Beyond the Text: Integrating Fake News Detection with Emotions and Contextual Cues

Aliyu Ibrahim Sulaiman[1*]; Dr. Fakhrun Jamal[2]; Abdurrazaq Jibril Baba[3]; Yahaya Salisu[4]; Mukhtar Yunusa Maijamaa[5]; Dr. Yusuf Aliyu Adam[6]

[1;2;3;4;5;6]Department of Computer Science and Engineering, Faculty of Engineering and Technology, Shobhit Institute of Engineering & Technology, Uttar Pradesh, India

[1][https://orcid.org/0009-0008-0061-9010]
[2][https://orcid.org/0000-0003-0004-4882]
[3][https://orcid.org/0009-0002-4042-0723]
[4][https://orcid.org/0009-0007-8273-7119]
[5][https://orcid.org/0009-0007-6849-4528]
[6][https://orcid.org/0009-0005-2520-6587]

Corresponding Author: Aliyu Ibrahim Sulaiman[1*]

Publication Date: 2025/12/16

Abstract: As the use of digital platforms increases, the proliferation of fake news plays a pivotal role in influencing trust in their use, both ethically and socially, as well as politically. Traditional fake news detection system focuses on Machine Learning (ML), Natural Language Processing (NLP), and Deep Learning (DL), which naturally rely on textual analysis. The review emphasizes an extensive approach to incorporating contextual cues and emotional influences on behavior, which influence the validity of the information. The research examined the existing literature and identified gaps within the model's architecture, including a lack of a proper emotional dataset, inadequate support for multiple languages, insufficient implementation of ethics in behavior training for user behavior, and limited real-time detection proficiency. The study suggests integrating explainable AI (XAI) for transparency in model prediction, multi-cultural emotion modeling, and differential privacy to protect users' data privacy, as well as addressing challenges in adopting the technology. The paper highlight the importance adoption of hybrid modelling, which also boosts the accuracy of the detection system. The goal is to contribute to the development of a transparent, effective, and robust model for detecting fake news, supporting multi-cultural and diverse linguistic contexts. The review contributes to advancing research on creating more computational fact-checking systems by integrating emotional and contextual cues as a way forward in alleviating fake news in the digital era.

Keywords: *Fake News; Contextual Analysis; Emotional Cues; Machine Learning; Natural Language Processing; Explainable AI.*

How to Cite: Aliyu Ibrahim Sulaiman; Dr. Fakhrun Jamal; Abdurrazaq Jibril Baba; Yahaya Salisu; Mukhtar Yunusa Maijamaa; Dr. Yusuf Aliyu Adam (2025) Beyond the Text: Integrating Fake News Detection with Emotions and Contextual Cues. *International Journal of Innovative Science and Research Technology*, 10(12), 811-821. https://doi.org/10.38124/ijisrt/25dec547

I. INTRODUCTION

In the era of the digital world and the increase of mediums of passing information globally, especially online blogs, digital newspapers (BBC, CNN, Al-Jazeera, VOA, etc.), social media platforms (Facebook, X, Instagram, etc.), and other digital media, which enable easy access to information publically has created ambiguity in determining the fate of news. The term "fake news "refers to a medium of sharing information that causes harm, intentionally or unintentionally, usually about the promotion of a particular moral or political cause or point of view [1], leading to the spread of a vast amount of news from different parts of the world, driven by various intentions, including recognition, trending topics, financial gain, political motivations, cybercrimes, and other means to mislead readers. According to a research, an estimate of more than 80% citizens of EU come across fake news at least once in a week, which is considered as treat to societal, economic, security, and political development of the region. To mitigate the occurrence of such in the society, half of EU people between aged 15 to 30 years need a reliable source of information to counteract fake news and extremism in the region. [2]. Similarly, in the likes of Africa, fake misinformation has created chaos in some advanced countries; in Nigeria, targeting the presidential candidate, Kenya a quote shared by Richard Quest of CNN, Ethiopia on ethnic violence, and the false resignation of South African president Jacob Zuma were among the news that created several chaos among the citizens and causing the damage against country in the view of the world, national security, and those involved [3].

With the evolution of Artificial Intelligence, Machine Learning, and Deep Learning Models in detecting fake misinformation, several techniques were introduced to filter and identify the tricks and patterns applied in a textual format related to keywords, grammar, language patterns, or sentence structure, which can lead to false detection of news accuracy. The model may also have a limitation in that it does not constantly develop to outright lies. It features real facts and information to mislead readers, who may believe the narrative is true if it effectively evokes emotional responses and employs other persuasive language techniques to convince them of its contents. The integration of Natural Language Processing (NLP), which utilises syntactic structures and statistical language models, makes it more challenging to falsify the information by ignoring other external factors that the information contained in the produced text may not account for. Addressing these challenges requires a different dimension to go beyond the text, with recent research beginning to explore contextual cues and emotions to determine user behaviour patterns, sentiment analysis, and emotion analysis.

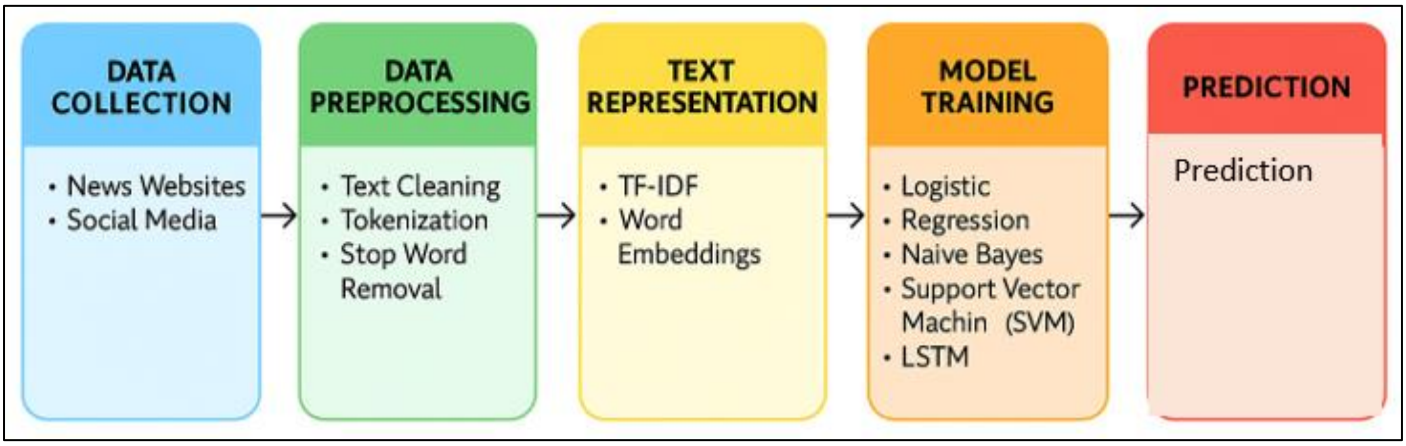

Fig 1 Textual Fake News Detection Approach

These processes must combined into fake news detection algorithms and models to provide reliable information globally, and seek to address five key questions:

- RQ1: How does the use of digital platforms contribute to promoting fake news?
- RQ2: What are the consequences of fake news in collaboration with digital platforms globally?
- RQ3: What models are used in detecting fake news, and what are their limitations?
- RQ4: How can AI/ML, Deep Learning, and NLP improved with the most trending techniques in processing fake news models?
- RQ5: How can emotional and contextual features contribute to the detection of fake news accuracy?

By exploring these questions, we aim to propose proper methods to advance fake news detection approaches, providing more reliable model development techniques that go beyond the text. We will review several approaches used in the processes, identify gaps and limitations, and propose a model-based system to enhance the future of fake news models.

The main contribution of the research is to:

- Contribute to eliminating gaps in the current models, such as the scarcity of emotional datasets, the lack of integration of multilingual models, and ethical concerns in modelling user behaviour.
- Providing a beyond-text approach to integrating emotional and contextual cues in building fake news detection systems.
- There is a need to develop and integrate real-time models for detecting misinformation on digital media platforms.
- Promoting the adaptation of hybrid techniques in modelling involves more user behaviour in detecting misinformation and suggesting future directions.

The structure of the paper was organized as follows: section 1: contained the introduction which give the general overview of fake news detection traditional process (AI/ML Models, NLP Models, and Deep Learning models), that focuses on textual extraction and pattern recognition, section 2: present a review of related publications, problems, and proposed extending the models to read beyond text. Section 3: Described the traditional techniques used in fake new detection. Section 4: Outline the challenges of the models including fewer linguistic datasets, poor real-time detection means, poor hybrid models, and unethical model used. Section 5: the direction toward development of the high standard model was highlighted, suggest development of multimodal fake news detection system, modelling the model on learning from multi-cultural emotions, explainable AI, and the need for a real-time fake news monitoring model. Finally, Section 6: conclusion of the work.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec547

➢ *Important Acronyms and their Meaning*

Table 1 Acronyms and their Definitions

Acronyms	Definition
NIST	National Institute of Standards and Technology
VPN	Virtual Private Network
GDPR	General Data Protection Regulation
NLP	Natural Language Processing
MFA	Multi-Factor Authentication
SHAP	SHapley Additive exPlanations
TF-IDF	Term Frequency Inverse Document Frequency
FANDC	Fake News Detection in Cloud
DL	Deep Learning
Bi-LSTM	Bidirectional Long Short-Term Memory
ML	Machine Learning
XAI	Explainable AI
CNN	Convolutional Neural Network
RoBERTa	Robustly Optimised BERT Pre-training Approach
SVM	Support Vector Machine
DCNN	Dynamic Convolutional Natural Network
KNN	K-Nearest Neighbor
LSTM	Long Short-Term Memory
MLP	Multi-Layer Perception
LIME	Local Interpretable Model-agnostic Explanation

II. RELATED WORK

A numerous research in fake news detection studies were conducted, which they have clearly defined the features for categorizing and extracting the source of information to mitigate the spread misinformation A review was done using TF-IDF, FastText embedding, and Word2Vec with deep learning (CNN) models and machine learning (SVM, MLP) models. This shows TF-IDF with SVM/CNN, which achieves the highest performance using the TruthSeeker dataset [4]. Another review demonstrates the application of LSTM in detecting fake news regularisation, achieving 98% accuracy on a balanced dataset [5]. Another review revealed that a text-based model was developed, which uses (BERT + CNN + Bi-LSTM) and recorded 98% accuracy. With the help of the TruthSeeker dataset, it also makes use of a Twitter bot for real-time detection [6].

Additionally, a model was developed which emphasis on enhancing the accuracy and integrity of digital data using the Logistic Regression algorithms and Naïve Bayes. The models includes data preprocessing, feature extraction with TF-IDF, and model evaluation with confusion matrix [7]. Similarly, a comprehensive survey was conducted, which encompasses psychological theories underlying the belief in fake news and machine learning techniques for detection, utilising content, context, and hybrid features [8]. A new approach was introduced to utilise a semi-supervised method for detecting fake news by embedding a small-sized BERT and employing active learning to reduce computational demand and data requirements [9]. However, another paper on fake news related to COVID-19, defined by the World Health Organization as an 'infodemic'. An infodemic is misleading information that confuses, which causes negative impact to the health. In Turkey, high number of information on COVID-19 which were caused a lot of stress and anxiety. This led to the development of a fake news detection model, which utilizes five conventional machine learning algorithms (i.e., Naïve Bayes, Random Forest, K-Nearest Neighbor, Support Vector Machine, Logistic Regression) in the Turkish language to identify the truth of Turkish COVID-19 news on social media [10]. The model lacks a multimodal approach and is used only in the Turkish language.

Furthermore, a deep learning model was proposed using federated learning to train models for recognising fake news [11]. Analysis was carried out, which assessed the methods and datasets used, outlining the limitations, and proposed the use of transformer model integrated with artificial intelligence and human cognitive for fake news detection [12]. A literature review was conducted on a deep learning approach for fake news detection, accompanied by a survey that emphasises the use of BERT, LSTM, and CNN in mitigating misinformation [13]. To overcome the challenges faced in detecting misinformation on a document, a model was developed that uses five (5) different datasets embedded, which focus on complex architectures in fake news detection [14]. A Neural Network was introduced that uses three layers of information (words, sentences, and headlines) to determine the language approach, which helps achieve high accuracy in detecting fake news, with a 6.77% success rate [15]. Similarly, an adaptive model was developed to distinguish between short and fake news, which utilises Capsule Networks in conjunction with DCNN and MLP to predict accuracy based on the input size [16].

A model with hybrid feature was developed, incorporating natural language processing (NLP), and machine learning (ML), deep learning (DL) techniques for detecting misinformation. The model provide processes of

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec547

identify misinformation in a large volume of text. The suggested model reads, learns, reduces, and improves the accuracy of the data using metadata, TF-IDF, and ML ensemble [17]. To enhance the use of social media, a sentiment analysis model and an emotional model were developed, which learn from users' comments and reactions to predict the accuracy of information using a bidirectional long short-term memory model (BLSTM) [18]. A review of compiled content-based, network, and linguistic methods was conducted, highlighting datasets such as BuzzFeed and LIAR and comparing feature extraction techniques (TF-IDF, Word2Vec) and classifiers (SVM, LSTM, KNN) [19]. A hierarchical approach was presented, which utilises dual attention on documents and words to enhance the model's ability to detect evidence of fake news [20]. Transformer models with Bayesian regularisation were assembled to provide a model that detects fake news on a dataset, achieving an F1 score of 98.92% [21]. A linguistic feature-based model was proposed to utilise semantics, readability, and syntax with a sequential neural network to classify news with 86% accuracy [22]. Another paper studied the use of social media in misleading the Turkish community regarding vaccines, which led to much unverified content being shared online, and the country is facing a future health crisis. To mitigate this, a model was developed using machine learning methods, including logistic regression, XGBoost and support vector machines were the tools used in identifying the falsified information against the anti-vaccine sentiments on social media post in Turkey. The model utilizes a Turkish media dataset, comprising 3778 verified posts, and the results show that Transformers can separate Turkish social media posts with anti-vaccine beliefs from other posts with a 75.9% area under the ROC curve rate [23].

Existing models for detecting fake news offer a variety of methods to identify misinformation, but lack a comprehensive approach that extends beyond the text to evaluate the accuracy of news sources, while also maintaining ethics and clarity in making decision regarding the validity of the fake news source. The proposed paper approach aim to look into beyond textual models by using multimodal systems to integrate images, audio, video and emotions for determining the validity of fake news, featuring the needs of multi-cultural emotions methods to handle the linguistics diversity globally, emphasise the need for real-time detectors for maintaining safety of the use of digital media, and explainable AI to (XAI) to for transparency, explain the decision techniques in validating the source, and introduces the use of differential privacy federated learning to protect privacy of the users while modelling the behavior of the news.

In summary, based on the existing literature review, the research was designed to comprehensively highlight the need for automating a fake news detection system by extending its capabilities beyond text analysis to enhance performance, accuracy, openness, user data protection, and improve prediction behavior.

III. TRADITIONAL FAKE NEWS DETECTION TECHNIQUES

Traditional news detection models were developed based on Natural Language Processing (NLP) models, Artificial Intelligence and Machine Learning Models, Deep Learning Models, and, most recently, Transformer-based models. The models focuses exclusively on detecting language structure, such as sentence composition, phrases, words, and linguistic patterns, to identify fake news.

➢ *Natural Processing Language (NPL)*

Natural language processing (NLP) can be described as a branch of computer science that utilises artificial intelligence to analyse and process human languages. It helps machines to read, analyse, interpret, and comprehend human speech [24].

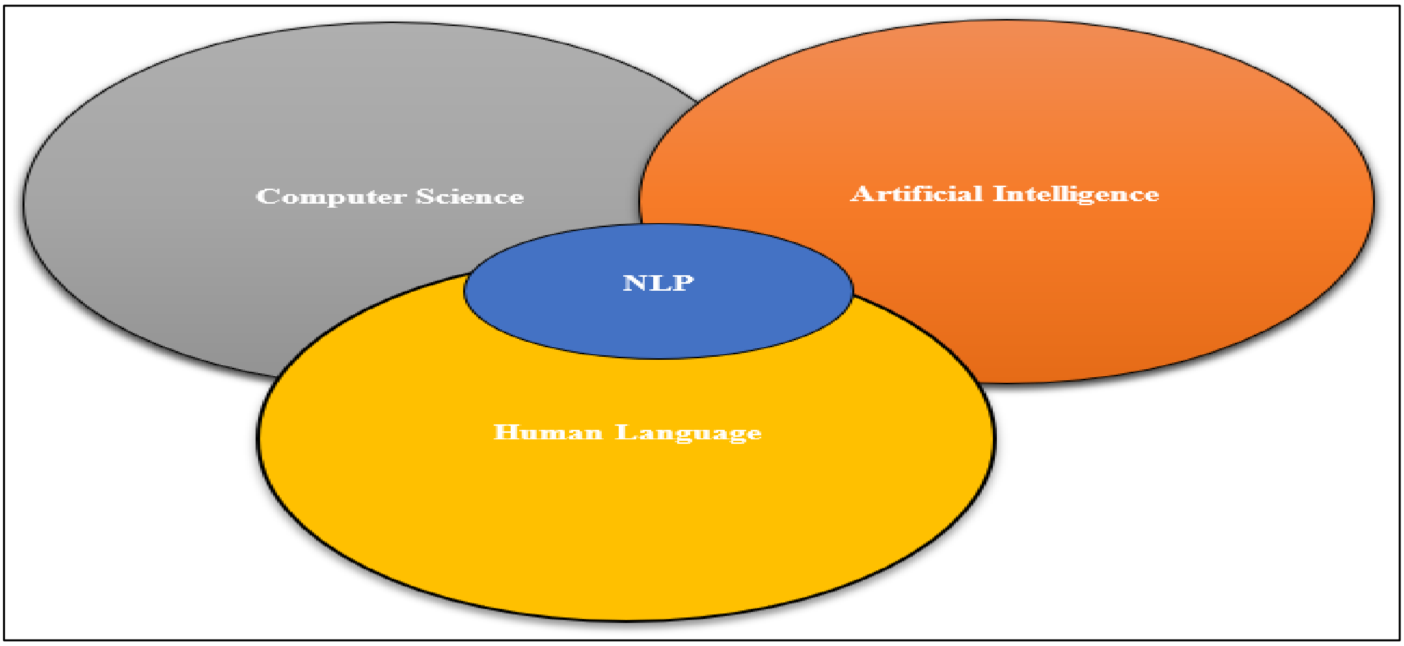

Fig 2 Natural Processing Language

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec547

The processes involved in natural language processing include text input (Preprocessing), Text Representation, Language Modelling and Understanding, NLP Tasks, Post-Processing, and a Feedback loop.

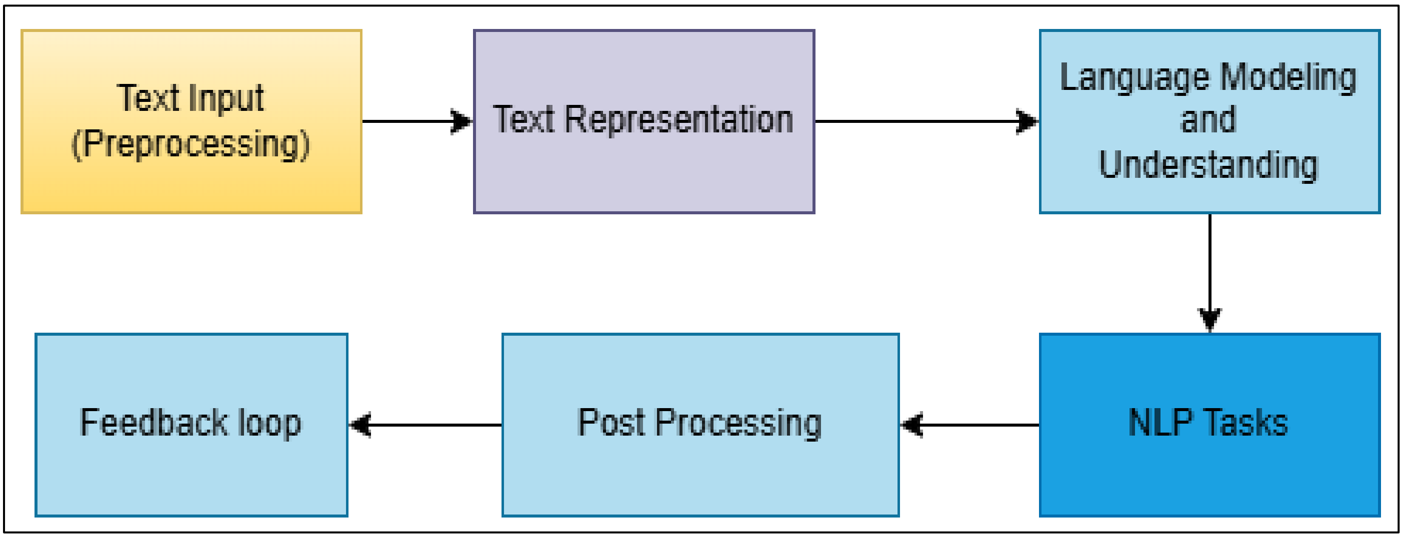

Fig 3 NLP Processes

These processes involve gathering, analyzing, and extracting the information required for a machine to understand the direction of human language, which is later integrated into different AI Models to learn and automate the process (e.g., Fake News Detection Models).

- *Machine Learning and Artificial Intelligence Models*

Machine Learning and Artificial Intelligence models, such as K-Nearest Neighbour (KNN), logistic regression, and Support Vector Machines (SVM), were developed and used in fake news detection. These models are used in text extraction from the structure of sentences, words, and linguistic approaches. Based on these approaches, the models achieved a performance of approximately 99% in terms of precision and accuracy. These models utilize classifiers to identify outliers within content and enable the machine to predict the result based on the extracted text.

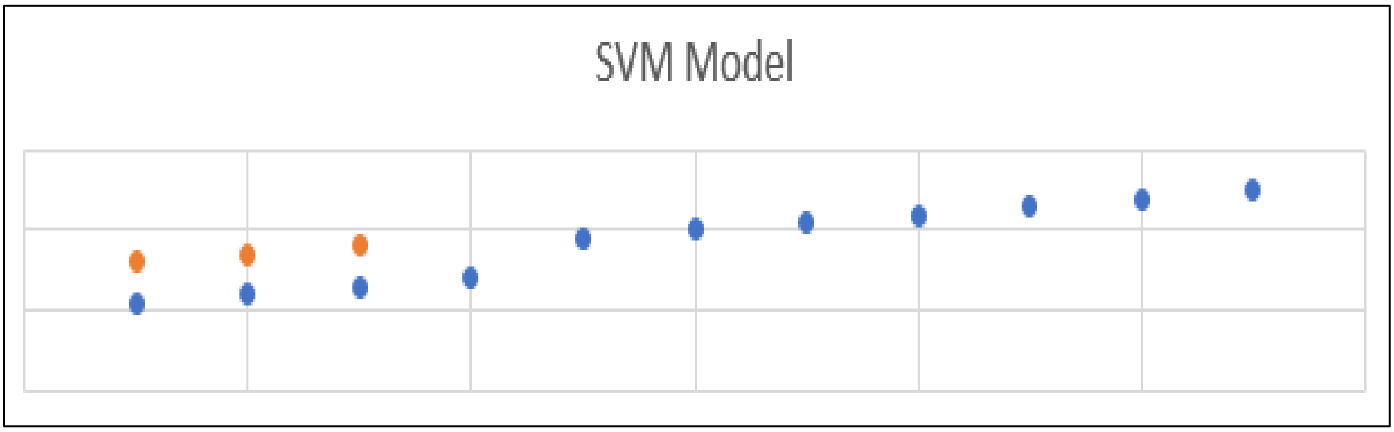

Fig 4 SVM Model

In the above model prediction, the information was categorised based on the outlier and the model accuracy, which generally reveal the processes on the model works to extract false news.

- *Deep Learning Models*

In deep learning models, Convolutional Neural Networks (CNN) with one and multiple convolution layers, Feedforward Neural Networks (FFNN), and Long Short-Term Memory (LSTM) Networks were among the different models developed to extract misinformation. These models were reported to provide 97.5% accuracy, recall, F1-score, and precision. Specifically, CNNs are highly effective for full-text analysis, and LSTM models are particularly suited for sequential data. All the models employ a textual form approach to detect the integrity of the provided data [25].

- *Transformer-Based Models*

Due to limitations in deep learning models and the vast amount of fake news, the transformer model was introduced to read, analyse, learn, and detect misinformation. Models like BERT, developed by Google, and RoBERTa, developed

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec547

by Facebook, use preprocessing steps that include tokenisation, stop word removal, and vectorisation to read the entire content of news before detecting falsified information [26].

All these approaches utilise textual and linguistic patterns learned by the models to predict the accuracy of the information. However, they fail to detect nuanced manipulations in the pattern, such as contextual cues analysis, user reactions, and sentiment analysis within the dataset.

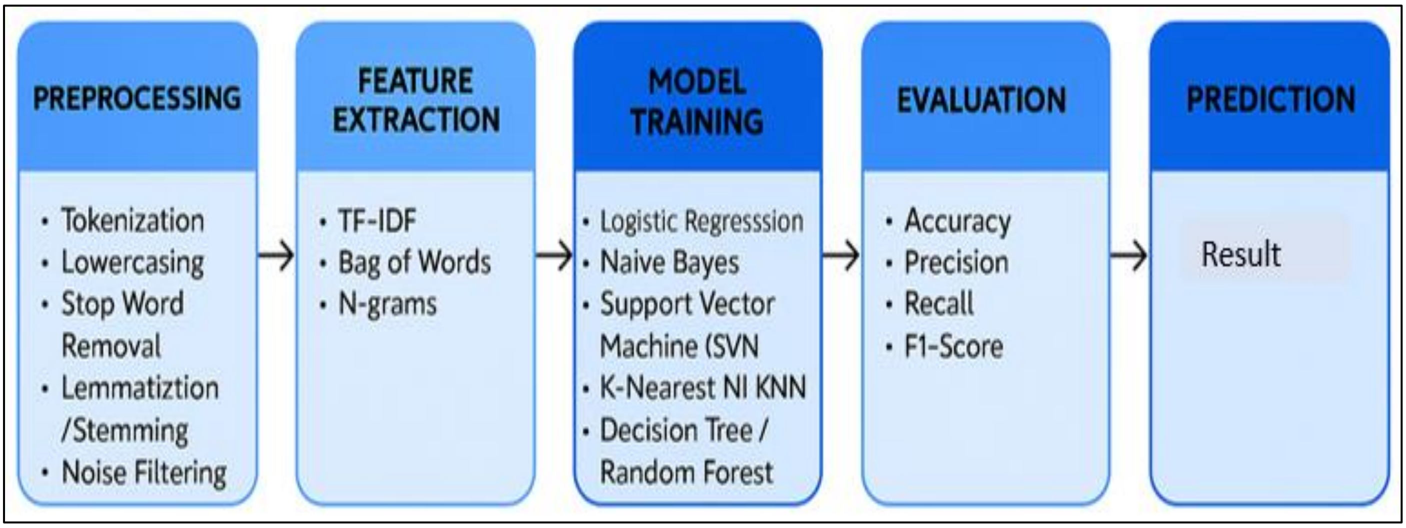

Fig 5 Method of Fake News Prediction

Table 2 Comparison Table Between the Techniques

Method	Definition	Advantages	Disadvantages
Natural Language Processing (NLP)	Uses statistical and language patterns to analyse the structure of text.	High standard in detecting language patterns and grammatical structure	Lack of contextual cues, emotions, and some features in the textual approach
Machine Learning (ML)	Uses well-defined algorithms like SVM, Logistic Regression, and Naïve Bayes to classify the validity of the news.	The labelled data accuracy is high and very fast in data classification	Requires a large data set, which can be complicated to understand due to the complex and nuanced data
Deep Learning (DL)	Uses Bi-LSTM, FFNN, DCNN, and CNN models in detecting text patterns.	Handled complex and unstructured data and provided high performance on sequential data	Slow in training the model, and a lack of transparency
Transformer-based Models	Uses pre-trained models like RoBERTa and BERT to go deeper in detecting text patterns.	Good at understanding contextual data and predicting with high accuracy	Limited in explaining the decision approach.

IV. KEY CHALLENGES OF THE STUDY

Machine learning, Artificial Intelligence, Deep Learning, and Transformer-based models play a vital role in mitigating the spread of misinformation; however, several limitations and restrictions have been identified that hinder the effectiveness and applications of these models. The following are key gaps outlined in the research challenges.

➢ *Fewer Emotional Datasets in the Existing Models.*

Despite the positive outcome in the effectiveness present in detecting fake news from the models, they were mostly limited to text-based approaches, which were the techniques on which the models were trained. The absence of emotions, contextual cues, and user behaviour analysis limits the models' ability to learn from users, hindering their ability to predict the implications of interpretability and accuracy, and to detect fake news [27]. The models were built to read from sub-categories of the content (usually, the heading, keywords, and linguistic approach) in determining the fake news. The technique has sparked a point of discussion among content creators, who utilize identifiable patterns and processes to outsmart a machine's intelligence and ensure readers' reliability.

➢ *Lack of Integration of Multilingual Signals*

The limited availability of low-resource languages in most detection models and datasets hinders the detection of fake news in multilingual contexts. High-resource languages, such as Bengali, Tamil, and Hindi in India, lack adequate interpretational data and tools for natural language processing, which other high-resource languages benefited from high-rated multimodal models. It helps in facilitating the circulation of false information in these languages, particularly through the use of emotive imagery and deceptive language. The majority of detection methods only

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec547

use text, ignoring code-mixed content or multimodal manipulations that are typical of Indic languages [28].

- *Ethical Concerns in Modelling User Behaviour*

Implementing user behavior in fake news models raised ethical concerns due to the involvement of data related to user profiling, data privacy, and surveillance. Comments, likes, and engagement patterns (behavioural cues) are rarely identified as important aspects in improving model detection performance, as their use requires adherence to data protection laws and ethical standards. This creates a significant gap in developing models with a profiling nature that balances the user's privacy protection with the effectiveness of the models in detecting fake news [29].

- *Poor of Integration of Real-Time Models in Detecting Misinformation on Digital Media Platforms*

Circulation of fake news has become a major issue worldwide, and researchers have explored various models using English-language datasets. This restricts the potential of multilingual and real-time emotion detection models, making them useless in other regions, particularly Asia, Africa, and Latin America. In addition, the models' responsiveness to crises has reduced due to the underdevelopment of real-time emotion detection in monitoring the spread on social media platforms. Additionally, the challenges that may arise during detection include processing a vast amount of data and ensuring a processing speed that provides instant feedback, thereby mitigating the spread of misinformation [30].

- *Lack of Adaptation of Hybrid Techniques in Implementing Modelling with More User Behaviour Involvement in Detecting Misinformation*

Hybrid techniques aim to combine multiple features, such as linguistic techniques and user behaviour involvement, in a single model capable of detecting fake news through various techniques to enhance accuracy. Moreover, most existing models focus on individual approaches and metrics (which includes the number of texts, likes, and shares), textual patterns, and lack behavioral insights. Few models thoroughly integrate these signals with textual or visual content within a single architecture, despite research indicating that user behaviour (such as engagement patterns, comment sentiment, and propagation networks) significantly influences the dissemination of false information. The performance of the detection is limited by this gap, particularly on most accessible platforms like Twitter and Facebook. This hindered the advancement of providing hybrid standard models and common frameworks for the platforms.

V. FUTURE DIRECTIONS

To address these challenges, several techniques must be implemented in enhancing the identification effectiveness of fake news from various sources. This includes methods to integrate these techniques into models, as well as measures for ensuring ethical and legal data protection. The following discusses the possible gaps to be added to achieve high performance and accuracy in fake news detection models.

- *Multimodal Fake News Detection (Text, Image, Video, Emotion).*

As fake news becomes a central point of view, relying on textual techniques or a single mode of detection proves to be insufficient. Multimodal detection techniques address this limitation by providing integrated means of reading data (images, videos, audio, and text) to improve misinformation detection processes. MisD-MoF is a framework introduced that utilises expert models (BERT for text, ViT for images, and Wav2Vec for audio) to enhance fake news accuracy, providing an additional 3.71% improvement on the dataset compared to unimodal systems [31]. Similarly, a real-time cloud-based detection model called FANDC, which works primarily on social media platforms and categorises news into propaganda, hoax, etc., recorded 99% accuracy on over 99 million tweets [30].

Moreover, the need for multimodal techniques is emphasised by integrating models with more visual mechanisms, emotional patterns, and other means of deepfake detection approaches that were not covered by text-only systems. SSRN introduces the Cross-modal Attention Mechanism, which utilises both textual and visual features to enhance model performance and interpretability [32]. Using this mechanism not only provides a means of detecting deceptive methods but also offers diverse scalability across content types and platforms. As fake news is increasing, multimodal mechanisms are beneficial and need to be adopted in detection systems to reduce the complexities in modern platforms.

- *Cross-Cultural Emotion Modelling.*

Fake news detection based on emotions determines its credibility and information accuracy, as it intensely exploits the content not just from what it carries but also in how people react. To achieve this, a global dataset from different languages and ethnic groups needs to be collected, which enables the model to learn and predict accurately. However, the current dataset, although it focuses on Western-centric perspectives, limits its prediction capability due to linguistic and multi-cultural aspects. Several findings have shown that the use of emotional cues on digital platforms as an effective medium of detecting news.

For instance, the emotional reaction to Western content may differ in other regions. However, the models were not equipped with the necessary mechanisms due to the diverse cultural backgrounds and limited multilingual data for training the model [33].

Furthermore, another study emphasises how effective real-time emotional analysis is obstructed by a lack of linguistic and cultural bias in classifying emotions from the models. The paper highlights the difficulty in recognising emotional manipulations in textual or visual content when training models on homogeneous cultural datasets [34]. This led to failure, particularly on social media platforms, in training models to recognise multi-cultural emotional reactions and determine accuracy.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec547

Integrating multi-cultural emotions is, therefore, important. It defines the means of working with emotionally rich datasets in social contexts and multiple languages to develop models for fake news that are adaptable to diverse emotional expressions. Adding this to the model will increase the robustness, performance, and equity of the systems in identifying fake news, supporting multi-cultural and linguistic deployment, which will make the world of AI competent enough to handle fake news.

- *Social Media Platforms Real-time emotion-aware detectors.*

Social media is one of the crucial media for disseminating fake news, which then gets massive traction before authenticating the source. Embedding a real-time monitoring system in the platforms enables it to validate the authenticity of the source, which in turn allows the modelling system to mitigate the global expansion of the news. A study revealed that [30] Twitter and Instagram show more content that is emotionally manipulated, and this content also generates high traffic globally, fueling the narrative across the world. Most of our social platforms lacked integration with such modules; instead, users had to submit a report primarily for content to be closed or removed. Real-time detectors will significantly and timely increase the performance of fake news detection models in reducing harm and building user trust in content and digital media platforms globally.

- *Explainable AI in Emotion/Context-Based Detection*

Detecting fake news with contextual cues and emotions by the models raises many questions, which in turn raises a concern that requires transparency in decision-making. To explain how emotions are used in determining the accuracy of fake news, considering factors such as sentiment, tone, and user interaction is highly effective but often regarded as a 'black box', making it difficult for developers, regulators, and make the users in understanding why certain content was flagged. To overcome this challenge, Explainable AI (XAI) was developed to bridge the limitation by integrating it into the models to interpret the result and justify the model's predictions.

In [33], a transformer-based system is highlighted that captures fake news by contextual patterns and emotional complexity, which often lack transparency, thereby limiting trust in adopting models for news verification. Moreover, real-time detection systems, particularly on social media, may flag a tweet, image, or video as fake but lack an explainable means to justify the classification. To overcome the challenge, incorporating Explainable AI (XAI) methods like LIME (Local Interpretable Model-agnostic Explanation) and SHAP (SHapley Additive exPlanations) provides more insight into visualizing how models work to fake news. This does not build trust but instead adds a means of regulatory compliance, aids in debugging, and bias detection.

- *Privacy-Preserving Modelling of User Behaviour (e.g., Via Federated Learning).*

User behaviour (such as comments, likes, and shares) has significant impact in shaping the world of AI in detecting, alerting, and alleviating the circulation of false information instantly. The research highlighted how information spreads by reacting to the content without verifying the source. Gathering and processing data to the model is essential, but at same time raises a concern over ethical use of the users data in which his can be accessed without his consent and accessing the data directly violates the standard of data protection and usage standards (National Institute of Standards and Technology (NIST), Nigerian Data Protection Commission (NDPC), Data Protection Board of India (DPBI).

Integrating federated learning and differential privacy techniques safeguards user privacy on these platforms. Differential privacy and federated learning are frameworks developed for machine learning that enable models to access and use a user's data in an ethical manner without revealing their identity. The framework provides the standard effectiveness and performance of the model in ensuring compliance with the rules and regulations for the use of digital data.

Differential privacy is a framework that protects the privacy of users by analyzing the datasets with noise, making it challenging to find the actual individual record used. The frameworks increase the performance of the models in detecting fake news, allowing them to be more effective and accurate in their performance while adhering to standards and maintaining data confidentiality.

- *Strengths, Ethics, and Contribution of the Future Direction Approach*

The table below highlights the advantages, ethics, and contributions toward advancing fake news detection models.

Table 3 Strengths, Ethics, and Contribution of Fake News: Future Direction

Directions	Strength	Ethics	Contribution
Multimodal Detection System	Provide additional means of detecting fake news (video, audio, text, and images)	Reduce reliance on a single form of data	Enhance semantic techniques and patterns in detecting fake content
Multi-cultural Emotions Model	Provide a worldwide standard model by adapting to multi-cultural approaches.	Enhance acceptance of diverse cross-cultural models and emotional expression.	Improve accuracy by addressing challenging linguistic contexts.
Integration of Explainable (XAI)	Promote transparency and user privacy protection trust	Provide support in understanding how models make decisions	Easy to validate the source of the news, whether it is positive/negative

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec547

Integration of privacy-preserving techniques	Enable users' data protection on decentralised data while training the model.	Uses regulations like (GDPR, NIST, FBI) in complying with privacy usage acts	Maintaining ethics and accuracy in accepting the data while predicting and using the data.
Privacy-preserving modelling of user behaviour	Enable models to preserve privacy by using the behaviour of the use	Reduces high number of false information on social media	Strengthen on real-time detection and of misinformation.

➢ *Future Research Topic*

Creative and Innovative methods should be more emphasized in the future of fake news detection, that enhance the prediction power, performance evaluation, adhering to ethical standard usage of the digital by the authorized bodies, real-time alert and prevention mechanisms, and Explainable AI (XAI) to provide highly productive and successful models in mitigating the spread of fake news globally. Some of the topics include:

- Using Deep Learning in Multimodal Fake News Detection: This topic involves creating an innovative model capable of combining text, video, images, and audio to analyse the techniques used on the source and identify the manipulations applied to the content.
- Federated Learning for Preserving Privacy in Detecting Fake News: To address ethical concerns in methodologies used to identify fake news, privacy protection models need to be developed, which will preserve users' privacy while detecting misinformation.
- A Combined Approach in Fake News Detection: Using Multimodal, Federated Learning, Explainable AI, and Real-time Detector. This topic focuses on integrating various techniques to develop a comprehensive model that promotes high performance.
- Real-time Emotionally Manipulated Fake News System on Social Media Platforms: The topic will emphasize on implementing the models that identify fake information and manipulate them accordingly by reporting to the authorities involve, automatic deletion, or alerting the users on the source credibility.
- Integrating cloud computing system with multimodal fake news detection model: Cloud computing is an adoptable technology with its dynamic nature of flexibility and usage of virtualized resources as services over the Internet [35]. Adopting cloud computing will enable fake news models to provide real-time data to detect and mitigate the spread of misinformation easily and concurrently.

VI. CONCLUSION

This review has extensively explored the gaps in misinformation detection systems and proposed a comprehensive review of applying emotional behavior and contextual cues in these systems to enhance performance and accuracy in predicting fake news on digital media and platforms. Traditional approaches, such as syntactic patterns and textual data, are insufficient for analysing and understanding the complex patterns of misinformation in the digital era. The paper highlighted that, the user engagement patterns, contextual metadata, and emotional behavior manipulation play an important role in determining the credibility of false news sources by evaluating the models, developed in machine learning, transformer-based model, and deep learning highlighted the crucial gaps which were used in addressing the challenges, the paper suggested key areas that need improvement to perfect the model accuracy and effectiveness, such as integration of multimodal systems (images, texts, audios, and videos), multi-cultural modelling to be world acceptable detection system, explainable AI (XAI) for transparency and casting trust to the users and suggest integrating user protection mechanism such federated learning and differential privacy. We applied these methodologies to offer a transformative approach to building an ethical, intelligent, and practical model for detecting fake news. Prioritising the development of such models enhances trust in using digital innovative systems through collaboration in research, data collection, and digital innovation.

ACKNOWLEDGMENT

The authors gratefully acknowledges the academic support provided by the Department of Computer Science and Engineering, Shobhit Institute of Engineering & Technology, Uttar Pradesh, India, during the course of this research.

➢ *Disclaimer:*

The views and conclusions expressed in this paper are solely those of the author and do not necessarily represent the official policies or positions of the affiliated institution.

➢ *Conflict of Interest:*

The author declares that there is no conflict of interest regarding the publication of this paper.

➢ *Informed Consent:*

This study did not involve human participants, and therefore informed consent was not required.

REFERENCES

[1]. Council of Europe, "Dealing with propaganda, misinformation and fake news - Democratic Schools for All," [Online]. Available: https://www.coe.int/en/web/campaign-free-to-speak-safe-to-learn/dealing-with-propaganda-misinformation-and-fake-news. Accessed: May 29, 2025.

[2]. European Union, "data.europa.eu," [Online]. Available:

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec547

https://data.europa.eu/data/datasets/s2183_464_eng?locale=en. Accessed: May 30, 2025.
[3]. BBC News, "A year in fake news in Africa," [Online]. Available: https://www.bbc.com/news/world-africa-46127868. Accessed: May 30, 2025.
[4]. M. A. B. Al-Tarawneh, O. Al-irr, K. S. Al-Maaitah, H. Kanj, and W. H. F. Aly, "Enhancing Fake News Detection with Word Embedding: A Machine Learning and Deep Learning Approach," Computers, vol. 13, no. 9, p. 239, Sept. 2024.
[5]. T. S. Camelia, F. R. Fahim, and M. M. Anwar, "A Regularized LSTM Method for Detecting Fake News Articles," arXiv preprint arXiv:2411.10713, 2024.
[6]. A. A. Olagunju and I. O. Awoyelu, "Performance Evaluation of Fake News Detection Models," Int. J. Inf. Technol. Comput. Sci., vol. 16, no. 6, pp. 89–100, Dec. 2024.
[7]. M. V. Rampurkar, "An Approach towards Fake News Detection using Machine Learning Techniques," unpublished.
[8]. J. Alghamdi, S. Luo, and Y. Lin, "A comprehensive survey on machine learning approaches for fake news detection," Multimed. Tools Appl., vol. 83, no. 17, pp. 51009–51067, Nov. 2023.
[9]. F. Folino, G. Folino, M. Guarascio, L. Pontieri, and P. Zicari, "Towards Data- and Compute-Efficient Fake-News Detection: An Approach Combining Active Learning and Pre-Trained Language Models," SN Comput. Sci., vol. 5, no. 5, p. 470, Apr. 2024.
[10]. M. Bozuyla and A. Özçift, "Developing a fake news identification model with advanced deep learning," Turk. J. Electr. Eng. Comput. Sci., 2024.
[11]. K. Machová, M. Mach, and V. Balara, "Federated Learning in the Detection of Fake News Using Deep Learning as a Basic Method," Sensors, vol. 24, no. 11, p. 3590, Jun. 2024.
[12]. S. Harris, H. J. Hadi, N. Ahmad, and M. A. Alshara, "Fake News Detection Revisited: An Extensive Review of Theoretical Frameworks, Dataset Assessments, Model Constraints, and Forward-Looking Research Agendas," Technologies, vol. 12, no. 11, p. 222, Nov. 2024.
[13]. M. Haqi Al-Tai, B. M. Nema, and A. Al-Sherbaz, "Deep Learning for Fake News Detection: Literature Review," Al-Mustansiriyah J. Sci., vol. 34, no. 2, pp. 70–81, Jun. 2023.
[14]. C.-O. Truică and E.-S. Apostol, "It's All in the Embedding! Fake News Detection Using Document Embeddings," Mathematics, vol. 11, no. 3, p. 508, Jan. 2023.
[15]. S. Singhania, N. Fernandez, and S. Rao, "3HAN: A Deep Neural Network for Fake News Detection," in Proc. Int. Conf., 2017, pp. 572–581.
[16]. M. H. Goldani, R. Safabakhsh, and S. Momtazi, "X-CapsNet For Fake News Detection," arXiv preprint arXiv:2307.12332, Jul. 2023.
[17]. D. P. Malik, R. Pandit, A. Chourasia, D. L. Singh, D. P. Rane, and D. Piyush, "Automated Fake News Detection: Approaches, Challenges, and Future Directions," unpublished.
[18]. S. Kh. Hamed, M. J. Ab Aziz, and M. R. Yaakub, "Fake News Detection Model on Social Media by Leveraging Sentiment Analysis of News Content and Emotion Analysis of Users' Comments," Sensors, vol. 23, no. 4, p. 1748, Feb. 2023.
[19]. T. Majeed, H. Farooq, T. Jan, and A. Nazir, "Fake News Detection: A Review," in Proc. 3rd Int. Conf. ICT for Digital, Smart, and Sustainable Development (ICIDSSD), New Delhi, India, Mar. 2022.
[20]. N. Vo and K. Lee, "Hierarchical Multi-head Attentive Network for Evidence-aware Fake News Detection," arXiv preprint arXiv:2102.02680, Feb. 2021.
[21]. S. D. Das, A. Basak, and S. Dutta, "A Heuristic-driven Uncertainty based Ensemble Framework for Fake News Detection in Tweets and News Articles," arXiv preprint arXiv:2104.01791, Dec. 2021.
[22]. A. Choudhary and A. Arora, "Linguistic feature based learning model for fake news detection and classification," Expert Syst. Appl., vol. 169, p. 114171, May 2021.
[23]. S. Sari and U. Bayram, "Vaccine hesitancy in Türkiye: A natural language processing study on social media," Turk. J. Electr. Eng. Comput. Sci., 2023.
[24]. A. M. Sanadi and V. Bhat, "A Review Paper on Natural Language Processing (NLP)," Int. J. Adv. Res. Sci. Commun. Technol., pp. 23–29, Mar. 2022.
[25]. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015.
[26]. S. F. N. Azizah, H. D. Cahyono, S. W. Sihwi, and W. Widiarto, "Performance Analysis of Transformer Based Models (BERT, ALBERT, and RoBERTa) in Fake News Detection," in Proc. 6th Int. Conf. on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, Nov. 2023, pp. 425–430.
[27]. N. Hoy and T. Koulouri, "An exploration of features to improve the generalisability of fake news detection models," Expert Syst. Appl., vol. 275, p. 126949, May 2025.
[28]. S. Bansal, N. S. Singh, S. S. Dar, and N. Kumar, "MMCFND: Multimodal Multilingual Caption-aware Fake News Detection for Low-resource Indic Languages," arXiv preprint arXiv:2410.10407, Oct. 2024.
[29]. L. Allein, M.-F. Moens, and D. Perrotta, "Preventing profiling for ethical fake news detection," Inf. Process. Manag., vol. 60, no. 2, p. 103206, Mar. 2023.
[30]. N. Cavus, M. Goksu, and B. Oktekin, "Real-time fake news detection in online social networks: FANDC Cloud-based system," Sci. Rep., vol. 14, no. 1, p. 25954, Oct. 2024.
[31]. M. Liu et al., "MisD-MoE: A Multimodal Misinformation Detection Framework with Adaptive Feature Selection," SSRN preprint, 2024.
[32]. S. S. Debata and S. Elango, "A Multimodal Attention-Driven Framework for Fake News Detection," SSRN preprint, 2025.
[33]. Z. Liu, T. Zhang, K. Yang, P. Thompson, Z. Yu, and S. Ananiadou, "Emotion Detection for Misinformation: A Review," Inf. Fusion, vol. 107, p. 102300, Jul. 2024.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/25dec547

[34]. J. M. Stella, K. Akhila, and B. N. Nishitha, "Real-Time Deepfake Detection and Emotion Analysis from Tweets, Images and Videos," Int. J. Eng. Res., vol. 14, no. 06, 2025.
[35]. F. Jamal, "Emerging Technologies and Developments in Cloud Computing: A Systematic Review," Int. J. Emerg. Trends Eng. Res., vol. 8, no. 3, pp. 894–905, Mar. 2020.