International Journal of Engineering Research Excellence and Applied Science (IJEREAS)

Volume-III (Issue 4) - December 2025 ISSN: 3048-5355

Fake News Detection Using Machine Learning Techniques: A Review

Jaspreet Singh Kamal

M.Tech Scholar

Department of Computer Science Engineering

Jaipur Institute of Technology, Group of Institutions, Jaipur, Rajasthan

Dr. Swapnil Singhal

Associate Professor

Department of Computer Science Engineering

Jaipur Institute of Technology, Group of Institutions, Jaipur, Rajasthan

Abstract- In recent years, consuming news from social media has become increasingly popular, gradually surpassing traditional sources such as newspapers, radio, and television. However, these very advantages also create a fertile ground for the spread of misinformation, resulting in the circulation of vast amounts of fake news. Unlike traditional news outlets that maintain editorial oversight and fact-checking processes, social media platforms often lack strict quality control mechanisms, which undermines the credibility and trustworthiness of the information being consumed. This has raised widespread concern, as fake news has the potential to distort public opinion, damage reputations, incite social unrest, and even influence political or economic decisions.

In nowadays's virtual world, the speedy growth of social media and online news systems has made it easier to get entry to data but has additionally raised severe issues about the spread of fake information. Fake news can impact public opinion, manage political debates, or even create social unrest. Therefore, growing efficient and accurate fashions to stumble on fake information has end up a crucial studies area in statistics technological know-how, device mastering, and natural language processing (NLP). One of the widely studied approaches in this domain combines Term Frequency–Inverse Document Frequency (TF-IDF) with the Passive-Aggressive Classifier, supplying a stability of simplicity, computational performance, and effectiveness in textual content classification

obligations.

The rapid growth of social media and online news platforms has significantly changed the way information is consumed, while simultaneously increasing the spread of fake news. The absence of strict editorial control on digital platforms allows misinformation to circulate widely, posing risks to public opinion, social stability, and decision-making processes. This paper presents a comprehensive review of fake news detection techniques using Natural Language Processing (NLP) and Machine Learning (ML) methods. It discusses commonly used approaches, algorithms, and text preprocessing strategies, with particular emphasis on feature extraction techniques such as TF-IDF. The review highlights the effectiveness, challenges, and limitations of existing models and outlines future research directions for improving fake news detection systems.

Keywords: Fake news detection, Natural Language Processing (NLP), Machine learning, TF-IDF, Text classification, Social media misinformation

I. INTRODUCTION

Fake News and Its Growing Concern

In the present-day virtual age, statistics have turned out to be the maximum effective device shaping societies, economies, and governance structures. The internet and social media systems have enabled human beings to access news inside seconds, making verbal exchange quicker than ever before. While this revolution in records sharing has added undeniable advantages, it has also created serious challenges. Among them, the upward thrust of faux information stands proud as a vital international subject. Fake news refers back to the planned advent and dissemination of false or misleading statistics, often with the reason to mislead, control, or provoke. Unlike misinformation, which may be shared by chance, fake news is generally produced with calculated motives such as political propaganda, financial benefit, or social disruption. The growing reach and influence of fake information have raised alarms in fields starting from politics and journalism to public fitness and countrywide safety.

Machine Learning for Fake News Detection

The upward thrust of digital media has converted the way humans eat statistics, however it has

additionally fueled the spread of faux news—false or deceptive statistics created to deceive, control, or generate interest. Fake news poses a extreme risk to democratic processes, public fitness, social consider, and financial stability. Traditional fact-checking techniques are gradual and can't maintain pace with the extent of misinformation circulating online. As a end result, gadget learning (ML) has emerged as one of the only answers for detecting and preventing fake information at scale. By leveraging algorithms that learn styles from records, ML enables the automated class of information content material into real or faux, thereby assisting quicker interventions and more reliable data ecosystems.

Why Machine Learning for Fake News Detection?

Machine learning is nicely-appropriate for fake information detection because it can manner massive volumes of unstructured textual content speedy, examine from styles in historic data, and continuously adapt to new kinds of incorrect information. Unlike rule-based totally structures, ML algorithms can manage the complexity of herbal language, locate diffused cues in writing styles, and generalize throughout distinct subjects. Moreover, the mixing of ML with herbal language processing (NLP) enables the extraction of linguistic, semantic, and contextual functions from information articles, headlines, and social media posts, allowing the version to distinguish among authentic and misleading content with high accuracy.

Another advantage of ML is its scalability and flexibility. With thousands and thousands of new posts, tweets, and articles generated each day, manual reality-checking cannot hold up. Machine mastering systems, but, can reveal statistics streams in real-time, flag suspicious content material, and assist truth-checkers with the aid of narrowing down instances requiring human assessment. In this way, ML acts as each a frontline defense towards misinformation and a assist device for newshounds, researchers, and policymakers.

II. APPROACHES AND ALGORITHMS IN FAKE NEWS DETECTION

Machine learning for fake news detection typically involves three key steps**:** feature extraction, model training, and classification**.**

Feature extraction Is a vital step in machine learning and natural language processing that transforms raw information into meaningful numerical representations that algorithms can interpret.

In the context of fake news detection, characteristic extraction makes a speciality of figuring out textual, linguistic, and contextual traits that distinguish actual news from fabricated content material. Common strategies encompass the Bag of Words (BoW) model, which counts phrase occurrences, and Term Frequency–Inverse Document Frequency (TF-IDF), which assigns higher weights to words which might be frequent in a file however rare across the corpus, making them greater discriminative. Advanced techniques including phrase embeddings (e.G., Word2Vec, GloVe, BERT) seize semantic and contextual relationships between words, imparting deeper insights into how phrases are used. Beyond textual functions, stylistic attributes together with sentence duration, punctuation, sentiment, clarity, and metadata (supply credibility, user engagement) also can be extracted to improve detection. By changing unstructured text into established vectors, feature extraction affords the inspiration upon which gadget mastering models can learn styles and classify news articles as real or fake.

Classification algorithms Are supervised gadget studying techniques used to categorize information into predefined instructions primarily based on learned styles. In fake news detection, these algorithms are trained on labeled datasets of actual and faux news to build predictive fashions that may become aware of incorrect information in unseen articles. Commonly carried out algorithms consist of Logistic Regression, which models the chance of a news object being genuine or fake; Naïve Bayes, which leverages probabilistic word distributions; and Support Vector Machines (SVMs), which separate instructions the usage of foremost hyperplanes. Ensemble methods consisting of Random Forests and Gradient Boosting improve accuracy by using combining multiple susceptible classifiers. In addition, online studying fashions like the Passive-Aggressive Classifier are specifically effective for real-time detection in dynamic news streams. More advanced processes integrate deep studying strategies including Recurrent Neural Networks (RNNs) and Transformers, which capture contextual and semantic features of language. By leveraging these algorithms, fake news detection systems can attain sturdy type overall performance, assisting efforts to counter incorrect information on digital platforms.

Deep learning approaches Have emerged as powerful techniques for fake information detection because of their capability to robotically study hierarchical and complex representations from raw textual content information. Unlike traditional gadget mastering fashions that depend closely on

guide characteristic extraction, deep learning fashions inclusive of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks can seize semantic, syntactic, and contextual styles inside news articles. More advanced architectures like Bidirectional LSTMs (BiLSTMs) and Gated Recurrent Units (GRUs) further decorate the potential to model long-variety dependencies in textual content. Recently, Transformer-based totally models consisting of BERT, RoBERTa, and GPT have achieved present day performance by using leveraging interest mechanisms to apprehend phrase relationships in context. These fashions now not handiest enhance category accuracy however additionally adapt efficaciously to multilingual and domain-precise fake news datasets. Despite their excessive accuracy, deep getting to know tactics require large datasets, large computational assets, and cautious dealing with of biases in education information. Nevertheless, they constitute a transformative step in automating the detection of incorrect information on social media and news structures.

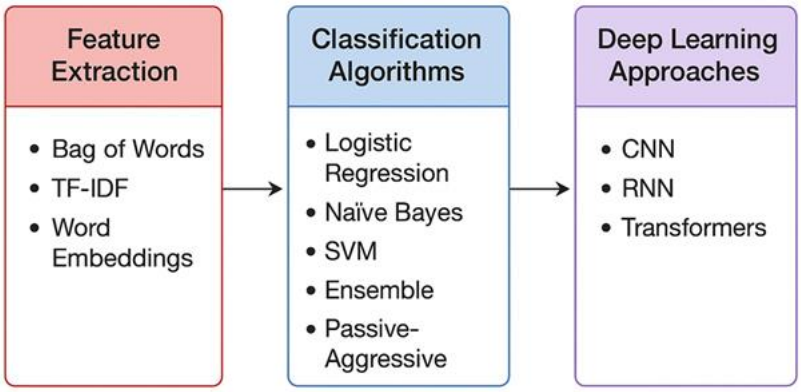

Fig. 1. Approaches and Algorithms for Fake News Detection

III.LITERATURE REVIEW

Patel, H. et.al (2025): The study "Fake News Detection Using Machine Learning Approaches" explores the use of machine learning algorithms for identifying and classifying false information. It uses models like naive Bayes classifier, cognitive systems backtracking, NLP techniques, CSI model, ensemble techniques, attention processes, LSTMs, and CNNs. The study also discusses machine learning ensemble approaches to categorize news articles as real or fake. It suggests a machine learning-based solution using SVM classifiers, text preprocessing, encoding, and attribute extraction to identify bogus news. The study also discusses the challenges of detecting fake news

with high precision, automated index scoring, and human detection.

Gunashekhar, S.et.al (2025): The spread of fake news has become a significant issue in the digital age, affecting societal trust and decision-making. This paper presents a machine learning-based approach to fake news detection using the Passive-Aggressive Classifier. By leveraging natural language processing (NLP) for text preprocessing and feature extraction, the system classifies news articles as real or fake. The model demonstrates high efficiency and scalability, making it suitable for real-time applications such as social media moderation. This study highlights the effectiveness of the Passive-Aggressive Classifier for dynamic environments and provides insights into future enhancements to improve detection accuracy and applicability.

Raj, A. et.al (2024): Fake news, propelled by the surge in social media and online platforms, poses a significant threat to public opinion, societal stability, and democratic processes. Its historical roots in misinformation, propaganda, and hoaxes have evolved with virtual age, exploiting the unregulated nature of the internet. Social media's accessibility and rapid information sharing contribute to the swift dissemination of false information, impacting public perception, decision-making, and even elections. The consequences include panic, confusion, and the potential incitement of violence, emphasizing the urgent need for robust detection systems.

Sadikot, J.et.al (2024): Determination of fake news has become a crucial research subject, considering the voluminous and misleading social media content, hence, important to determine the authenticity of the news on the internet. This paper conducts a thorough analysis of current research endeavors in fake news detection, particularly emphasizing the effectiveness of traditional machine learning techniques. The primary aim is to develop a robust model using supervised machine learning algorithms to accurately classify news articles as authentic or deceptive. To achieve this objective, various tools and technologies have been used. Through this interdisciplinary approach, it has been tried to contribute to the ongoing efforts to combat the proliferation of fake news and maintain the integrity of information dissemination in the digital realm. In all the graphs developed for accuracy metrics like precision, recall, and F1-score, the Passive Aggressive Classifier consistently shows the best values. It proves to be the most effective classifier for distinguishing between real and fake news articles which proves to be beneficial to society.

Shabani, V.et.al (2023): The rapid growth of fake news, as well as its damaging effects on every area of our lives, has increased the demand for detecting and combating fake news. As a result, distinguishing between real and fake news is critical. However, due to the massive amount of information generated every minute on the Internet, making this distinction manually is extremely difficult. This study will suggest an approach for detecting fake news and a mechanism for implementing it on social media. In this paper, the Naive Bayes Classifier and Passive Aggressive Classifier techniques will be used to detect fake news. The results will prove that the problem of identifying fake news is possible if Machine learning and Natural Language Processing algorithm are used.

*Samarpita, S.*et.al (2023):Majority of the people get affected with misleading stories spread through different posts on social media and forward them assuming that it is a fact. Nowadays, Social media is used as a weapon to create havoc in the society by spreading fake news. Such havoc can be controlled by using machine-learning algorithms. Various methods of machine learning and deep learning techniques are used to identify false stories. There is a need for identification and controlling of fake news posts that have increased in alarming rate. Here we use Passive-Aggressive Classifier for fake news identification. Two datasets, Kaggle fake news dataset and as well as dynamically web scrapped dataset from politifact.com website. We achieved 88.66% accuracy using Passive Aggressive Classifier.

Trivedi, N. et.al (2022): Nowadays, extensive use of social media and the internet has given rise to fake news that is fabricated articles intentionally build to mislead readers. Generally, Humans are not good at distinguishing facts and fakes that may cause serious damage to government, market, and society. We made our dataset from 5 publicly available da-tasets to avoid bias towards Indian context-based news articles. By using various feature extraction techniques and Machine learning algorithms, we were able to get an accuracy of 93.90%. The proposed method can be used to debunk false information.

Mohd., N. et.al (2021): During this modern time when people have such easy access to social media and internet, they have become very connected and sometimes it takes less than a minute for any piece of information to get viral. Consequently, fake news is spreading very easily and quickly on those platforms which lead to inappropriate actions and consequences. Sometimes this fake

information is spread by people intentionally to create chaos among the people. The aim of this paper is to create a model wherein it classifies each piece of information provided to it into two categories, either non-hostile, fake, offensive or defamation. In order to accomplish the project's objective, we have suggested the algorithm term frequency - inverse document frequency. It is a statistical technique for determining how pertinent a word is in a group of documents. It is widely used for information retrieval and summarization. This method is used to calculate the term frequency of each word and further classify the words according to the occurrence in each particular type of document in the training dataset. Furthermore, the testing dataset uses this information to find the nature of the text document. This method is proposed for the natural language “Hindi” because of the unavailability of such fake news classifiers in this language.

Kudari, J.et.al (2020) Consuming news from social media is becoming increasingly popular nowadays. Social media brings benefits to users due to the inherent nature of fast dissemination, cheap cost, and easy access. However, the quality of news is considered lower than traditional news outlets, resulting in large amounts of fake news. Detecting fake news becomes very important and is attracting increasing attention due to the detrimental effects on individuals and the society. The performance of detecting fake news only from content is generally not satisfactory, and it is suggested to incorporate user social engagements as auxiliary information to improve fake news detection. Thus it necessitates an in- depth understanding of the correlation between user profiles on social media and fake news. We perform a comparative analysis over explicit and implicit profile features between these user groups, which reveals their potential to differentiate fake news.

Kanavos.et.al (2019): The detection of fake news is a crucial task in today’s society, given the widespread use of social media and online platforms. In this study, we investigate the application of Machine Learning (ML) algorithms for the detection of fake news. We consider two different datasets of categorized news articles of various sizes and apply various ML algorithms, along with two methods of text vectorization. Specifically, w e e xamine B ag of Words and Tf-Idf, with the use of stemming and with different ngram values. The resulting vectors are processed by Naive Bayes algorithms, Linear algorithms, Support Vector Machines, and Random Forest classifiers. F 1-score and computational t ime for each algorithm-vectorization combination were recorded. Our results have shown that Linear Algorithms and Support Vector Machines combined with Tf-Idf vectors and

n-gram value of (1,2) produced the highest accuracies, with an F1-score up to 96.8%. Index Terms—Machine Learning, Text Mining, Information Retrieval, Fake News Detection

IV. SOCIAL MEDIA AS A MEDIUM FOR FAKE NEWS

Social media has emerged as one of the most influential platforms for communication, statistics sharing, and public engagement in these days's virtual technology, but it has additionally become a powerful medium for the unfold of fake news. The ease of content creation and on the spot dissemination, mixed with the dearth of strict regulatory mechanisms, has enabled incorrect information to unfold rapidly throughout platforms like Facebook, Twitter, Instagram, and WhatsApp. Fake information flourishes on the algorithms of social media that prioritize engagement over accuracy, allowing sensational or emotionally charged content to attain thousands and thousands within seconds, regularly faster than reality-checked news from credible assets. Moreover, the participatory nature of social media, where users can like, share, and remark, amplifies the movement of fake information with out verification, creating echo chambers that reinforce pre-current ideals and biases. The anonymity furnished online additionally encourages malicious actors, which include political corporations, business entities, or even overseas companies, to deliberately unfold misleading narratives for propaganda, financial benefit, or social disruption. Fake news on social media has considerable effects, inclusive of influencing public opinion, inciting violence, undermining believe in journalism, or even affecting democratic strategies like elections. Additionally, the usage of manipulated photos, deepfakes, and misleading headlines makes it increasingly difficult for the overall public to distinguish among reality and fiction. While social media companies have started introducing measures which include fact-checking partnerships, content moderation, and reporting mechanisms, the assignment stays colossal because of the quantity and pace of statistics shared online.

Therefore, the function of media literacy, important wondering, and cognizance among customers is as crucial as technological interventions in preventing the menace of fake news. Ultimately, whilst social media offers unprecedented opportunities for connectivity and data change, its role as a breeding ground for faux information highlights the urgent want for responsible usage, more potent policies, and collaborative efforts among governments, tech organizations, and civil society to guard

the integrity of records inside the digital age.

TABLE 1: SOCIAL MEDIA AS A MEDIUM FOR FAKE NEWS WITH CATEGORIES, CHARACTERISTICS, AND IMPACTS

Aspect	Description	Impact on Fake News Spread
Accessibility	Social media platforms are open and easily accessible to billions of users worldwide.	Enables rapid dissemination of fake news with minimal barriers.
Virality & Sharing	Features like likes, shares, and retweets amplify content visibility.	Fake news spreads faster than corrections due to emotional or sensational appeal.
User-Generated Content	Content is created by individuals rather than verified institutions.	Increases the volume of unverified information and misinformation.
Algorithmic Amplification	Recommendation systems prioritize trending and engaging content.	Fake news often receives higher visibility due to high engagement.
Echo Chambers	Users interact within like-minded communities.	Reinforces pre-existing beliefs, reducing exposure to fact-checks.
Anonymity & Pseudonymity	Users can hide identity or create fake accounts.	Encourages malicious actors to spread false content without accountability.
Low Regulation	Limited monitoring of posts compared to traditional media.	Creates opportunities for coordinated misinformation campaigns.

Speed vs. Verification	News spreads instantly, while fact-checking takes time.	Fake narratives often gain traction before truth is established.

V. IMPORTANCE OF PREPROCESSING IN TEXT DATA

In herbal language processing (NLP) and textual content mining, preprocessing is a essential step that drastically affects the accuracy, efficiency, and interpretability of machine gaining knowledge of fashions. Raw text facts is inherently unstructured, frequently containing noise including spelling errors, unique characters, redundant information, and variations in word paperwork. Without right preprocessing, algorithms can also war to apprehend significant styles, main to bad version overall performance. Thus, preprocessing acts as a bridge between uncooked textual content and effective device gaining knowledge of packages.

One of the number one goals of text preprocessing is standardization. Techniques which includes tokenization, lowercasing, forestall-word removal, and stemming or lemmatization assist make certain that phrases with similar meanings are represented continually.

Another important aspect is handling noisy and high-dimensional data. Text datasets often contain thousands of unique terms, many of which are irrelevant or occur infrequently. Preprocessing techniques like filtering rare words, applying n-grams, and converting text into numerical representations such as term frequency–inverse document frequency (TF-IDF) or word embeddings (Word2Vec, GloVe, BERT) help capture semantic meaning while minimizing computational complexity (Cambria & White, 2014). By transforming text into structured numerical features, models can more effectively learn relationships within the data.

Preprocessing also plays a vital role in improving generalization and model accuracy. For example, in sentiment analysis or spam detection, unprocessed data may lead to overfitting due to inconsistencies like capitalization or punctuation patterns. Normalization ensures that models focus on meaningful features rather than superficial differences. Furthermore, techniques like text cleaning, spelling correction, and handling emojis or hashtags are particularly important in social media analytics, where informal and unstructured language dominates (Krouska et al., 2016).

The importance of preprocessing extends past model overall performance to interpretability and

efficiency. Clean and properly-established statistics permits simpler visualization, topic modeling, and clustering, imparting greater insightful consequences for decision-making. Additionally, lowering needless complexity lowers computational price and training time, which is important in huge-scale programs such as search engines like google, recommendation structures, and real-time chatbots.

VI. CONCLUSION & DISCUSSION

This paper presented a comprehensive review of fake news detection using Natural Language Processing (NLP) and Machine Learning (ML) techniques. The discussion highlighted various approaches and algorithms, showing that both traditional ML models and advanced deep learning methods play a crucial role in identifying misleading information. The literature review revealed that no single technique is universally effective, and hybrid models often achieve better performance. Social media platforms were identified as the primary medium for the rapid spread of fake news due to their high user engagement and ease of content sharing. The categorization and characteristics of fake news further emphasized its serious social, political, and economic impacts. The study also discussed the importance of text preprocessing, as effective cleaning, normalization, and feature extraction significantly improve model accuracy. Overall, this review underlines the need for robust, scalable, and adaptive fake news detection systems. Future research should focus on multilingual datasets, real-time detection, and explainable AI models to enhance reliability. Integrating domain knowledge with NLP and ML techniques can further strengthen fake news detection frameworks.

References

[1] Gupta, S., & Meel, P. (2020). Fake news detection using passive-aggressive classifier. *Lecture Notes in Networks and Systems*, 145, 155–164.

[2] Gunashekhar, S., Nishanthi, S., Gayathri, P., Ramya, S., & Prasath, G. J. (2025). Fake news detection using passive aggressive classifier. *International Journal of Progressive Research in Engineering Management and Science*, 5(1), 1233–1235.

[3] A comprehensive guide to fake news detection using passive-aggressive classifier and TF-IDF. (2023). *Python PlainEnglish*.

[4] Fake news detection using passive aggressive and TF-IDF vectorizer. (2020). *International*

Research Journal of Engineering and Technology, 7(12), 903.

[5] AI-driven fake news detection: A machine learning-based approach. (2025). *International Journal of Research Publication and Reviews*, 6(3), 6225–6232.

[6] Ali Adeeb, R., & Mirhoseini, M. (2023). The impact of affect on the perception of fake news on social media: A systematic review. *Social Sciences*, 12(12), Article 674.

[7] False news or fake news may have limited effects beyond increasing beliefs in false claims. (2020). *Misinformation Review*.

[8] AI-generated content raises risks of more bank runs, UK study shows. (2025, February 14). *Reuters*.

[9] It's only a matter of time before disinformation leads to disaster. (2024). *Financial Times*.

[10] Karnataka fake news draft bill raises concerns. (2025). *The Economic Times*.

[11] Fake news. (n.d.). In *Wikipedia*.

[12] Shah, S., & Patel, S. (2025). A comprehensive survey on fake news detection using machine learning. *Journal of Computer Science*, 21(4), 982–990.

[13] Choras, M., Demestichas, K., Gielczyk, A., Herrero, Á., Ksieniewicz, P., Remoundou, K., & Wozniak, M. (2020). Advanced machine learning techniques for fake news detection: A systematic mapping study. *arXiv*.

[14] Khan, J. Y., Khondaker, M. T. I., Afroz, S., Uddin, G., & Iqbal, A. (2019). A benchmark study of machine learning models for online fake news detection. *arXiv*.

[15] Ahmed, S., Hinkelmann, K., & Corradini, F. (2022). Combining machine learning with knowledge engineering to detect fake news in social networks: A survey. *arXiv*.

[16] Gong, S., Sinnott, R. O., Qi, J., & Paris, C. (2023). Fake news detection through graph-based neural networks: A survey. *arXiv*.

[17] Exploring machine learning for fake news detection: Techniques and trends. (2025). *Springer*.

[18] Approaches to identify fake news: A systematic literature review. (2020). *Procedia Computer Science*.

[19] A brief survey for fake news detection via deep learning models. (2022). *Procedia Computer Science*.

[20] Singh, R. (2022). A review of machine learning-based fake news analysis. *International*

Journal of Scientific Research in Science and Technology, 9(2), 252–261.

[21] A review of fake news detection approaches: A critical analysis. (2023). *ScienceDirect*.

[22] Khan, Z. A., & V., R. (2023). Fake news detection using TF-IDF weighted with Word2Vec: An ensemble approach. *International Journal of Intelligent Systems and Applications in Engineering*, 11(3), 1065–1076.

[23] Li, K. (2021). A fake news detection method based on TF-IDF and ensemble machine learning. *CEUR Workshop Proceedings*, 2943.