in the random forest every tree's prediction provides a classification.

The outcome is the 'class label' that receives the most of votes in all the trees.

Mathematically; the prediction of a 'Random Forest' RF for a new instance with features $X_1, X_2, X_3, \ldots \ldots \ldots X_n$

Can be expressed as:

RF Prediction =

Majority Vote ($Tree_1$ $(X_1,X_2,\ldots\ldots.X_n)$,$Tree_2$$(X_1, X_2,\ldots..X_n)$,…..,$Tree_N$$(X_1,X_2,\ldots..X_n)$)

Where Tree shows the predictions of i-th decision tree in random forest. The Majority Vote function selects the class label that occurs most frequently among the 'individual tree' predictions.

Gradient Boosting (GB) Classifier:

It's another ensemble learning technique used for tasks in the classification, including fake news detection. Specifically, here it describes the concept of Gradient Boosting using decision trees, as it is a common implementation. Here's a representation of fake news using gradient boosting classifier:

It is a method that builds a set of decision trees sequentially. The ensemble's forecast is the weighted average of the projections from each individual tree.

For a classification problem:

Prediction = Class Label with the most votes

For a regression problem:

Prediction = $\sum_{i=1}^{n} a_i \, . \, f_i(x)$

Where:

"a_i": is the weight assigned to the ith tree.

"$f_i\ (x)$": is the prediction of the ith tree.

6.1.7 Training the Model

Separate the dataset in validation, training and testing sets, train the training data with the model, and fine-tune hyper-parameters and architecture for improved performance (Zhang et al., 2020).

6.1.8 Evaluation

Using metrics like F1-score, precision, accuracy, recall and others, the outcome of models on the validation set will be assessed and further fine-tuning may be necessary. They are used as:-

TP (True Positive): forecasted as fake news and actually also a fake news sets.

TN (True Negative): forecasted as real news and actually also a real news sets.

FN (False Negative): forecasted as real news and actually as fake news sets.

FP (False Positive): forecasted as fake news and actually as real news sets.

		Actual	
		Positive	Negative
Predicted	Positive	'True Positive'	'False Positive'
	Negative	'False Negative'	'True Negative'

Fig. 4 Confusion Matrix Table

Accuracy:

Accuracy proportions the correctly classified values, indicating the frequency of the classifier's accuracy by dividing the sum of real or true values with the total values.

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

Precision: Precision is the model's accuracy in correctly classifying positive values, calculated by dividing all the predicted values by the actual values:

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

Recall:

The model's ability to predict positive value outcome is calculated by dividing the true positive outcome values by the total number of actual positive outcomes to determine the frequency of correct predictions.

$$\text{Recall } = \frac{TP}{FN + TP}$$

F1-Score:

F1-Score:

The harmonic mean of the precision and the recall is a useful tool when considering both of the precision and recall.

$$\text{F1 Score } = \frac{2 * Precision * Recall}{Precision + Recall}$$

Support:

It is a measure of the frequency of the product or items in the dataset, calculated by dividing the number of transactions containing an item or product set by the total number of transactions.

$$\text{Supp(x)} = \frac{Freq(x)}{T}$$

6.1.9 Testing and Deployment:

The final model being tested on the test dataset to ensure its generalizability to new data, and then deployed to a production environment or integrated into an application for fake news detection in real-time.

7. Experimental Analysis

The following were the experimental hardware platforms: An Intel Core i3 serves as the CPU. There were 32 GB of RAM. The following were the platforms for experimental software: Windows 11 × 64-bit was the operating system (Ma et al., 2018). The Seaborn sklearn.model_selection, sklearn.mertces machine learning framework was used. Python 3.7 was the programming language used. The software was created using Jupyter 7.0 and Anaconda 3 5.2.0.The models that are employed include the random forest, the 'logistic regression', the 'gradient boosting' and the 'decision tree' (Rubin et al., 2016).

The sample experimental data were accessed from the Kaggle datasets. Sample dataset for both true fake news has been taken from kaggle (Wang et al., 2018). Further the true fake news dataset has been combined into a single dataset and then processes for the training and testing purpose.

```
data_true.head()
```

	title	text	subject	date
0	As U.S. budget fight looms, Republicans flip t...	WASHINGTON (Reuters) - The head of a conservat...	politicsNews	December 31, 2017
1	U.S. military to accept transgender recruits o...	WASHINGTON (Reuters) - Transgender people will...	politicsNews	December 29, 2017
2	Senior U.S. Republican senator: 'Let Mr. Muell...	WASHINGTON (Reuters) - The special counsel inv...	politicsNews	December 31, 2017
3	FBI Russia probe helped by Australian diplomat...	WASHINGTON (Reuters) - Trump campaign adviser ...	politicsNews	December 30, 2017
4	Trump wants Postal Service to charge 'much mor...	SEATTLE/WASHINGTON (Reuters) - President Donal...	politicsNews	December 29, 2017

```
data_fake.head()
```

	title	text	subject	date
0	Donald Trump Sends Out Embarrassing New Year'...	Donald Trump just couldn t wish all Americans ...	News	December 31, 2017
1	Drunk Bragging Trump Staffer Started Russian ...	House Intelligence Committee Chairman Devin Nu...	News	December 31, 2017
2	Sheriff David Clarke Becomes An Internet Joke...	On Friday, it was revealed that former Milwauk...	News	December 30, 2017
3	Trump Is So Obsessed He Even Has Obama's Name...	On Christmas day, Donald Trump announced that ...	News	December 29, 2017
4	Pope Francis Just Called Out Donald Trump Dur...	Pope Francis used his annual Christmas Day mes...	News	December 25, 2017

Fig. 5 True and fake news dataset.

Another separate dataset has been accessed from kaggle where the true and fake news are included in a single data set. This information is further use for the training and testing purpose. Logistic regression is used to prepare a model for checking whether news is true or fake

```
data_merge = pd.concat([data_fake,data_true], axis = 0)
data_merge.head()
```

	title	text	subject	date	class
0	Donald Trump Sends Out Embarrassing New Year'...	Donald Trump just couldn t wish all Americans ...	News	December 31, 2017	0
1	Drunk Bragging Trump Staffer Started Russian ...	House Intelligence Committee Chairman Devin Nu...	News	December 31, 2017	0
2	Sheriff David Clarke Becomes An Internet Joke...	On Friday, it was revealed that former Milwauk...	News	December 30, 2017	0
3	Trump Is So Obsessed He Even Has Obama's Name...	On Christmas day, Donald Trump announced that ...	News	December 29, 2017	0
4	Pope Francis Just Called Out Donald Trump Dur...	Pope Francis used his annual Christmas Day mes...	News	December 25, 2017	0

```
data_merge.tail()
```

	title	text	subject	date	class
21402	Exclusive: Trump's Afghan decision may increas...	ON BOARD A U.S. MILITARY AIRCRAFT (Reuters) - ...	worldnews	August 22, 2017	1
21403	U.S. puts more pressure on Pakistan to help wi...	WASHINGTON (Reuters) - The United States sugge...	worldnews	August 21, 2017	1
21404	Exclusive: U.S. to withhold up to $290 million...	WASHINGTON (Reuters) - The United States has d...	worldnews	August 22, 2017	1
21405	Trump talks tough on Pakistan's 'terrorist' ha...	ISLAMABAD (Reuters) - Outlining a new strategy...	worldnews	August 22, 2017	1
21406	U.S., North Korea clash at U.N. forum over nuc...	GENEVA (Reuters) - North Korea and the United ...	worldnews	August 22, 2017	1

Fig. 6 Merged data of true and fake news dataset.

Another separate dataset has been accessed from kaggle where the true and fake news are included in a single data set. This information is further use for the training and testing purpose. Logistic regression is used to prepare a model for checking whether news is true or fake.

A few models of the "Machine Learning" (ML) for classifying the 'articles as real or fake'. These models included the 'Logistic Regression', the 'Gradient Boosting Classifier', the 'Decision Tree Classifier' and the' Random Forest Classifier'. The models trained on a 'subset' of the dataset and then evaluated on a test sets.

The evaluated results are shown below in the 'Table 2'

Table-2: Performance tabular view of machine learning models for detecting 'fake news'.

Model	Accuracy	Precision	Recall	Support	F1 Score
GB	0.995276293	0.99	1	5318	1
DT	0.995900178	1	1	5318	1
RF	0.987344029	0.99	0.99	5318	1
LR	0.977584029	1	1	5318	1

As displayed details on Table-2, the Decision Tree Classifier model achieved the higher value in overall performance, having the accuracy of 0.995900178 as well as an F1-score of 1. GB model also performed well, having the accuracy of 0.995276293 and the F1 -score as 1.

For testing purpose different sample text from the dataset has been taken to check if the inputted news is either true or fake.

'T' signifies that the sample news tested by the four models was predicted to be TRUE news.

'F' signifies that the sample news tested by the four models was predicted to be FALSE news

After checking for the eight sample news taken from the dataset, it displayed the following output:

Table-3: Output predicted by models.

News Samples	Actual True News	Models LR	DT	GB	RF
'News Sample-1'	Yes	'T'	'T'	'T'	'T'
'News Sample-2'	No	'F'	'F'	'F'	'F'
'News Sample-3'	No	'F'	'F'	'F'	'F'
'News Sample-4'	Yes	'T'	'T'	'T'	'T'
'News Sample-5'	No	'F'	'F'	'F'	'F'
'News Sample-6'	Yes	'T'	'T'	'T'	'T'
'News Sample-7'	Yes	'T'	'T'	'T'	'T'
'News Sample-8'	No	'F'	'F'	'F'	'F'

Table 4 Improvements in Existing Models

Model	Pros	Cons	Handling Cons
LR	Simple to interpret, Relatively fast training	Limited feature handling, Prone to over-fitting	Feature engineering, Regularization techniques
DT	Easy to interpret, Handles various data types	Prone to over-fitting, Sensitive to small changes in data	Pruning, Ensemble methods
GB	Powerful and flexible, Less prone to over-fitting	Black box model, Computationally expensive	Feature importance analysis, Use simpler models for interpretability
RF	Highly accurate, Less prone to over-fitting, Handles various data types	Black box model, Computationally expensive	Feature importance analysis, Use simpler models for interpretability

8. Conclusion

This paper summarizes the essential findings and trends in fake news(text) disturbing the social as well as environmental aspects, over the past years emphasizing the need for 'innovative solutions' to tackle the evolving threat of fake news, acknowledging its impact and challenges across different domains.

In this paper it has been discussed about the four different models i.e. the 'decision tree classifier', the 'random forest classifier', the' logistic regression' and the 'gradient boosting classifier' which is used to check whether the news is real or fake and respond on the basis of four model evaluation for checking of fake news is done.

Mitigation of 'fake news' is a blooming concern in today's information-driven world, as 'fake news' can influence public opinion, disrupt social harmony, and undermine media credibility. To combat this issue, various strategies and mitigation efforts have been implemented by individuals, organizations, and technology. These include fact-checking, AI algorithms, media literacy, and legal measures. The on-going battle to uphold the integrity of the information landscape is crucial for fostering a more informed and discerning society.

Mitigation techniques are also important and a serious research scope, which have proved to be a boon to the cyber community to combat fake news and deal with it in the right way. Models to mitigate spread of fake news, which might include techniques to crumble the entire communication of fake news, beginning from source to platforms, from reporting to removal of fake news real quick.

Acknowledgement

The editors and reviewers are highly appreciated by the authors for their insightful and necessary comments, which significantly enhanced the overall manuscript as a whole.

Author Contribution

Ritika Sharma and Gourab Das has drafted the manuscript as well as done the machine learning model training and testing section, Ashish Kumar Rauniyar and Gauri Kumari has contributed in completing the literature review section. The whole 'manuscript' has been edited and reviewed by all the authors according to the requirements and Dr. Pawan Kumar Chaurasia has supervised the works.

Conflict of Interest

Author's state that there is not any recognized conflict' of interest for them.

References

Agarwal, A., Mishra, S., Ahmad, S. 2023. Fake News Detection Using Machine Learning. Lecture Notes in Electrical Engineering, 1068 LNEE2, 51–59. https://doi.org/10.1007/978-981-99-5358-5_4

Ahmad, I., Yousaf, M., Yousaf, S., Ahmad, M.O. 2020. Fake News Detection Using Machine Learning Ensemble Methods. Complexity, 2020. https://doi.org/10.1155/2020/8885861

Ahmed, H., Traore, I., Saad, S. 2017. Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques. Lecture Notes in Computer Science Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, 10618 LNCSOctober, 127–138. https://doi.org/10.1007/978-3-319-69155-8_9

Aïmeur, E., Amri, S., Brassard, G. 2023. Fake news, disinformation and misinformation in social media: a review. In Social Network Analysis and Mining Vol. 13, Issue 1. Springer Vienna. https://doi.org/10.1007/s13278-023-01028-5

Ajao, O., Bhowmik, D., Zargari, S. 2018. Fake news identification on Twitter with hybrid CNN and RNN models. ACM International Conference Proceeding Series, June, 226–230. https://doi.org/10.1145/3217804.3217917

Alkhalil, Z., Hewage, C., Nawaf, L., Khan, I. 2021. Phishing Attacks: A Recent Comprehensive Study and a New Anatomy. Frontiers in Computer Science, 3March, 1–23. https://doi.org/10.3389/fcomp.2021.563060

Castillo, C., Mendoza, M., Poblete, B. 2011. Information credibility on Twitter. Proceedings of the 20th International Conference Companion on World Wide Web, WWW 2011, May 2014, 675–684. https://doi.org/10.1145/1963405.1963500

Chaurasia, P. K. 2020. Paradigms of Machine Learning and Data Analytics. 156–174. https://doi.org/10.4018/978-1-7998-2491-6.ch009

DiFonzo, N., Bordia, P. 2007. Rumor, gossip and urban legends. Diogenes, 541, 19–35. https://doi.org/10.1177/0392192107073433

Dutta, S., Bandyopadhyay, S.K. 2020. Fake job recruitment detection using machine learning approach. International Journal of Engineering Trends and Technology, 684, 48–53. https://doi.org/10.14445/22315381/IJETT-V68I4P209S

Fazil, M., Abulaish, M. 2018. A Hybrid Approach for Detecting Automated Spammers in Twitter. IEEE Transactions on Information Forensics and Security, 1311, 2707–2719. https://doi.org/10.1109/TIFS.2018.2825958

Gupta, S.K., Shrivastava, A., Upadhyay, S.P., Chaurasia*, P. K. 2021. A Machine Learning Approach for Heart Attack Prediction. International Journal of Engineering and Advanced Technology, 106, 124–134. https://doi.org/10.35940/ijeat.f3043.081062

Jee, K., Kaur, A., Joshi, A. 2022. Machine Learning in Healthcare Industry: An Analysis. International Conference on Cyber Resilience, ICCR 2022, June. https://doi.org/10.1109/ICCR56254.2022.9995941

Jwa, H., Oh, D., Park, K., Kang, J. M., Lim, H. 2019. exBAKE: Automatic fake news detection model based on Bidirectional Encoder Representations from Transformers BERT. Applied Sci (Switz.) 919, 1–9. https://doi.org/10.3390/app9194062

Kaliyar, R. K., Goswami, A., Narang, P. 2021. FakeBERT: Fake news detection in social media with a BERT-based deep learning approach. Multimed. Tools Appl. 808, 11765–11788. https://doi.org/10.1007/s11042-020-10183-2

Khanam, Z., Alwasel, B. N., Sirafi, H., Rashid, M. 2021. Fake News Detection Using Machine Learning Approaches. IOP Conference Series: Materials Science and Engineering, 10991, 012040. https://doi.org/10.1088/1757-899x/1099/1/012040

Liu, Y., Wu, Y.F.B. 2018. Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks. 32nd AAAI Conference on Artificial Intelligence, AAAI 2018, December, 354–361. https://doi.org/10.1609/aaai.v32i1.11268

Ma, J., Gao, W., Wong, K. F. 2018. Detect Rumor and Stance Jointly by Neural Multi-task Learning. The Web Conference 2018 - Companion of the World Wide Web Conference, WWW 2018, April, 585–593. https://doi.org/10.1145/3184558.3188729

Martínez-Sala, A.-M., Monserrat-Gauchi, J., Segarra-Saavedra, J. 2019. El influencer 2.0 turístico: de turista anónimo a líder de opinión. Latina, Revista de Comunicación, 74, 734–747. https://doi.org/10.4185/RLCS-201

Moreno-Castro, C., Vengut-Climent, E., Cano-Orón, L., Mendoza-Poudereux, I. 2021. Exploratory study of the hoaxes spread via WhatsApp in Spain to prevent and/or cure COVID-19. Gaceta Sanitaria, 356, 534–541. https://doi.org/10.1016/j.gaceta.2020.07.008

Nightingale, S. J., Wade, K. A., Watson, D. G. 2017. Can people identify original and manipulated photos of real-world scenes? Cognitive Research: Principles and Implications, 21. https://doi.org/10.1186/s41235-017-0067-2

Olan, F., Jayawickrama, U., Arakpogun, E. O., Suklan, J., Liu, S. 2022. Fake news on Social Media: the Impact on Society. Information Systems Frontiers, January. https://doi.org/10.1007/s10796-022-10242-z

Palani, B., Elango, S., Vignesh Viswanathan, K. 2022. CB-Fake: A multimodal deep learning framework for automatic fake news detection using capsule neural network and BERT. In Multimedia Tools and Applications Vol. 81, Issue 4. Springer US. https://doi.org/10.1007/s11042-021-11782-3

Pérez-Rosas, V., Kleinberg, B., Lefevre, A., Mihalcea, R. 2018. Automatic detection of fake news. COLING 2018 - 27th International Conference on Computational Linguistics, Proceedings, September, 3391–3401.

Peters, U. 2022. What Is the Function of Confirmation Bias? Erkenntnis, 873, 1351–1376. https://doi.org/10.1007/s10670-020-00252-1

Rubin, V., Conroy, N., Chen, Y., Cornwell, S. 2016. Fake News or Truth? Using Satirical Cues to Detect Potentially Misleading News. 7–17. https://doi.org/10.18653/v1/w16-0802

Sharma, P. 2021. Review on Heart Disease Prediction using Machine Learning. Int. J. Res. Appl Sci. Eng. Technol. 94, 1053–1060. https://doi.org/10.22214/ijraset.2021.33848

Sharma, P., and Chaurasia, P. K. 2011. A Critical Review on Financial Fraud Detection Techniques. IRACST- Int. J. Comput. Netw. Wirel. Commun. 8(3), 91-98.

Shu, K., Sliva, A., Wang, S., Tang, J., Liu, H. 2017. Fake News Detection on Social Media. ACM SIGKDD Explorations Newsletter 191, 22–36. https://doi.org/10.1145/3137597.3137600

Singh, J. P., Kumar, A., Rana, N. P., Dwivedi, Y. K. 2022. Attention-Based LSTM Network for Rumor Veracity Estimation of Tweets. Inf. Syst. Front. 242, 459–474. https://doi.org/10.1007/s10796-020-10040-5

Therriault, D. J., Zaboski, B. A., Jankovsky A. 2022. What Can Fake News, Politics, and Religion Tell Us About Pseudoscience? J. Educ. 91

Verma, P. K., Agrawal, P., Amorim, I., Prodan, R. 2021. WELFake: Word Embedding over Linguistic Features for Fake News Detection. IEEE Trans. Comput. Soc. Syst. 84, 881–893. https://doi.org/10.1109/TCSS.2021.3068519

Verma, P. K., Agrawal, P., Madaan, V., Prodan, R. 2023. MCred: multi-modal message credibility for fake news detection using BERT and CNN. J. Ambient Intell. Humaniz. Comput. 148, 10617–10629. https://doi.org/10.1007/s12652-022-04338-2

Wang, Y., Ma, F., Jin, Z., Yuan, Y., Xun, G., Jha, K., Su, L., Gao, J. 2018. EANN: Event adversarial neural networks for multi-modal fake news detection. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, September, 849–857. https://doi.org/10.1145/3219819.3219903

Westerlund, M. 2019. The emergence of deepfake technology: A review. Technology Innovation Management Review, 911, 39–52. https://doi.org/10.22215/TIMREVIEW/1282

Zhang, T., Wang, D., Chen, H., Zeng, Z., Guo, W., Miao, C., Cui, L. 2020. BDANN: BERT-Based Domain Adaptation

Neural Network for Multi-Modal Fake News Detection. Proceedings of the International Joint Conference on Neural Networks, July 2020. https://doi.org/10.1109/IJCNN48605.2020.9206973

Cite this article:
Sharma, R., Kumari, G., Rauniyar, A.K., Das, G., Mishra, R.K., Chaurasia, P.K., 2024. Unmasking Social Fake News: Machine Learning Approach. J. Appl. Sci. Innov. Technol. 3 (1), 20-32.