ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VIII, August 2025

Advanced Techniques for Fake News Detection on Twitter Using NLP and AI: A Comprehensive Review

Mrs. T. Poornima[1], Dr. S. Kumaravel[2], Mr. A. Thottarayaswamy[3*]

[1]*Research Scholar (Part time) Department of Computer Science, Sri Ramakrishna Mission Vidyalaya College of Arts and Science, (Affiliated to Bharatiyar University) Periyanaickenpalayam, Coimbatore – 641020.*

[2]*Head of the department, Department of Computer Science, Sri Ramakirshna Mission College of Arts and Science,Vidyalaya(Affiliated to Bharatiyar University) Periyanaickenpalayam, Coimbatore – 641020.*

[3]*Department of Computer Science, Vivekam Matriculation Higher Secondary School, Veerapandi Pirivu, Coimbatore - 641019*

DOI: https://doi.org/10.51583/IJLTEMAS.2025.1408000001

Abstract: This review paper explores the critical aspects of pre-processing, feature extraction, feature selection, and classification in fake news detection on Twitter. Pre-processing involves cleaning text data by removing unwanted elements and normalizing the text. Feature extraction transforms raw text into analysable formats, identifying significant aspects to distinguish between fake and real news. Feature selection focuses on identifying relevant features to reduce dimensionality and noise. Various classification techniques are evaluated for their effectiveness, demonstrating the potential of machine learning in achieving high accuracy. The paper also examines Twitter's role in rapid information dissemination and the challenges posed by fake news. It reviews content-based methods like text analysis, user-based methods such as behaviour and network analysis, hybrid methods integrating multiple features, and the use of external knowledge through fact-checking and knowledge graphs. Addressing the challenges of high information volume, rapid spread, and evasion techniques, the paper concludes that effective fake news detection on Twitter requires advanced analytical techniques and real-time monitoring to manage and mitigate misinformation.

Keywords: Fake News Detection, Twitter, Pre-processing, Feature Extraction, Machine Learning.

I. Introduction

In the digital age, social media platforms have become primary sources for news consumption and information dissemination. Among these platforms, Twitter stands out due to its real-time nature and widespread user engagement. However, the rapid spread of information on Twitter also makes it a fertile ground for the propagation of fake news, which can have significant societal impacts. Consequently, detecting and mitigating fake news on Twitter is a critical area of research. This review paper delves into the essential components of fake news detection, focusing on pre-processing, feature extraction, feature selection, and classification.

Pre-processing is the foundational step that involves cleaning and normalizing text data, which is crucial for enhancing the performance of machine learning models. This stage includes removing unwanted characters, punctuation, URLs, HTML tags, and converting text to lowercase to ensure uniformity. Feature extraction follows, transforming raw text into formats that can be effectively analysed, and identifying key aspects that help distinguish between fake and real news. Feature selection then narrows down the relevant features, reducing dimensionality and minimizing noise, thereby improving model performance (E. Cueva et al).

The paper evaluates various classification techniques, demonstrating the potential of machine learning to achieve high accuracy in detecting fake news. Beyond technical methodologies, the paper also examines the unique challenges posed by Twitter's platform, such as the high volume and rapid spread of information, as well as the sophisticated evasion techniques used by fake news creators. It reviews different approaches to fake news detection, including content-based methods like text analysis, user-based methods such as behaviour and network analysis, hybrid methods that combine multiple features, and the use of external knowledge through fact-checking and knowledge graphs.

Addressing these challenges, the paper concludes that a combination of advanced analytical techniques and real-time monitoring is essential for effective fake news detection on Twitter. This comprehensive approach not only enhances the accuracy of detection algorithms but also helps in managing and mitigating the spread of misinformation, ultimately contributing to a more reliable information ecosystem on social media platforms.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VIII, August 2025

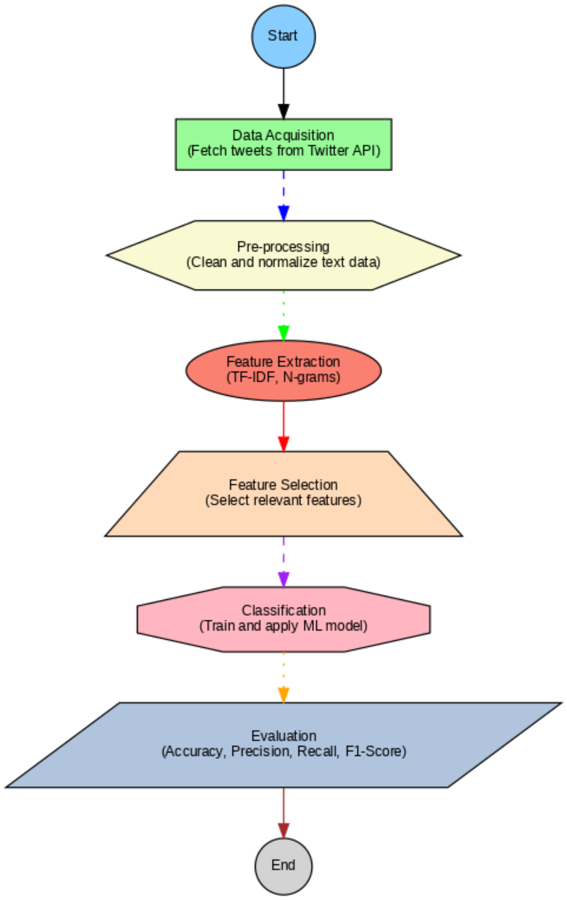

Fig.1. Flow Diagram for Classifying Fake and Real News on Twitter

The flowchart outlines a process for classifying Twitter data as fake or real news. It starts with acquiring tweets, then cleans and prepares the data, extracts and selects relevant features, applies classification models, and evaluates their performance.

Social media

Social media includes online platforms that enable users to create, share, and exchange content. Popular examples are Facebook, Twitter, Instagram, LinkedIn, and YouTube (V. Kumar. et al., 2021).

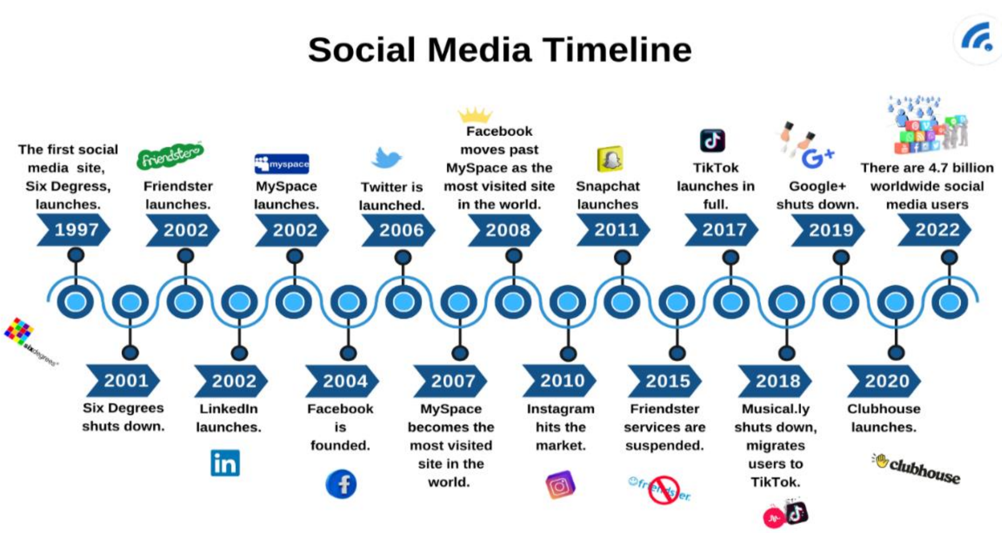

Figure.1 Popular social media platforms with Timeline

Here are some key characteristics and aspects of social media:

- ***User-generated content:*** Driven by text, images, and videos created by users.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VIII, August 2025

- ***Interaction and engagement:*** Allows comments, likes, shares, and direct messages.
- ***Personalization:*** Uses algorithms to tailor content to users' preferences.
- ***Global reach:*** Connects people worldwide for information and idea sharing.
- ***Concerns:*** Issues with privacy, misinformation, and online harassment.

Overall, social media has become an increasingly important part of modern communication and has had a significant impact on various aspects of society, from politics and activism to marketing and entertainment (Akhtar, P. et al., 2023). Some samples of social media platforms are listed below:

Table.1. Social media platforms

Platform	Description	Founded	Active Users (as of 2022)	Primary Function
Facebook	A social networking site where users can connect with others and share content	2004	2.8 billion	Social networking
Twitter	A microblogging site where users can share short messages	2006	330 million	Social networking / Microblogging
Instagram	A photo and video sharing platform	2010	1.2 billion	Photo and video sharing
LinkedIn	A professional networking site where users can connect with other professionals	2003	740 million	Professional networking
YouTube	A video sharing platform	2005	2 billion	Video sharing / Entertainment / Education / Marketing

Fake news on social media

Fake news on social media involves intentionally misleading or fabricated information aimed at influencing public opinion. Often featuring sensational headlines, it spreads rapidly on platforms like Facebook, Twitter, and WhatsApp, leading to misinformation, reputation damage, and social unrest (Youness Madani et al, 2021). It impacts various sectors:

- **Business**: False claims can harm reputations and affect stock market performance.
- **Education**: Misleading information can lead to misunderstandings and poor academic decisions.
- **Healthcare**: Incorrect health information can result in harmful self-diagnosis or treatment choices.
- **Activism**: Fake news can distort social issues and influence public opinion negatively.
- **Elections**: Misinformation can sway voting behaviour and impact election outcomes.
- **Banks**: False rumours about financial stability can damage a bank's reputation.

1. Opinion analysis

Opinion analysis involves analyzing large volumes of user-generated content to determine the sentiment or polarity of feedback about products or services. Computer processing formalizes this knowledge, making it useful for further applications. Figure 1.2 depicts the object model of opinion analysis (Akhtar, P et al., 2023).

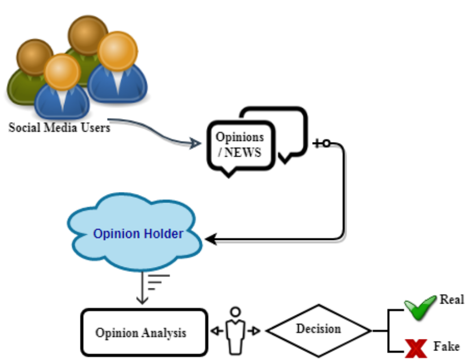

Figure.2 Opinion Analysis Model

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VIII, August 2025

Opinion Mining (OM) is a Natural Language Processing (NLP) technique that extracts and analyses subjective information from text, like social media posts or customer reviews, to determine sentiment positive, negative, or neutral. It uses methods such as part-of-speech tagging and machine learning to handle large text datasets.

OM is widely used in marketing, customer service, and public opinion research to gain insights and improve strategies. However, fake news can skew results and undermine system accuracy, highlighting the need for reliable data sources and ongoing verification.

- Fake news can generate a false or misleading perception of public opinion, which can skew the results of sentiment analysis and lead to inaccurate conclusions.
- Fake news can also impact the credibility of the sentiment analysis system itself.
- To mitigate the impact of fake news on sentiment analysis, it is important to ensure that the data sources used for sentiment analysis are reliable and trustworthy.
- This can be achieved by carefully selecting the sources of data and using techniques such as fact-checking to verify the accuracy of the data before it is used for sentiment analysis.
- Additionally, it is important to continually monitor for the presence of fake news and take appropriate action to mitigate its impact on sentiment analysis results (Akhtar, P et al., 2023).

Twitter:

Twitter is a popular social media platform that allows users to post and interact with short messages known as tweets. Founded in 2006, it enables real-time communication and sharing of news, opinions, and media among users globally. With features such as hashtags, mentions, and rewets, Twitter facilitates the rapid spread of information and engagement on various topics, including news, politics, and entertainment. It serves as a key tool for personal expression, networking, and information dissemination.

Fake News

Fake news detection on Twitter involves identifying and filtering out false or misleading information shared on the platform. This process uses various techniques, including machine learning algorithms, natural language processing, and network analysis, to assess the credibility of tweets and the authenticity of their sources. The goal is to prevent the spread of misinformation by analysing content patterns, user behaviour, and contextual factors to distinguish between accurate and deceptive news.

Approaches to Fake News Detection on Twitter

Fake news detection on Twitter focuses on identifying and mitigating misleading or false information spread through tweets. Given the platform's rapid information dissemination and high user engagement, addressing fake news is crucial to preventing misinformation.

i. **Content-Based Methods**

- **Text Analysis:** Utilizes NLP to detect linguistic patterns and anomalies in tweet content indicative of fake news.
- **Feature Extraction:** Employs techniques such as term frequency-inverse document frequency (TF-IDF) and sentiment analysis to evaluate text credibility.

ii. **User-Based Methods**

- **User Behaviour:** Analyses user profiles, historical behaviour, and interaction patterns to identify potentially unreliable sources.
- **Social Network Analysis:** Examines interaction networks to detect misinformation spread patterns among accounts (Kaliyar et al, 2021).

iii. **Hybrid Methods**

- **Combining Features:** Integrates content, user behavior, and social context features using machine learning (ML) and deep learning (DL) models to improve detection accuracy.
- **Model Types:** Implements algorithms like Support Vector Machines (SVM), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks to classify tweets as real or fake.

iv. **External Knowledge**

- **Fact-Checking:** Cross-references tweets with credible news sources and databases to verify information accuracy.
- **Knowledge Graphs:** Uses structured data about entities and relationships to detect inconsistencies.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VIII, August 2025

v. Challenges

- **Volume and Velocity:** High information volume and rapid spread complicate the timely detection and management of fake news.
- **Evasion Techniques:** Fake news creators use sophisticated methods, such as altering language patterns and employing bots, to evade detection.

Effective fake news detection on Twitter requires a combination of advanced analytical techniques and real-time monitoring to manage and mitigate misinformation (Q. Abbas et al, 2022).

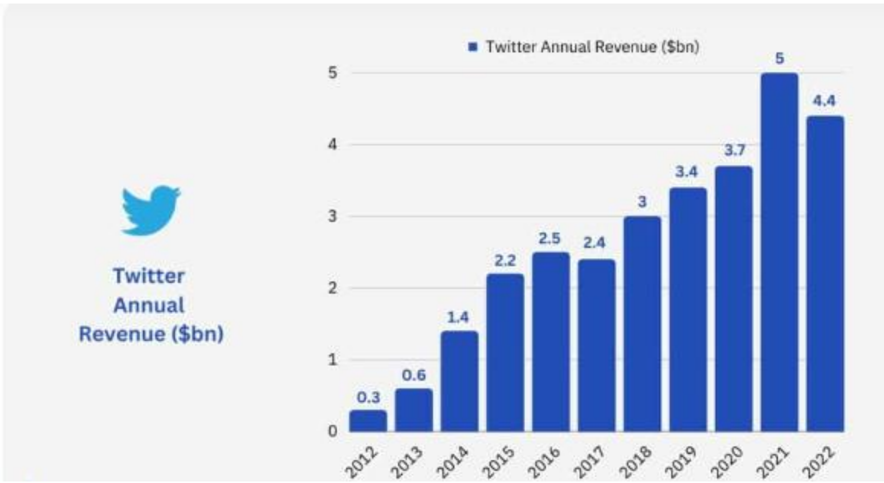

Figure.3. Twitter's Annual Revenue (2012-2022)

The image is a graph depicting Twitter's annual revenue from 2012 to 2022. The y-axis represents the revenue in billions of dollars (USD), while the x-axis denotes the years. The graph illustrates a generally upward trend in Twitter's revenue over the years, with some notable exceptions.

There was a slight dip in revenue in 2016, followed by a more significant decline in 2022. The peak revenue occurred in 2021, reaching $5 billion. By 2022, Twitter's revenue had decreased to $3.4 billion.

Pre-processing:

Pre-processing in fake news detection involves cleaning and preparing text data to enhance the performance of machine learning models. Key steps include removing unwanted characters, punctuation, URLs, and HTML tags; converting text to lowercase for uniformity; tokenizing text into individual words; eliminating common stop words; and applying stemming or lemmatization to reduce words to their root forms. Additionally, pre-processing addresses missing values, removes duplicates, and handles numerical and special characters. These steps ensure that the text data is clean, consistent, and ready for accurate and reliable analysis by fake news detection algorithms (Balshetwar et al, 2024).

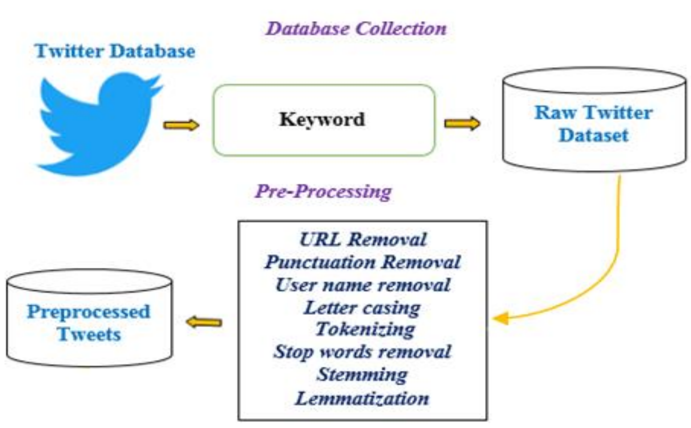

Figure.4. Twitter Data Collection and Pre-Processing Workflow

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VIII, August 2025

The above figure shows the process of gathering raw Twitter data using a specified keyword, then performing pre-processing steps to clean and ready the data for further analysis.

Table 1: Pre-processing technique used in fake news detection

Pre-processing Technique	Description	Example
Removing Punctuation	Eliminating punctuation marks to avoid their impact on text analysis.	"Hello, world!" → "Hello world"
Tokenization	Splitting text into individual words or tokens for easier analysis.	"Fake news spreads quickly." → ["Fake", "news", "spreads", "quickly"]
Lowercasing	Converting all text to lowercase to ensure uniformity.	"FAKE News" → "fake news"
Removing URLs	Stripping out web addresses that can distract from the main text content.	"Read more at https://example.com" → "Read more at"
Removing Stop Words	Removing common words that do not carry significant meaning.	"This is an example" → "example"
Stemming	Reducing words to their base or root form to standardize different forms of the same word.	"running" → "run"
Lemmatization	Reducing words to their base or dictionary form, similar to stemming but more sophisticated.	"better" → "good"

Cueva et al. (2020) investigated fake news detection on Twitter using various AI methods, focusing on LSTM, GRU, and NLP networks. They pre-processed data and trained these models to evaluate their accuracy in predicting fake news. Their comparative analysis revealed that while all three models performed well, the NLP model uniquely excelled at identifying satire as fake news, making it the most effective choice for fake news detection on Twitter.

Kumar et al. (2021) addressed the critical issue of fake news on social media platforms like Facebook and Twitter, proposing an AI and ML-based approach for binary classification to detect fake news. Their goal was to empower users to identify fake news and verify the legitimacy of publishing websites. This study was highlighted at the 2021 International Conference on Technological Advancements and Innovations (ICTAI). Akhtar et al. (2023) tackled the challenge of fake news and disinformation (FNaD) disrupting decision-making and supply chain disruptions (SCDs). They developed an FNaD detection model using AI and ML, supported by case studies from Indonesia, Malaysia, and Pakistan.

Madani et al. (2021) introduced a novel approach for detecting fake news on Twitter during the Covid-19 pandemic, focusing on Moroccan tweets. By integrating NLP, ML, and deep learning with Apache Spark, their method achieved an accuracy of 79%, highlighting the role of tweet sentiment in fake news detection. Kaliyar et al. (2021) proposed FakeBERT, a BERT-based deep learning approach for fake news detection on social media, combining CNN with varying kernel sizes and filters. This method achieved an accuracy of 98.90%, as detailed in their publication in Multimedia Tools and Applications.

Agarwal et al. (2019) used NLP and ML techniques for fake news detection, employing bag-of-words, n-grams, count vectorizer, and TF-IDF for preprocessing and training on different classifiers to identify the most effective model. Their evaluation focused on precision, recall, and F1 scores. Dahlan and Suyanto (2023) developed a Transformer-based fake news detection system using BERT and three augmentation methods: spell-checking-based, acronym-based, and typography-based.

Bhatia et al. (2023) proposed a technique for detecting fake news sources on Twitter using tweet content, published time, and social graph features to develop a deep neural network (DNN) predictive model. Their approach achieved an accuracy of 98.7% on the PolitiFact dataset. Chauhan et al. (2023) addressed fake news dissemination in India, proposing a machine learning algorithm for detection and comparing it with other methods developed to date. Heidari et al. (2021) utilized BERT and transfer learning to detect bot accounts related to Covid-19, introducing a model that integrated new features to enhance fake news detection. Chakraborty et al. (2022) combined feature engineering, network representation, and NLP techniques to distinguish between genuine and spam accounts on Twitter, offering a comprehensive strategy for identifying fake user accounts.

Gupta et al. (2021) employed social network analysis and graph classification techniques to study the spread of fake versus real news on Twitter during the Covid-19 pandemic. They created and analyzed graphs of fake and real news to predict news authenticity based on graph connectivity. Bhatia et al. (2022) developed a machine learning-based system for categorizing news as fake or real, implementing a comprehensive process including data collection, preprocessing, feature extraction, and classification with various ML algorithms.

Sultan et al. (2023) proposed a fake news detection model using Logistic Regression, Decision Tree, Random Forest, and Gradient Boosting Classifier. Their evaluation showed that the Decision Tree model achieved the highest accuracy of 99.60%. Abbas et al.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VIII, August 2025

(2022) combined CNN and RNN (LSTM) networks in a deep learning approach to mitigate the impact of fake news. They used word vectors for pre-processing and LSTM output for classification. Balshetwar et al. (2023) incorporated sentiment analysis as a key feature for fake news detection, achieving an accuracy of 99.8% by comparing classifiers like Naïve Bayes, Passive-Aggressive, and Deep Neural Networks (DNN).

Feature Extraction:

Feature extraction in fake news detection involves transforming raw text data from news articles or social media posts into a format that machine learning models can analyse effectively. It focuses on identifying and quantifying important aspects of the text to help distinguish between fake and real news.

Table 2: Feature extraction algorithms for fake news detection on Twitter

Feature Extraction Algorithm	Description	Examples
Term Frequency-Inverse Document Frequency (TF-IDF)	Measures the importance of words in a tweet relative to the entire corpus.	Words like "scandal" or "expose" in a tweet might have high TF-IDF scores in fake news.
Bag of Words (BoW)	Represents text by the frequency of each word without considering word order.	A tweet might be represented as a vector with counts of words like "news", "government", etc.
Word Embeddings (Word2Vec, GloVe, FastText)	Converts words into dense vectors based on their semantic meanings and relationships.	Words like "fraud" and "deception" are mapped to vectors close to each other in the embedding space.
N-grams	Sequences of 'n' contiguous words or characters.	"Breaking news", "government scandal" are 2-grams that capture context in a tweet.
Named Entity Recognition (NER)	Identifies and classifies entities like people, organizations, or locations.	Detecting "Donald Trump" or "United Nations" in tweets to assess credibility.
Sentiment Analysis	Analyses the emotional tone of the tweet (positive, negative, neutral).	Tweets with high negative sentiment might be more likely to be fake news.
Topic Modeling (Latent Dirichlet Allocation, LDA)	Identifies topics or themes present in the tweets.	Detecting topics like "elections" or "pandemic" in tweets and comparing them with reliable sources.
Textual Entailment	Measures if one sentence logically follows from another.	Verifying if the claims in a tweet are supported by facts in reputable sources.
Character N-grams	Sequences of 'n' characters, useful for detecting spelling variations and manipulations.	"scndal" vs. "scandal" can be captured with character n-grams.
Stylometric Features	Analyses writing style features such as punctuation, sentence length, and word choice.	Unusual writing patterns or excessive exclamation marks could indicate fake news.

This table provides a quick overview of the different feature extraction methods, their descriptions, and how they might be applied or identified in tweets.

N. F. Baarir and A. Djeffal propose a system for identifying fake news through machine learning techniques. They use term frequency-inverse document frequency (TF-IDF) with bag of words and n-grams for feature extraction and apply a Support Vector Machine (SVM) as the classifier. The study includes the development of a dataset consisting of both fake and true news for training purposes. The results validate the effectiveness of their system in distinguishing between authentic and false news, demonstrating the utility of machine learning in improving news accuracy.

J. Shaikh and R. Patil address the challenge of fake news detection by employing various classification techniques to enhance accuracy. Their work highlights the impact of fake news across diverse domains, including politics and education, and proposes a solution involving different classifiers such as Support Vector Machine (SVM), Naïve Bayes, and Passive Aggressive Classifier. The study utilizes Term Frequency-Inverted Document Frequency (TF-IDF) for feature extraction and achieves an impressive accuracy of 95.05% with the SVM classifier. This research underscores the potential of machine learning techniques in improving the reliability of news detection despite the challenges posed by limited resources.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VIII, August 2025

G. Rawat et.al. explore the escalating issue of fake news on social networks and the potential of machine learning (ML) techniques to address it. The chapter presents a comprehensive mapping analysis of various ML approaches for detecting fake news, highlighting the widespread concern over the impact of false propaganda. The authors review classic machine learning models and emphasize the importance of feature extraction and vectorization techniques. They recommend utilizing Python tools like the count vectorizer and TF-IDF vectorizer for effective text tokenization and feature extraction. By applying feature selection methods and analyzing Confusion Matrix results, the study aims to optimize accuracy in fake news classification, contributing valuable insights to the field of automated misinformation detection.

D. R. Collen et. al address limitations in previous fake news detection models related to feature extraction and text classification. The study introduces a 5-layer Convolutional Neural Network (5L-CNN) to analyze both titles and content of news articles, utilizing inbuilt tokenizers for word embedding. By comparing traditional machine learning algorithms—such as Decision Trees, Random Forest, Logistic Regression, and Naive Bayes with deep learning techniques including Recurrent Neural Networks (RNN), 5L-CNN, and Long Short-Term Memory (LSTM) networks, the authors demonstrate that the 5L-CNN model achieves superior accuracy of 99.99%. This research highlights the efficacy of deep learning methods over conventional approaches in enhancing fake news detection.

Feature Selection:

Feature selection in fake news detection on Twitter involves identifying and selecting the most relevant features from tweet data to improve the performance of detection models. This process helps in reducing dimensionality, minimizing noise, and focusing on the most informative aspects of the data (J. Shaikh,et al, 2020).

Table: 3: Feature selection algorithms used in fake news detection

Feature Selection Algorithm	**Description**	**Examples**
Principal Component Analysis (PCA)	Reduces dimensionality by transforming features into a set of linearly uncorrelated components.	Reduces the number of tweet features while preserving the variance that contributes to fake news detection.
Chi-Square Test	Measures the dependency between features and the target variable by comparing observed and expected frequencies.	Identifies significant words or phrases in tweets that are most indicative of fake news.
Information Gain	Evaluates the worth of features based on their ability to improve the purity of class labels in subsets.	Determines which words or phrases in tweets contribute most to distinguishing fake from real news.
Mutual Information	Measures the amount of information obtained about one feature through another feature.	Identifies key features that provide significant information about the authenticity of tweets.
Recursive Feature Elimination (RFE)	Iteratively removes the least important features based on model performance to improve accuracy.	Helps in selecting the most relevant words or topics in tweets for accurate fake news detection.
Genetic Algorithms (GA)	Uses evolutionary strategies to select the best subset of features by simulating natural selection.	Optimizes feature selection by evolving feature subsets over generations to enhance model performance.
Embedded Methods (e.g., LASSO)	Integrates feature selection within the model training process, penalizing less relevant features.	LASSO (Least Absolute Shrinkage and Selection Operator) selects important tweet features while penalizing others.
Filter Methods (e.g., Correlation Coefficient)	Evaluates the relevance of features by measuring their correlation with the target variable.	Selects tweet features that show strong correlations with fake news indicators.
Wrapper Methods (e.g., Forward Selection)	Uses model performance to evaluate feature subsets and select the best combination of features.	Iteratively adds features to improve model accuracy in detecting fake news in tweets.
Embedded Feature Selection (e.g., Tree-based methods)	Uses algorithms like Random Forests or Gradient Boosting to rank features based on their importance in the model.	Identifies key features from tweets that contribute to high accuracy in fake news detection.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VIII, August 2025

These feature selection algorithms help enhance the effectiveness of fake news detection models by focusing on the most relevant features in Twitter data, improving both accuracy and efficiency in identifying misleading information.

Ravish et. al. tackle the challenge of detecting fake news, which can mislead public opinion and undermine democratic processes. They critique manual detection methods, such as those employed by the House of Commons and the Crosscheck project, noting their limited scalability. To address this, the paper presents a novel approach combining Multi-layered Principal Component Analysis (PCA) for feature selection with a firefly-optimized algorithm. The proposed system employs Multi-Support Vector Machines (MSVM) for classification and is tested on ten different datasets. The study highlights that while feature extraction and selection methods significantly enhance accuracy, datasets with fewer features yielded lower performance. This research underscores the importance of effective feature management in improving fake news detection accuracy.

Z. Tian and S. Baskiyar explore the challenge of identifying fake news on social media using machine learning techniques. They propose a detection system utilizing a k-nearest neighbors (KNN) model enhanced by Genetic and Evolutionary Feature Selection (GEFeS), achieving an accuracy of 91.3%. Additionally, they investigate the potential of quantum machine learning by employing a quantum KNN (QKNN) model, which yielded an accuracy of 84.4%. This study highlights the effectiveness of feature selection in improving traditional machine learning approaches and explores innovative quantum techniques for tackling the complexities of fake news detection.

O. Ngada and B. Haskins investigate the effectiveness of various machine learning algorithms in detecting fake news by leveraging features related to article structure, readability, and the alignment between titles and bodies. Their study evaluates six algorithms such as AdaBoost (AB), Decision Tree (DT), K-Nearest Neighbour (KNN), Random Forest (RF), Support Vector Machine (SVM), and XGBoost (XGB)—using performance metrics and confusion matrix analysis. The results reveal that the Support Vector Machine classifier outperformed the others, providing the most promising accuracy in distinguishing fake news from genuine content.

B. Hawashin et.al address the challenge of detecting fake news in Arabic social media posts, a topic that has been relatively underexplored. The study emphasizes the importance of optimizing the feature selection process to enhance detection accuracy. By refining feature selection, their experimental results demonstrate significant improvements in the performance of traditional machine learning methods for detecting fake news in Arabic, highlighting the critical role of feature optimization in advancing detection capabilities.

M. Smith et. al. investigate how evolutionary-based feature selection can enhance the accuracy of fake news detection while significantly reducing the number of features required. The paper highlights the increasing problem of fake news in a technologically driven society and demonstrates that employing evolutionary algorithms for feature selection not only improves detection performance but also simplifies the feature set, making the detection process more efficient. The study underscores the benefits of advanced feature selection techniques in addressing the challenges of fake news detection.

M. Narra et al. examine the impact of various subset feature selection techniques on the performance of fake news detection models. They evaluate Principal Component Analysis (PCA) and Chi-square for feature selection, applied to both machine learning and pre-trained deep learning models. Their findings indicate that the Extra Tree classifier achieves the highest accuracy of 0.9474 when using a combination of Term Frequency-Inverse Document Frequency (TF-IDF) and Bag of Words features. The study also highlights the critical role of preprocessing, as models without adequate preprocessing show significantly lower performance. While Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, Residual Neural Networks (ResNet), and InceptionV3 also perform well, they fall short of the Extra Tree classifier's accuracy. Overall, the study underscores the effectiveness of subset feature selection and preprocessing in enhancing the robustness of fake news detection models.

Classification:

The classification of Twitter data as fake or real involves collecting tweets using relevant keywords, followed by pre-processing steps like text cleaning, normalization, tokenization, and removing stop words. Features are extracted using techniques like TF-IDF or word embeddings. Machine learning models such as Logistic Regression, Naive Bayes, or deep learning models like LSTM are then trained on labelled datasets. The model's performance is evaluated using metrics like accuracy, precision, recall, and F1 score (G. Rawat et al,2023). Once trained and validated, the model can classify new tweets in real-time, identifying fake and real content effectively.

The below studies collectively demonstrate the potential of machine learning and deep learning techniques in effectively classifying Twitter data as fake or real, highlighting the importance of advanced text processing and feature extraction methods.

J. Shaikh and R. Patil address the challenge of fake news by implementing various classification techniques to evaluate their effectiveness. Their study utilizes Support Vector Machine (SVM), Naïve Bayes, and Passive Aggressive Classifier to detect fake news, with Term Frequency-Inverse Document Frequency (TF-IDF) as the primary feature extraction method. The proposed model achieves a notable accuracy of 95.05%, demonstrating the potential of machine learning techniques in combating fake news. The work highlights the ongoing issues of resource limitations in fake news detection and provides a robust solution for improving accuracy in diverse contexts, particularly in the realms of politics and education.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VIII, August 2025

V. Kumar et. al. address the challenge of fake news proliferation on social media platforms like Facebook and Twitter. Their paper focuses on implementing binary classification techniques using artificial intelligence (AI) and machine learning (ML) to differentiate between fake and genuine news content. The study aims to equip users with tools to identify fake news and assess the credibility of news sources. By leveraging natural language processing (NLP) and machine learning principles, the authors provide a solution to mitigate the spread of misleading information and combat click-bait tactics used by spammers.

N. Mittal et. al.explore the pervasive issue of fake news across various platforms, including traditional media and social networks like Facebook, Twitter, and YouTube. The paper emphasizes the need for precise characteristics to develop effective machine learning algorithms for fake news detection. It provides a comprehensive overview of different classifiers employed in the detection process, aiming to enhance the accuracy of distinguishing genuine news from false content. This work contributes to improving user perception and the reliability of news content by focusing on the implementation and analysis of machine learning techniques in fake news detection.

P. Sharma and R. Sahu explore the application of deep learning models for detecting fake news using the ISOT dataset, which includes 137,176 tweets and news data from various sources. The study employs feature extraction techniques such as Term Frequency-Inverse Document Frequency (TF-IDF) and Bag of Words, with word vectors generated through GloVe embeddings. The deep learning models examined include Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, Bidirectional LSTM (BiLSTM), and a hybrid CNN-BiLSTM approach. The research finds that CNN and LSTM models achieve the highest accuracy of 83%, demonstrating the effectiveness of deep learning techniques in distinguishing fake news from genuine content.

T. Bhatia et. al present a novel approach for identifying fake news sources on Twitter by leveraging deep neural networks (DNN). Their method incorporates three primary feature types: tweet content, publication time, and social graph attributes. Evaluating their technique on the PolitiFact dataset from FakeNewsNet, the authors achieve an impressive accuracy of 98.7%, significantly outperforming traditional baseline methods. This research highlights the effectiveness of deep learning in distinguishing fake news sources and demonstrates the potential for advanced models to enhance the reliability of information on social media platforms.

Linmei Hu.et.al. provide an extensive review of deep learning-based methods for fake news detection (FND), focusing on various features including news content, social context, and external knowledge. The survey categorizes existing methods into supervised, weakly supervised, and unsupervised approaches, systematically evaluating representative techniques in each category. The paper also introduces commonly used FND datasets and presents a quantitative analysis of the performance of deep learning methods across these datasets. Finally, the authors identify current limitations and propose future research directions, offering valuable insights into advancing FND methodologies and improving detection accuracy.

Table: 4: Classification of fake news detection on Twitter:

Author	Algorithm	Description	Features Used	Accuracy	Output
J. Shaikh and R. Patil, 2020	**Support Vector Machine (SVM)**	Classifies data by finding the hyperplane that best separates classes.	TF-IDF, Bag of Words	95.05%	Classifying tweets as fake or real based on word frequency features.
J. Shaikh and R. Patil, 2020	**Naïve Bayes**	Probabilistic classifier based on applying Bayes' theorem with strong independence assumptions between features.	TF-IDF, Bag of Words	Varies	Predicting the likelihood of a tweet being fake based on word probabilities.
O. Ngada and B. Haskins, 2020	**Decision Trees**	Models data by splitting it into subsets based on feature values, forming a tree-like structure.	TF-IDF, Bag of Words	Varies	Building a decision tree to categorize tweets based on extracted features like word counts.
O. Ngada and B. Haskins, 2020	**Random Forest**	Ensemble method using multiple decision trees to improve classification accuracy and robustness.	TF-IDF, Bag of Words	Varies	Aggregating predictions from multiple decision trees to identify fake news in tweets.
O. Ngada and B. Haskins, 2020	**Logistic Regression**	Statistical model that uses a logistic function to model binary outcome variables.	TF-IDF, Bag of Words	Varies	Estimating the probability of a tweet being fake based on features like term frequencies.
Z. Tian and S. Baskiyar, 2021	**K-Nearest Neighbors (KNN)**	Classifies data based on the majority label among the nearest neighbors in the feature space.	TF-IDF, Bag of Words	91.3%	Classifying a tweet as fake or real by comparing it to the

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VIII, August 2025

					nearest labelled examples in the feature space.
J. Shaikh and R. Patil, 2020	**Passive Aggressive Classifier**	Incremental classifier that updates its model aggressively when predictions are wrong.	TF-IDF, Bag of Words	Varies	Adapting the model to new data by adjusting based on misclassified tweets.
P. Sharma and R. Sahu, 2023	**Convolutional Neural Network (CNN)**	Deep learning model that uses convolutional layers to extract features from text and detect patterns.	TF-IDF, Bag of Words, GloVe embeddings	83%	Using CNN to identify patterns in text sequences from tweets to classify them as fake or real.
P. Sharma and R. Sahu, 2023	**Long Short-Term Memory (LSTM)**	Recurrent neural network capable of learning long-term dependencies in sequence data.	TF-IDF, Bag of Words, GloVe embeddings	83%	Leveraging LSTM to analyze the sequence of words in tweets for detecting fake news.
P. Sharma and R. Sahu, 2023	**Bidirectional LSTM (BiLSTM)**	LSTM model that processes data in both forward and backward directions for better context understanding.	TF-IDF, Bag of Words, GloVe embeddings	83%	Using BiLSTM to capture context from both past and future words in a tweet for improved classification.
P. Sharma and R. Sahu, 2023	**Hybrid CNN-BiLSTM**	Combines CNN and BiLSTM to leverage both spatial and temporal feature extraction for improved performance.	TF-IDF, Bag of Words, GloVe embeddings	83%	Integrating CNN and BiLSTM to enhance feature extraction and classification accuracy for tweets.
T. Bhatia et al., 2023	**Deep Neural Network (DNN)**	General deep learning architecture used to model complex patterns in data.	Tweet content, Published time, Social graph	98.7%	Detecting fake news sources on Twitter by analysing tweet content, timing, and social connections.
D. R. Collen et al., 2022	**5L-CNN**	Deep learning model with five convolutional layers for feature extraction and classification.	TF-IDF, Bag of Words	99.99%	Using a deep CNN with five layers to detect fake news from tweets with high accuracy.

This table describes the examples of how each algorithm is applied to fake news detection on Twitter.

Proposed model for fake news detection

In this state, to detect such types of fake news and to provide well verified news to our society, the Natural Language Processing (NLP) and Artificial Intelligence (AI) such as Machine learning (ML) and Deep Learning (DL) Techniques are applied in this research. After preprocesses over the actual dataset these methods are effectively identifies the fake news with collected dataset and evaluated by the metrics such as accuracy, precision, recall and F1-measure. The proposed model can be described in below figure,

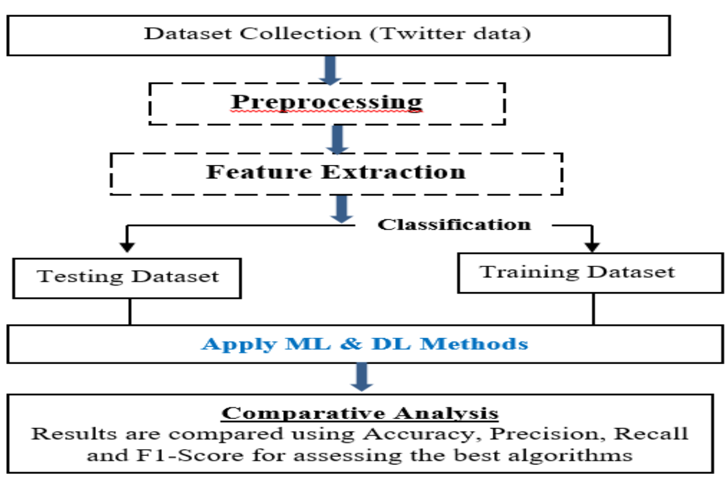

Figure.5. Proposed framework for Online Fake News detection

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue VIII, August 2025

i. ***Dataset Collection:*** The first step is to collect Twitter data for real or fake news analysis. This can be done by using the Twitter API to stream data in real-time or by using a third-party tool to scrape historical data. The collected data may contain text, images, video, and other metadata (Kaliyar, R.K, et al., 2021).

ii. ***Preprocessing:*** The next step is to preprocess the data to ensure it is ready for analysis. This includes tasks such as cleaning the data by removing irrelevant characters, URLs, and mentions, as well as removing stop words and stemming or lemmatizing the text. Also, data labeling needs to be done to differentiate between real and fake news.

iii. ***Feature Extraction:*** After preprocessing, the next step is to extract relevant features from the data. This can be done by using techniques such as Bag-of-Words, TF-IDF, or word embedding's to represent the data in a numerical format that can be fed into ML & DL algorithms.

iv. ***Apply ML & DL Methods:*** With the preprocessed and feature extracted data, ML & DL algorithms can be applied to perform various tasks such as classification or clustering. For real or fake news analysis, supervised learning algorithms such as Naive Bayes, Support Vector Machines, Random Forest, or Deep Learning models such as Convolutional Neural Networks (CNNs) can be used for classification.

v. ***Comparative Analysis:*** Once the ML & DL models have been trained and evaluated, comparative analysis can be done to compare the performance of different algorithms. Metrics such as accuracy, precision, recall, F1-score, and AUC-ROC can be used to evaluate the performance of the models.

Overall, the process of collecting, preprocessing, feature extraction, applying ML & DL methods, and conducting comparative analysis on Twitter data for real or fake news analysis can provide valuable insights into the identification of fake news on social media platforms. The Scikit-learn module in the Python programming language compiled into a script was used to test given dataset by using classification methods, which correctly differentiate the given data into two classes of twitter dataset as 'real' and 'fake. The Scikit-learn module was selected as Python is considered globally to be one of the most popular programming languages for scientific computing and has extensive scientific libraries available.

II. Conclusion

This survey underscores the necessity of integrating advanced pre-processing, feature extraction, and classification techniques to enhance fake news detection on Twitter. The effectiveness of machine learning and deep learning in identifying misinformation highlights their crucial role in improving detection accuracy. The proposed framework, combining NLP and AI, provides a robust approach for managing the rapid spread of fake news. Continued innovation and real-time analytical advancements are essential for effectively combating misinformation and ensuring the reliability of information on social media platforms.

References:

1. E. Cueva, G. Ee, A. Iyer, A. Pereira, A. Roseman and D. Martinez, "Detecting Fake News on Twitter Using Machine Learning Models," *2020 IEEE MIT Undergraduate Research Technology Conference (URTC)*, Cambridge, MA, USA, 2020, pp. 1-5, doi: 10.1109/URTC51696.2020.9668872.
2. V. Kumar, A. Kumar, A. K. Singh and A. Pachauri, "Fake News Detection using Machine Learning and Natural Language Processing," 2021 International Conference on Technological Advancements and Innovations (ICTAI), Tashkent, Uzbekistan, 2021, pp. 547-552, doi: 10.1109/ICTAI53825.2021.9673378.
3. Akhtar, P., Ghouri, A.M., Khan, H.U.R. et al. Detecting fake news and disinformation using artificial intelligence and machine learning to avoid supply chain disruptions. Ann Oper Res 327, 633–657 (2023).
4. Youness Madani, Mohammed Erritali, Belaid Bouikhalene,Using artificial intelligence techniques for detecting Covid-19 epidemic fake news in Moroccan tweets,Results in Physics,Volume 25,2021,104266,ISSN 2211-3797.
5. Kaliyar, R.K., Goswami, A. & Narang, P. FakeBERT: Fake news detection in social media with a BERT-based deep learning approach. Multimed Tools Appl 80, 11765–11788 (2021).
6. Vasu Agarwal, H. Parveen Sultana, Srijan Malhotra, Amitrajit Sarkar,Analysis of Classifiers for Fake News Detection,Procedia Computer Science,Volume 165,2019,Pages 377-383,ISSN 1877-0509,https://doi.org/10.1016/j.procs.2020.01.035.
7. F. Dahlan and S. Suyanto, "Data Augmentations to Improve BERT-based Detection of Covid-19 Fake News on Twitter," 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE), Jakarta, Indonesia, 2023, pp. 140-145, doi: 10.1109/ICCoSITE57641.2023.10127796.
8. T. Bhatia, B. Manaskasemsak and A. Rungsawang, "Detecting Fake News Sources on Twitter Using Deep Neural Network," 2023 11th International Conference on Information and Education Technology (ICIET), Fujisawa, Japan, 2023, pp. 508-512, doi: 10.1109/ICIET56899.2023.10111446.
9. R. Chauhan, S. Upadhyay and H. Vaidya, "Fake News Detection Based on Machine Learning Algorithm," 2023 3rd International Conference on Innovative Sustainable Computational Technologies (CISCT), Dehradun, India, 2023, pp. 1-5, doi: 10.1109/CISCT57197.2023.10351408.
10. M. Heidari et al., "BERT Model for Fake News Detection Based on Social Bot Activities in the COVID-19 Pandemic," 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 2021, pp. 0103-0109, doi: 10.1109/UEMCON53757.2021.9666618.

11. M. Chakraborty, S. Das and R. Mamidi, "Detection of Fake Users in Twitter Using Network Representation and NLP," 2022 14th International Conference on COMmunication Systems & NETworkS (COMSNETS), Bangalore, India, 2022, pp. 754-758, doi: 10.1109/COMSNETS53615.2022.9668371.
12. K. Gupta and K. Potika, "Fake News Analysis and Graph Classification on a COVID-19 Twitter Dataset," 2021 IEEE Seventh International Conference on Big Data Computing Service and Applications (BigDataService), Oxford, United Kingdom, 2021, pp. 60-68, doi: 10.1109/BigDataService52369.2021.00013.
13. T. K. Bhatia, K. Chauhan and R. Suden, "A Novel Technique to Detect the Fake News by Using the Machine Learning Approaches," 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2022, pp. 1-6, doi: 10.1109/ICRITO56286.2022.9964886.
14. Sultana, M. Islam, M. Hasan and F. Ahmed, "Fake News Detection Using Machine Learning Techniques," 2023 IEEE/ACIS 21st International Conference on Software Engineering Research, Management and Applications (SERA), Orlando, FL, USA, 2023, pp. 98-103, doi: 10.1109/SERA57763.2023.10197712.
15. Q. Abbas, M. U. Zeshan and M. Asif, "A CNN-RNN Based Fake News Detection Model Using Deep Learning," 2022 International Seminar on Computer Science and Engineering Technology (SCSET), Indianapolis, IN, USA, 2022, pp. 40-45, doi: 10.1109/SCSET55041.2022.00019.
16. Balshetwar, S.V., RS, A. & R, D.J. Fake news detection in social media based on sentiment analysis using classifier techniques. Multimed Tools Appl 82, 35781–35811 (2023).
17. J. Shaikh and R. Patil, "Fake News Detection using Machine Learning," 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), Gunupur Odisha, India, 2020, pp. 1-5, doi: 10.1109/iSSSC50941.2020.9358890.
18. N. F. Baarir and A. Djeffal, "Fake News detection Using Machine Learning," 2020 2nd International Workshop on Human-Centric Smart Environments for Health and Well-being (IHSH), Boumerdes, Algeria, 2021, pp. 125-130, doi: 10.1109/IHSH51661.2021.9378748.
19. G. Rawat, T. Pandey, T. Singh, S. Yadav and P. K. Aggarwal, "Fake News Detection Using Machine Learning," 2023 International Conference on Artificial Intelligence and Smart Communication (AISC), Greater Noida, India, 2023, pp. 759-762, doi: 10.1109/AISC56616.2023.10085488.
20. D. R. Collen, L. K. Nyandoro and K. Zvarevashe, "Fake News Detection using 5L-CNN," 2022 1st Zimbabwe Conference of Information and Communication Technologies (ZCICT), Harare, Zimbabwe, 2022, pp. 1-7, doi: 10.1109/ZCICT55726.2022.10045981.
21. Ravish, R. Katarya, D. Dahiya and S. Checker, "Fake News Detection System Using Featured-Based Optimized MSVM Classification," in IEEE Access, vol. 10, pp. 113184-113199, 2022, doi: 10.1109/ACCESS.2022.3216892.
22. Z. Tian and S. Baskiyar, "Fake News Detection using Machine Learning with Feature Selection," 2021 6th International Conference on Computing, Communication and Security (ICCCS), Las Vegas, NV, USA, 2021, pp. 1-6, doi: 10.1109/ICCCS51487.2021.9776346.
23. Hawashin, A. Althunibat, T. Kanan, S. AlZu'bi and Y. Sharrab, "Improving Arabic Fake News Detection Using Optimized Feature Selection," 2023 International Conference on Information Technology (ICIT), Amman, Jordan, 2023, pp. 690-694, doi: 10.1109/ICIT58056.2023.10225974.
24. O. Ngada and B. Haskins, "Fake News Detection Using Content-Based Features and Machine Learning," 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), Gold Coast, Australia, 2020, pp. 1-6, doi: 10.1109/CSDE50874.2020.9411638.
25. M. Smith, A. Richardson, B. Brown, G. Dozier, M. King and J. Morris, "A Study of the Impact of Evolutionary-Based Feature Selection for Fake News Detection," 2020 IEEE Symposium Series on Computational Intelligence (SSCI), Canberra, ACT, Australia, 2020, pp. 1859-1865, doi: 10.1109/SSCI47803.2020.9308613.
26. M. Narra et al., "Selective Feature Sets Based Fake News Detection for COVID-19 to Manage Infodemic," in IEEE Access, vol. 10, pp. 98724-98736, 2022, doi: 10.1109/ACCESS.2022.3206963.
27. J. Shaikh and R. Patil, "Fake News Detection using Machine Learning," 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), Gunupur Odisha, India, 2020, pp. 1-5, doi: 10.1109/iSSSC50941.2020.9358890.
28. V. Kumar, A. Kumar, A. K. Singh and A. Pachauri, "Fake News Detection using Machine Learning and Natural Language Processing," 2021 International Conference on Technological Advancements and Innovations (ICTAI), Tashkent, Uzbekistan, 2021, pp. 547-552, doi: 10.1109/ICTAI53825.2021.9673378.
29. P. Sharma and R. Sahu, "Fake News Detection Using Deep Learning Based Approach," 2023 International Conference on Circuit Power and Computing Technologies (ICCPCT), Kollam, India, 2023, pp. 651-656, doi: 10.1109/ICCPCT58313.2023.10245933.
30. T. Bhatia, B. Manaskasemsak and A. Rungsawang, "Detecting Fake News Sources on Twitter Using Deep Neural Network," 2023 11th International Conference on Information and Education Technology (ICIET), Fujisawa, Japan, 2023, pp. 508-512, doi: 10.1109/ICIET56899.2023.10111446.
31. lINMEI Hu, Siqi Wei, Ziwang Zhao, Bin Wu,Deep learning for fake news detection: A comprehensive survey,AI Open,Volume 3,2022,Pages 133-155, ISSN 2666-6510,https://doi.org/10.1016/j.aiopen.2022.09.001.