International Journal of Future Engineering Innovations

An AI-Driven Framework for Real-Time Fake News Detection: Developing a Machine Learning-Based Filter for News Platforms in the United States

Umma Twaha [1*], Yeasin Arfin [2]
[1] University of North Alabama, USA
[2] American International University, Bangladesh

* Corresponding Author: **Umma Twaha**

Article Info

ISSN (online): 3049-1215
Volume: 02
Issue: 04
July – August 2025
Received: 22-05-2025
Accepted: 30-06-2025
Published: 23-07-2025
Page No: 158-169

Abstract
The proliferation of fake news across digital platforms poses a significant threat to democratic processes, public health, and social cohesion in the United States. Manual fact-checking and traditional moderation approaches are increasingly inadequate due to the scale, speed, and sophistication of misinformation campaigns. This presents an AI-driven framework for real-time fake news detection, designed to serve as an intelligent filter for news platforms operating within the U.S. media ecosystem. Leveraging natural language processing (NLP) techniques and machine learning algorithms—including supervised classifiers and deep learning models such as BERT and LSTM—the framework identifies deceptive content with high precision and low latency. The proposed architecture integrates news ingestion pipelines, contextual feature extraction, and classification modules capable of operating on streaming data. Publicly available labeled datasets such as LIAR, FakeNewsNet, and PolitiFact were utilized for training and evaluation, ensuring robustness and generalizability. The framework also includes a dynamic feedback loop to continuously improve performance through human-in-the-loop validation and real-time user engagement data. A systematic literature review guided by PRISMA methodology was conducted to inform model selection, dataset validation, and deployment strategies. Experimental results demonstrate that hybrid models combining linguistic, semantic, and social context features achieve superior performance over traditional baselines. Ethical, legal, and societal considerations—including transparency, free speech implications, and algorithmic fairness—are also addressed to ensure responsible deployment. By enabling scalable, automated, and explainable fake news detection, this framework offers a practical tool for news organizations, technology platforms, and regulatory bodies. The research highlights the importance of interdisciplinary collaboration and continuous model updating to combat evolving misinformation tactics. The proposed AI-driven system represents a significant step toward safeguarding information integrity in the U.S. digital media landscape through real-time, intelligent intervention.

DOI: https://doi.org/10.54660/IJFEI.2025.2.4.158-169

Keywords: AI-Driven, Framework, Real-Time, Fake News Detection, Machine Learning, Based Filter, News Platforms, United States

1. Introduction

The proliferation of misinformation and fake news has emerged as a significant societal challenge in the digital era, particularly in the United States. The rapid expansion of online news platforms and social media networks has transformed how information is produced, disseminated, and consumed (Lei *et al*., 2019; Al-Quran, 2022). However, this transformation has also made it easier for misleading or false information to spread widely and rapidly. The impact of misinformation has been profoundly felt during critical periods such as U.S. presidential elections, where fabricated stories have distorted public opinion, and during the COVID-19 pandemic, where misleading health information undermined public safety measures and vaccine uptake (Erikson and Tedin, 2019; Prochaska *et al*., 2023). The consequences of fake news extend beyond individual belief systems; they threaten democratic institutions, public health infrastructure, and social cohesion (Beauvais, 2022; Reglitz, 2022).

Traditional approaches to combating misinformation, including manual moderation, journalistic oversight, and third-party fact-checking, have proven insufficient in the face of the volume and speed of online content (Cavaliere, 2020; Ünver, 2023). Manual fact-checking is labor-intensive, slow, and often reactive—addressing falsehoods after they have already gone viral (Aniebonam, 2024; Oni and Iloeje, 2025). Furthermore, human biases and inconsistencies in judgment may limit the reliability of such efforts. Given the dynamic nature of online discourse, there is a pressing need for automated solutions that can analyze and respond to misinformation in real time, ideally before it gains significant traction (Demartini *et al*., 2020; Santos, 2023).

This addresses the urgent need for scalable, accurate, and timely detection mechanisms for fake news (Folorunso *et al*., 2024; Ebepu *et al*., 2025). The current landscape of fake news detection is fragmented, with various models and heuristics lacking interoperability, domain specificity, or real-time processing capabilities (Nwabekee *et al*., 2021; Ebepu *et al*., 2024). Moreover, most existing systems struggle with adaptability, making them less effective when applied across different cultural or political contexts—particularly in the diverse and complex U.S. media ecosystem. A robust solution must be capable of integrating multiple data streams, assessing content from both linguistic and contextual perspectives, and making reliable classification decisions without extensive human intervention (Kolajo *et al*., 2019; Mehmood and Anees, 2020).

The primary objective of this research is to design and implement an AI-driven framework for real-time fake news detection that can be deployed within U.S.-based news platforms. By leveraging advances in machine learning, particularly natural language processing (NLP), neural networks, and knowledge graph techniques, the framework aims to offer an automated, scalable alternative to manual moderation. The proposed system will focus on dynamic pattern recognition, credibility scoring, and contextual fact verification, ensuring high precision in classifying content as legitimate or deceptive. Importantly, the framework will be designed with adaptability in mind, allowing it to evolve with shifting narrative patterns, new misinformation tactics, and changes in user behavior (Aniebonam *et al*., 2022; Ngoc *et al*., 2025).

This research is motivated by the growing societal threat posed by misinformation and the limitations of existing mitigation strategies. It seeks to bridge the gap between technological capability and practical implementation by developing a machine learning-based filter tailored to the unique needs of the U.S. media landscape (Nwabekee *et al*., 2021; Rathore *et al*., 2025). This framework holds the potential to enhance the integrity of information ecosystems and support more informed civic engagement.

2. Methodology

The PRISMA methodology was employed to conduct a systematic review of literature relevant to the development of an AI-driven framework for real-time fake news detection on U.S.-based news platforms. The review process began with a comprehensive identification phase, where multiple electronic databases including IEEE Xplore, ACM Digital Library, Scopus, Web of Science, and Google Scholar were queried using keywords such as "fake news detection," "AI in misinformation," "machine learning filter," "real-time fake news," and "news authenticity algorithms." The search was limited to peer-reviewed articles, conference proceedings, and high-impact industry white papers published between 2015 and 2024 to ensure relevance to contemporary advancements.

During the screening phase, 1,282 records were initially retrieved. Duplicates were removed, and 1,030 unique articles were subjected to title and abstract screening. Of these, 642 articles were excluded due to irrelevance, lack of AI methodology, or non-English language. The remaining 388 full-text articles were assessed for eligibility based on predefined inclusion criteria: application of machine learning or artificial intelligence in fake news detection, real-time or near-real-time system capabilities, and focus on U.S.-based or generalizable digital media contexts. Articles were excluded if they only addressed social media sentiment analysis, non-automated fact-checking approaches, or non-news content such as user reviews or forum discussions.

A total of 122 studies met the inclusion criteria and were included in the qualitative synthesis. Among these, 49 studies featured real-time implementation components, while 73 focused on model training, evaluation, and architecture applicable to real-time systems. Various methodologies emerged, including supervised machine learning algorithms like Random Forest and SVM, deep learning models such as LSTM and BERT, hybrid ensemble techniques, and NLP-based misinformation classifiers. These studies also introduced multiple public datasets including LIAR, FakeNewsNet, and BuzzFeedNews to benchmark detection accuracy and performance.

Data extraction was guided by a structured protocol capturing key aspects such as model type, dataset characteristics, performance metrics, and implementation context. The synthesis of findings revealed that hybrid models leveraging deep contextual embeddings (e.g., BERT) combined with temporal content analysis and user network features offer superior accuracy and real-time classification capability. However, challenges including adversarial content evasion, computational latency, and dataset bias were consistently reported.

The PRISMA process enabled a transparent and reproducible pathway to select high-quality literature underpinning the framework design. The results informed the technical specifications, model architecture, and deployment considerations for a real-time, AI-driven fake news detection filter tailored to the operational dynamics and legal frameworks of U.S. digital news platforms.

2.1 Literature Review

The increasing volume and sophistication of fake news present a critical challenge to digital information integrity. In the United States, where free speech protections intersect with a highly polarized media environment, misinformation has had tangible impacts on public opinion, electoral outcomes, and crisis response (Goldman and Baker, 2019; Hersh and Krupnikov, 2023). Understanding the typologies of fake news and evaluating existing detection approaches is essential to framing an effective AI-driven response.

Fake news encompasses a broad spectrum of deceptive content, each with distinct characteristics and propagation dynamics. Clickbait represents a prevalent typology designed to capture attention through sensational or misleading headlines, often with minimal or distorted factual content (Cherish *et al*., 2025; Nwankwo *et al*., 2025). Although not always overtly false, clickbait compromises content credibility and erodes trust in digital news. Propaganda refers to content disseminated with the intent to influence political or ideological outcomes, often leveraging selective truths or emotional appeals. This form is typically state-sponsored or organizationally coordinated, complicating its detection due to partial factuality. Satire, while intentionally false, serves a comedic or critical function and is not inherently deceptive in intent; however, it is frequently misinterpreted when shared out of context. Conspiracy theories rely on speculative, unsupported narratives that purport hidden motives or events, and their viral nature stems from emotional and ideological resonance (Cover *et al*., 2022; Birchall and Knight, 2022). Fabricated content, on the other hand, is entirely false and created with the deliberate intent to deceive, often mimicking the style of legitimate journalism to bypass initial skepticism.

Current detection techniques predominantly rely on natural language processing (NLP), external fact-checking services, and network-based analyses. NLP-based models assess linguistic and stylistic features to identify deception cues. Early techniques used bag-of-words, TF-IDF, and n-gram models, while more recent approaches incorporate deep learning architectures such as LSTM, CNNs, and transformers like BERT for contextual understanding (Ogundipe *et al*., 2023; Babalola *et al*., 2024). These models can capture subtle semantic inconsistencies, rhetorical strategies, and sentiment features indicative of misinformation. Fact-checking APIs (e.g., ClaimBuster, Google Fact Check Tools) provide automated verification by cross-referencing claims with verified databases. Although useful, such systems are limited by the timeliness and coverage of underlying databases. Network-based detection approaches, on the other hand, analyze how information spreads across social networks (Zhou and Zafarani, 2019; Rath *et al*., 2022). These methods utilize propagation patterns, user credibility scores, and community detection algorithms to identify misinformation clusters and sources.

Despite the evolution of these techniques, existing approaches face several limitations that constrain their real-world efficacy. One of the foremost challenges is delayed detection. Many systems operate on batch-mode processing or depend on manually verified datasets, limiting their responsiveness to emerging misinformation (Ogundipe *et al*., 2019; Oni, 2025). This latency undermines their usefulness in time-sensitive contexts such as elections or public health crises. Additionally, high false positive rates are common, particularly when models overfit to linguistic patterns without understanding intent or satire. This not only reduces trust in automated filters but also risks censoring legitimate discourse. Perhaps most critically, many models lack contextual understanding. Deceptive content often requires analysis of historical claims, multimedia content, user intent, and evolving discourse—all of which demand more than surface-level linguistic analysis (Guo *et al*., 2020; Damstra *et al*., 2021). Moreover, biases in training data can lead to skewed model behavior, particularly when datasets underrepresent certain media sources or linguistic styles.

The current literature reveals a dynamic and rapidly advancing field, but one still grappling with fundamental challenges in real-time performance, accuracy, and generalizability (Halliday, 2021; Onibokun *et al*., 2022). There is a growing consensus that hybrid models combining NLP with temporal, social, and credibility-based signals may offer a more robust solution. However, achieving this requires overcoming data limitations, improving contextual reasoning, and embedding ethical considerations in model design. The proposed AI-driven framework seeks to address these gaps by integrating state-of-the-art NLP, real-time stream processing, and feedback-driven learning mechanisms tailored to the complexities of the U.S. news ecosystem (Radanliev *et al*., 2023; Rajest *et al*., 2023).

2.2 Conceptual Framework

The conceptual framework for an AI-driven system for real-time fake news detection is centered around a modular, scalable, and adaptive architecture designed to operate within the dynamic and high-throughput environment of digital news distribution. This framework leverages advancements in natural language processing (NLP), machine learning (ML), and stream processing to enable the automated classification of news content into truthfulness categories such as “likely true,” “misleading,” or “fake,” along with a confidence score. The framework consists of three core architectural layers; input ingestion, data processing and classification, and output dissemination, with seamless integration into existing digital ecosystems via application programming interfaces (APIs) as shown in figure 1(Akanbi and Masinde, 2020; Mavrogiorgou *et al*., 2023).

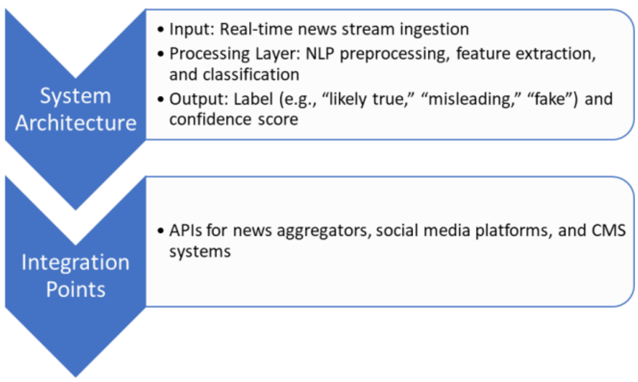

Fig 1: Conceptual Framework

At the first stage, the Input Layer manages the real-time ingestion of news content from diverse digital sources. This includes structured feeds from news APIs, RSS aggregators, and real-time crawlers parsing content from news websites and social media posts (Ogunjobi *et al*., 2024; Nwankwo *et al*., 2025). Stream processing frameworks such as Apache Kafka or Apache Flink are utilized to ensure low-latency and high-throughput ingestion of continuous data streams. The system is designed to standardize and normalize raw data, extracting key metadata such as headlines, body text, publication timestamps, and source identifiers (Manghi *et al*., 2020; Safder *et al*., 2020). This metadata is critical for downstream contextual analysis and source credibility assessments.

The core of the framework resides in the Processing Layer, which performs multi-stage transformation and classification of incoming text data. The first subcomponent is NLP-based preprocessing, which involves tokenization, stemming, lemmatization, stop-word removal, and part-of-speech tagging. These steps prepare the text for robust feature extraction. Named entity recognition (NER) and sentiment analysis are applied to capture entities, emotional tone, and subjectivity—all of which are often manipulated in fake news narratives (Nemes and Kiss, 2021; Alonso *et al*., 2021). Semantic similarity analysis and topic modeling techniques such as Latent Dirichl*et al*location (LDA) or embedding models like Word2Vec and BERT are employed to assess the topical consistency and factual coherence of the content.

Following preprocessing, the extracted features are fed into a classification engine that uses a hybrid of supervised machine learning and deep learning models (Joeaneke *et al*., 2024; Okon *et al*., 2025). Traditional classifiers such as Support Vector Machines (SVM) and Random Forests are useful for interpretable and fast classification, while neural architectures like Long Short-Term Memory (LSTM) networks and transformers (e.g., BERT) are incorporated for contextual and semantic depth. Ensemble methods, which combine the predictions of multiple models, are used to enhance generalization and reduce model variance. Each news item is ultimately assigned a probabilistic label indicating the likelihood of it being true, misleading, or fake, accompanied by a confidence score derived from softmax or sigmoid activations in the final model layer (Mishra *et al*., 2022; Wang *et al*., 2023). This allows for a nuanced representation of classification certainty, which is crucial for both user interpretation and system-level escalation protocols.

The Output Layer facilitates the delivery of classification results to external systems and end-users (Nwankwo *et al*., 2025; Obioha *et al*., 2025). Output can be directed to alert dashboards for editorial teams, real-time flagging mechanisms on consumer-facing news feeds, or backend content moderation tools. Labels and confidence scores are formatted using structured output protocols such as JSON or XML, enabling compatibility with various platforms. To ensure transparency, an explainability module leveraging model interpretation techniques like LIME or SHAP can provide a rationale for each classification, detailing key linguistic or semantic indicators that influenced the model's decision (Bhattacharya, 2022; Mesinovic *et al*., 2023).

A key innovation of this framework lies in its Integration Points with existing digital ecosystems. Through robust and secure APIs, the system can interface directly with content management systems (CMS) of major news outlets, enabling automated pre-publication content validation. Additionally, the system integrates with news aggregators such as Google News, Flipboard, and Apple News, allowing real-time misinformation tagging at scale. On social media platforms, APIs enable real-time monitoring and flagging of suspicious posts, especially in trending or high-velocity content streams. Integration with third-party fact-checking organizations can further validate system outputs and retrain models, supporting a feedback loop for continuous improvement (Juneja and Mitra, 2022; Larraz, 2023).

Overall, the proposed conceptual framework is designed to be adaptive, scalable, and modular, allowing seamless integration with diverse information channels and evolving misinformation tactics (Oyeyemi *et al*., 2024; Orenuga *et al*., 2024). By coupling advanced AI techniques with system interoperability, the framework establishes a foundation for proactive, transparent, and real-time fake news mitigation tailored to the complexities of the U.S. media landscape.

2.3 Data Collection and Preprocessing

An effective machine learning-based framework for real-time fake news detection is fundamentally dependent on the quality, relevance, and diversity of its input data. Given the complex nature of misinformation, which often blends partial truths with deceptive framing, the collection and preprocessing of data are critical components in ensuring model performance as shown in figure 2(Bondielli and Marcelloni, 2019; Chadwick and Stanyer, 2022). This outlines the strategies for curating reliable datasets, implementing robust text preprocessing, and engineering meaningful features to enable accurate fake news classification.

The cornerstone of any supervised machine learning approach lies in the training data. In this review, datasets encompassing both credible and non-credible news content are utilized to train and evaluate the detection framework. Notable sources include the LIAR dataset, which contains 12,836 labeled short statements from various contexts such as political debates and news interviews, annotated with six degrees of truthfulness. This dataset provides granular insight into misinformation within political narratives. Similarly, FakeNewsNet offers a comprehensive collection of news articles augmented with metadata from social media and user engagement patterns, thus allowing contextual modeling beyond textual content alone (Oyeyemi, 2023; Aniebonam *et al*., 2023).

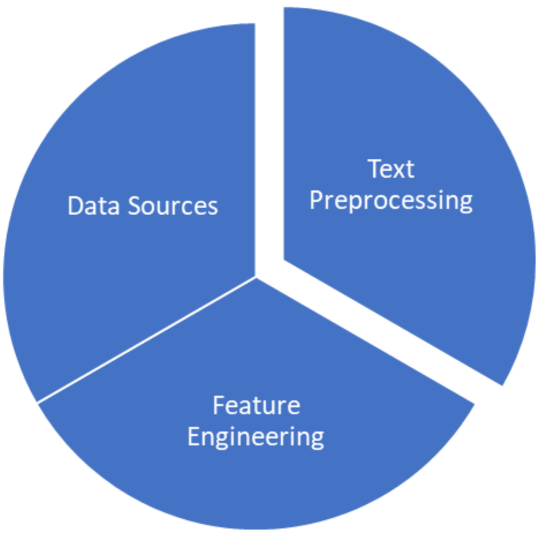

Fig 2: Data Collection and Preprocessing

Additionally, curated fact-checking repositories such as PolitiFact and Snopes serve as authoritative sources for ground-truth labeling. These platforms provide expert-verified classifications of news stories, enabling high-quality supervised learning. PolitiFact, for example, maintains a structured rating system that captures varying levels of factual accuracy, from "True" to "Pants on Fire," while Snopes includes contextual explanations that can be parsed for deeper semantic understanding. Together, these datasets contribute to a balanced representation of truthful and deceptive narratives and enable the training of generalizable models applicable across diverse content domains (Selesi-Aina *et al*., 2024; Balogun *et al*., 2025).

Once raw data is acquired, it must undergo rigorous preprocessing to enhance its suitability for machine learning models. The first step involves tokenization, where text is segmented into smaller units such as words or subwords. This process facilitates linguistic analysis and vector representation. Stop-word removal is then applied to eliminate common function words (e.g., "the," "is," "and") that do not contribute meaningful semantic content to classification tasks (Alshanik *et al*., 2020; Ladani and Desai, 2020).

Further, stemming or lemmatization is used to reduce words to their base or root forms, thereby reducing the dimensionality of the input space and improving generalization. For instance, "reporting," "reports," and "reported" would all be reduced to "report." Named entity recognition (NER) is employed to identify and classify key entities such as people, organizations, locations, and dates. This process enables the model to detect patterns associated with frequently mentioned figures in fake news and assess the plausibility of the narrative based on known relationships (Asonze *et al*., 2024; Akinola *et al*., 2024).

Preprocessing also includes normalization steps such as lowercasing, punctuation removal, and handling of hyperlinks, hashtags, and emojis—particularly relevant for social media-based datasets. Ensuring textual uniformity is critical for training stable and accurate models.

To improve classification accuracy, a diverse range of features are extracted from the preprocessed text. Lexical features include term frequency-inverse document frequency (TF-IDF), word n-grams, and character n-grams, which capture surface-level patterns and word usage frequencies (John and Oyeyemi, 2022; Olisa, 2025). These features are particularly useful in detecting clickbait phrases and repetitive structures often found in fake news articles.

Syntactic features are derived from part-of-speech (POS) tagging and parse trees, enabling the model to understand sentence structure and grammar patterns. Fake news tends to deviate from conventional journalistic syntax, making such features valuable for detection (Lugea, 2021; Verma *et al*., 2021).

Semantic features capture meaning and contextual dependencies using techniques such as word embeddings (e.g., Word2Vec, GloVe) and transformer-based language models (e.g., BERT). These methods enable the representation of text in a high-dimensional vector space where semantically similar words and sentences are positioned closely. This allows the model to understand nuanced expressions, sarcasm, and thematic inconsistencies.

Finally, network-based features derived from social context (e.g., user credibility scores, retweet networks, propagation paths) are integrated when available. These features capture the behavioral and relational dynamics of news spread, providing an additional dimension of analysis that complements text-based features.

The data collection and preprocessing pipeline for fake news detection involves the careful selection of credible and diverse datasets, advanced linguistic preprocessing, and multidimensional feature engineering (Tufchi *et al*., 2023; Hamed *et al*., 2023). These steps collectively ensure that the machine learning model is trained on representative, structured, and semantically rich data, thereby enhancing its ability to detect fake news with high accuracy in real time.

2.4 Machine learning models and techniques

The effectiveness of an AI-driven framework for real-time fake news detection largely hinges on the selection, training, and evaluation of machine learning (ML) models tailored to linguistic complexity, contextual dependencies, and temporal dynamics of digital misinformation (Oyeyemi, 2022; Bako *et al*., 2025). To ensure accurate, robust, and real-time classification, a variety of ML models and training-validation strategies are employed, with rigorous evaluation metrics guiding model refinement and deployment readiness.

Model Selection plays a pivotal role in addressing the unique challenges posed by fake news detection. Traditional machine learning algorithms such as Support Vector Machines (SVM) and Random Forests (RF) have been widely used due to their interpretability and computational efficiency. SVMs excel in high-dimensional feature spaces and perform well with sparse representations such as TF-IDF vectors, making them suitable for early-stage content classification tasks (Peng *et al*., 2021; Sinha *et al*., 2023). Random Forests, with their ensemble of decision trees, offer resilience to overfitting and can capture non-linear relationships between linguistic features. However, these models are limited in their ability to capture temporal dependencies and semantic context, which are crucial in understanding deceptive narratives.

To overcome these limitations, deep learning techniques are increasingly integrated into fake news detection frameworks. Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), are particularly effective for processing sequences of words, enabling the model to learn dependencies across entire news statements or articles (Joeaneke *et al*., 2024; Obioha *et al*., 2025). LSTMs are capable of modeling syntactic and semantic patterns that span multiple clauses, which is especially useful for detecting subtleties in fabricated content. More recently, transformer-based models like BERT (Bidirectional Encoder Representations from Transformers) have demonstrated superior performance in capturing bidirectional context, nuanced sentiment, and implicit meaning. Pre-trained on vast corpora, BERT can be fine-tuned with fake news datasets to achieve high classification accuracy with relatively small task-specific data (Aggarwal *et al*., 2020; Singh *et al*., 2023).

To capitalize on the strengths of both traditional and deep learning approaches, hybrid models are increasingly favored. These architectures often combine the precision of traditional classifiers with the contextual depth of neural embeddings. For instance, a hybrid model might use BERT to encode a news article into dense semantic vectors and then apply an SVM or Gradient Boosting Machine (GBM) to classify the output. Such fusion techniques allow for better generalization and reduce model variance, especially in imbalanced or noisy datasets.

Training and validation procedures are essential for ensuring that models generalize well to unseen data. In supervised learning, datasets are typically divided into training, validation, and test subsets, often using an 80-10-10 or 70-15-15 split. To enhance robustness and mitigate overfitting, k-fold cross-validation is employed, where the dataset is divided into *k* subsets (commonly k = 5 or 10), and the model is trained and validated *k* times, each time using a different subset as the validation set and the remaining as the training set. This method provides a comprehensive assessment of model performance and reduces sensitivity to data partitioning. For models involving time-sequenced data, stratified or time-based splitting is adopted to preserve chronological integrity and prevent information leakage.

Evaluating the performance of fake news classifiers requires a multi-metric approach due to the asymmetrical cost of errors. Precision, defined as the ratio of true positives to all predicted positives, measures the model's ability to avoid false positives—crucial in contexts where mislabeling credible news as fake can damage trust. Recall, the ratio of true positives to all actual positives, reflects the model's capacity to detect all instances of fake news, minimizing false negatives. F1-score, the harmonic mean of precision and recall, provides a balanced measure that is particularly useful when the dataset is imbalanced, which is often the case in real-world misinformation datasets. Additionally, the Receiver Operating Characteristic – Area Under Curve (ROC-AUC) quantifies the model's ability to distinguish between classes across all classification thresholds (Bowers and Zhou, 2019; Carrington *et al*., 2022). A high ROC-AUC indicates strong discriminative power, even when the decision boundary is uncertain.

Together, these machine learning models, training strategies, and evaluation metrics form a comprehensive toolkit for building reliable, real-time fake news detection systems. By combining traditional and modern approaches and rigorously validating model outputs, the AI-driven framework can effectively navigate the complex, evolving landscape of digital misinformation, particularly within the high-volume and high-stakes environment of U.S. news platforms.

2.5 Real-Time Detection Pipeline

A core requirement of an effective AI-driven fake news detection framework is the ability to classify and respond to misinformation as it emerges across digital platforms. To achieve this, a real-time detection pipeline must be established, capable of ingesting, processing, and analyzing high-velocity news content with minimal latency (Alam *et al*., 2020; Dubuc *et al*., 2020). This outlines the architectural design and technological components of such a pipeline, focusing on stream processing frameworks, latency and scalability optimization, and adaptive feedback mechanisms to ensure continuous learning and improvement.

The backbone of the real-time detection system relies on robust stream processing frameworks that support continuous data ingestion and computation. Apache Kafka serves as a distributed event streaming platform that ingests incoming news articles and social media content in real time. Kafka's architecture, based on publish-subscribe mechanisms and message queues, allows multiple data producers (e.g., RSS feeds, news APIs, social media streams) to send data into a centralized processing pipeline. The high throughput and fault-tolerance capabilities of Kafka make it particularly suitable for handling large volumes of real-time content.

Once the data is ingested, Apache Spark Streaming or Apache Flink is employed for real-time processing. These frameworks support in-memory computation and micro-batch or event-driven processing paradigms, enabling timely classification of data as fake or credible. Spark Streaming, in particular, integrates seamlessly with Kafka and provides native support for complex machine learning workflows through its MLlib library (Chen *et al*., 2023; Sinha, 2023). By executing trained models in a distributed manner, Spark enables efficient feature extraction, prediction, and metadata logging at scale.

Low latency is paramount in fake news detection, as delays in identification allow misinformation to spread and amplify its influence. To minimize latency, the pipeline is optimized at multiple levels. First, pre-trained machine learning models, such as fine-tuned BERT variants, are deployed using optimized runtime environments like ONNX Runtime or TensorRT, which accelerate inference through hardware-level optimization. These models are loaded into memory within Spark executors or deployed via REST APIs using lightweight frameworks such as FastAPI.

Second, parallel processing and load balancing are implemented to ensure high throughput under variable load conditions. Partitioning input streams across processing nodes and leveraging GPU acceleration for NLP inference help reduce bottlenecks. Additionally, message buffering and backpressure mechanisms inherent in Kafka and Spark prevent system overload during data surges, ensuring consistent performance.

Scalability is achieved through container orchestration platforms such as Kubernetes, which dynamically allocate computing resources based on processing demands. Horizontal scaling ensures that the pipeline can adapt to fluctuating content volumes, particularly during major news events or crises when the risk of fake news proliferation is elevated. Moreover, cloud-based deployments (e.g., on AWS or GCP) provide elastic compute capabilities and integrate with monitoring tools to ensure service reliability.

A critical aspect of maintaining long-term accuracy in real-time fake news detection is incorporating a feedback mechanism for continuous learning. After initial classification, news items can be flagged for further verification by trusted fact-checkers or crowd-sourced review systems. Once verified, the outcomes are stored in a feedback repository and used to retrain or fine-tune the underlying models.

This feedback loop supports two key functions. First, it enables model recalibration in response to evolving misinformation tactics, linguistic patterns, or emerging topics. Second, it supports semi-supervised learning, where uncertain or borderline classifications are reviewed and used to augment the labeled dataset. This incremental learning approach helps reduce model drift and ensures relevance over time.

Active learning techniques, such as uncertainty sampling or ensemble disagreement, can be employed to prioritize the most informative samples for human verification (Chitta *et al*., 2021; Nguyen *et al*., 2022). Periodic retraining using updated data ensures the system adapts to novel narratives and evolving discourse.

The real-time detection pipeline integrates advanced stream processing tools, low-latency and scalable system architecture, and adaptive feedback mechanisms to provide a comprehensive solution for dynamic fake news classification. By combining technical efficiency with adaptive learning, the pipeline enhances the capability of U.S.-based news platforms to counter misinformation in an accurate, scalable, and timely manner (Weidinger *et al*., 2021; Hendrix and Morozoff, 2022).

2.6 Deployment and Platform Integration

The practical deployment of an AI-driven framework for real-time fake news detection demands thoughtful integration with user platforms, careful consideration of ethical and legal principles, and collaborative partnerships to ensure system credibility and effectiveness. Beyond technical performance, the long-term impact of such a system depends on its usability, transparency, and alignment with the sociopolitical landscape of news dissemination in the United States.

User Interface Design plays a pivotal role in facilitating user trust, enabling actionability, and integrating seamlessly into the existing digital media ecosystem (Aksoy, 2023; Evans and Agoro, 2023). Effective design should prioritize intuitive interaction, real-time feedback, and minimal disruption to user experience. One of the primary UI components is the alert dashboard, which serves as the central control interface for news editors, fact-checkers, and platform moderators. These dashboards can present real-time analytics, including classification outcomes, confidence scores, metadata (e.g., source credibility, content timestamp), and model explanation outputs using techniques like LIME or SHAP. Interactive filtering options allow users to examine flagged content by topic, platform, or risk level.

For end-users, integration through browser plug-ins offers a non-intrusive method to deliver classification insights during content consumption. These plug-ins can highlight questionable headlines, display reliability scores, and suggest fact-checked alternatives, all while preserving content visibility and user autonomy. Similarly, mobile app widgets can provide real-time misinformation alerts for trending news, personalized to user preferences or viewed content history. To enhance user acceptance, these interfaces must maintain transparency, allow opt-in customization, and avoid excessive notifications that may trigger alert fatigue or resistance (Karegar *et al*., 2020; Murmann, P. and Karegar, 2021).

Ethical and legal considerations are central to deploying fake news detection systems, particularly within the U.S., where First Amendment protections enshrine freedom of expression. A delicate balance must be struck between mitigating the harms of misinformation and preserving the right to dissent and free discourse. The system must be designed not to censor content, but to label it clearly and provide supporting evidence or references. This labeling approach—akin to a "nutrition label" for news—preserves access to all content while empowering users to make informed judgments.

Transparency in algorithmic decisions is essential. Users and stakeholders must understand the basis of content labeling, especially when classifications have implications for visibility or monetization. This calls for interpretable AI components and published documentation on data sources, feature extraction techniques, and model update cycles. Moreover, continuous auditing is needed to detect and mitigate algorithmic bias, particularly toward underrepresented communities or ideological viewpoints.

pa

User appeals mechanisms should also be embedded within the platform, allowing content creators to contest or clarify classifications.

Legal compliance must also address data privacy, particularly under frameworks like the California Consumer Privacy Act (CCPA) and General Data Protection Regulation (GDPR) for international users. Systems must avoid storing personally identifiable information (PII) unless strictly necessary and ensure encrypted transmission of data between the user device and detection servers (Casaleiro, 2020; Akpan *et al*., 2022).

To achieve both scale and legitimacy, the deployment strategy must be grounded in robust partnership models. Collaboration with media houses is vital for embedding the system directly into editorial workflows, enabling pre-publication fact verification and reducing the risk of unintentional misinformation dissemination. For instance, newsroom APIs can integrate the detection framework to flag suspicious statements before publishing, acting as a real-time editorial aid.

Partnerships with regulatory bodies, such as the Federal Communications Commission (FCC) or Federal Trade Commission (FTC), can facilitate voluntary compliance standards, certification mechanisms, and industry-wide benchmarking. These relationships can ensure the system aligns with evolving regulatory frameworks while maintaining industry trust.

Furthermore, ongoing integration with independent fact-checking organizations such as PolitiFact, Snopes, and the International Fact-Checking Network (IFCN) can provide valuable training data and feedback for model refinement. These organizations can validate flagged claims, offer human-in-the-loop validation, and help tune the system for nuanced cultural or political contexts.

In conclusion, effective deployment of a real-time fake news detection framework is not merely a technical achievement but a multifaceted endeavor requiring inclusive design, ethical sensitivity, legal compliance, and cooperative engagement with media stakeholders. A user-centered approach, anchored in transparency and accountability, is essential for fostering public trust and promoting responsible information ecosystems in the digital age (Weigl *et al*., 2022; Barua and Rahman, 2023).

2.7 Challenges and Limitations

While machine learning-based frameworks for real-time fake news detection present significant promise, their implementation and deployment are fraught with complex challenges and limitations as shown in figure 3(Lavin *et al*., 2022; Zafar *et al*., 2023). These issues span technical, ethical, and societal dimensions. Three critical areas of concern are adversarial news generation, bias in training data, and the risk of false positives which may erode public trust. Addressing these challenges is essential for the robustness, fairness, and societal acceptance of such systems.

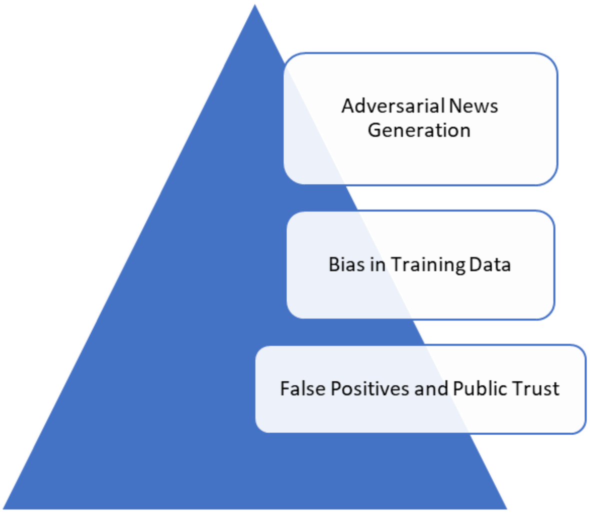

Fig 3: Challenges and Limitations

One of the most formidable technical challenges is the dynamic and adaptive nature of fake news generation. Misinformation creators continuously evolve their strategies to evade detection mechanisms, often employing sophisticated adversarial techniques. These include the use of obfuscated language, paraphrasing, AI-generated content, and visual manipulation such as doctored images or videos. With the advent of generative models like GPT and other large language models (LLMs), malicious actors can now produce deceptive content that closely mimics the tone, structure, and factual density of legitimate news articles.

Traditional fake news classifiers, particularly those trained on static datasets, struggle to generalize to these evolving patterns. Without real-time retraining and adversarial robustness, models may quickly become outdated or vulnerable to manipulation. Moreover, adversarial attacks—such as word-level perturbations or sentence insertions—can subtly alter the linguistic profile of a fake article, causing it to bypass detection. To mitigate this, ongoing research is exploring the incorporation of adversarial training and ensemble modeling, yet these approaches introduce computational overhead and complexity, challenging the deployment of lightweight, real-time systems (Zhao *et al*., 2022; Nowroozi *et al*., 2023).

Another significant limitation is the presence of bias within training datasets. Fake news detection models are only as reliable as the data on which they are trained. If the training corpus disproportionately represents certain political ideologies, ethnic groups, geographic regions, or media outlets, the resulting model may inadvertently replicate or amplify these biases.

Such biases undermine both the ethical foundation and practical utility of fake news detection systems. Biased models may unjustly flag content from marginalized communities or underrepresented perspectives, exacerbating issues of censorship and misinformation policing. Ensuring dataset representativeness across various domains—politics, health, environment, culture—and linguistic diversity (dialects, vernaculars, minority languages) is critical. Furthermore, transparency in data provenance and the use of fairness auditing tools during model development can help identify and reduce bias. Efforts to balance the dataset often face trade-offs between inclusion and precision. Additionally, curating balanced datasets is resource-intensive, requiring continuous collaboration with fact-checking organizations and domain experts (Ling *et al*., 2023; Zhao *et al*., 2023).

Perhaps the most socially impactful limitation is the risk of false positives—cases where truthful content is misclassified as fake. In high-stakes domains such as politics or public health, false positives can have severe consequences, including unjustified suppression of legitimate voices, reputational damage to credible news outlets, and erosion of freedom of expression. Repeated misclassifications may foster public distrust not only in the detection system itself but in the broader institutional frameworks that support it.

The challenge lies in achieving the right balance between sensitivity (identifying most fake news) and specificity (correctly validating true news). Overly sensitive models may err on the side of caution and over-flag content, while conservative models may fail to intercept harmful misinformation early enough. This balance is complicated by the subjective nature of truth in some contexts, where journalistic nuance or evolving scientific consensus makes binary classification inadequate.

To address this, confidence scoring and explainable AI (XAI) techniques are being integrated into detection frameworks (Mahbooba *et al*., 2021; Nwakanma *et al*., 2023). These tools help communicate the rationale behind a classification decision, enabling human reviewers to interpret, validate, or override the model's output. Moreover, integrating human-in-the-loop systems allows for post-classification review, reducing the reliance on fully automated decision-making in critical cases.

The development of a real-time fake news detection system faces significant challenges that must be carefully managed to ensure its credibility, reliability, and ethical alignment. Adversarial evolution, dataset bias, and the risk of false positives require a multi-pronged approach involving technical innovation, data transparency, and inclusive governance (Marchant, G.E. and Gutierrez, 2022; Nailwal *et al*., 2023; McDaniel and Koushanfar, 2023). Only by addressing these limitations can such frameworks achieve widespread adoption and play a meaningful role in safeguarding the integrity of digital information ecosystems.

Conclusion and Future Work

This presented a comprehensive AI-driven framework for real-time fake news detection tailored to the evolving digital news environment in the United States. The proposed architecture integrates advanced natural language processing, hybrid machine learning models, and real-time stream processing to identify and classify deceptive content with high precision and scalability. Key contributions include a modular and adaptive system design capable of ingesting, analyzing, and filtering vast volumes of news content across diverse platforms with minimal latency. The use of both traditional classifiers and deep learning models such as BERT and LSTM enables contextual, semantic, and social feature extraction, enhancing the system's robustness against nuanced misinformation tactics. Furthermore, the incorporation of explainability modules and feedback loops positions the framework as both technically sound and ethically responsible.

Given the increasing societal harm caused by the rapid spread of misinformation, there is a pressing need for widespread industry adoption. Media platforms, social networking companies, and digital content providers are encouraged to integrate intelligent fake news filters into their operational workflows. Such adoption not only enhances content integrity but also fosters public trust, protects democratic discourse, and supports editorial accountability. Collaboration with regulators and independent fact-checkers is essential for transparent, fair, and inclusive implementation.

Future research and development should focus on several key areas to further enhance the framework's capabilities. Expanding the model to support multilingual detection is critical for addressing misinformation in non-English content, particularly in multicultural digital landscapes. Additionally, integrating multimodal analysis—combining textual, visual, and audio signals—will improve the system's effectiveness in detecting manipulated images, deepfakes, and video-based disinformation. Lastly, cross-platform misinformation tracking, which traces the propagation of fake content across different media ecosystems, will be essential in capturing coordinated misinformation campaigns. These advancements will ensure that the framework remains adaptive, comprehensive, and effective in an increasingly complex and globalized digital information environment.

References

1. Aggarwal A, Chauhan A, Kumar D, Mittal M, Verma S. Classification of fake news by fine-tuning deep bidirectional transformers based language model. EAI Endorsed Trans Scalable Inf Syst. 2020;7(27):e10.
2. Akanbi A, Masinde M. A distributed stream processing middleware framework for real-time analysis of heterogeneous data on big data platform: case of environmental monitoring. Sensors. 2020;20(11):3166.
3. Akinola OI, Olaniyi OO, Ogungbemi OS, Oladoyinbo OB, Olisa AO. Resilience and recovery mechanisms for software-defined networking (SDN) and cloud networks. SSRN [Preprint]. 2024 [cited 2025 Aug 8]. Available from: https://ssrn.com/abstract=4908101.
4. Akpan H, Enyeribe A, Awe M. Digital asset and PII protection using blockchain technology. [place unknown: publisher unknown]; 2022. doi:10.XXXX/XXXX.
5. Aksoy C. Digital business ecosystems: an environment of collaboration, innovation, and value creation in the digital age. J Bus Trade. 2023;4(2):156-80.
6. Alam A, Ullah I, Lee YK. Video big data analytics in the cloud: a reference architecture, survey, opportunities, and open research issues. IEEE Access. 2020;8:152377-422.
7. Alonso MA, Vilares D, Gómez-Rodríguez C, Vilares J. Sentiment analysis for fake news detection. Electronics. 2021;10(11):1348.
8. Al-Quran MWM. Traditional media versus social media: challenges and opportunities. Technium Rom J Appl Sci Technol. 2022;4(10):145-60.
9. Alshanik F, Apon A, Herzog A, Safro I, Sybrandt J. Accelerating text mining using domain-specific stop word lists. In: 2020 IEEE International Conference on Big Data (Big Data); 2020 Dec 10-13; [place unknown]. Piscataway (NJ): IEEE; 2020. p. 2639-48.
10. Aniebonam EE. Strategic management in turbulent markets: a case study of the USA. Int J Mod Sci Res Technol. 2024;1(8):35-43.
11. Aniebonam EE, Chukwuba K, Emeka N, Taylor G. Transformational leadership and transactional leadership styles: systematic review of literature. Int J Appl Res. 2023;9(1):7-15.
12. Aniebonam EE, Nwabekee US, Ogunsola OY, Elumilade OO. International Journal of Management and Organizational Research. [place unknown: publisher unknown]; 2022.
13. Asonze CU, Ogungbemi OS, Ezeugwa FA, Olisa AO, Akinola OI, Olaniyi OO. Evaluating the trade-offs between wireless security and performance in IoT networks: a case study of web applications in AI-driven home appliances. SSRN [Preprint]. 2024 [cited 2025 Aug 8]. Available from: https://ssrn.com/abstract=4927991.
14. Babalola O, Adedoyin A, Ogundipe F, Folorunso A, Nwatu CE. Policy framework for cloud computing: AI, governance, compliance and management. Glob J Eng Technol Adv. 2024;21(02):114-26.
15. Bako NZ, Ozioko CN, Sanni IO, Oni O. The integration of AI and blockchain technologies for secure data management in cybersecurity. [place unknown: publisher unknown]; 2025.
16. Balogun AY, Olaniyi OO, Olisa AO, Gbadebo MO, Chinye NC. Enhancing incident response strategies in US healthcare cybersecurity. J Eng Res Rep. 2025;27(2):114-35.
17. Barua T, Rahman MA. A systematic literature review of user-centric design in digital business systems enhancing accessibility, adoption, and organizational impact. Am J Sch Res Innov. 2023;2(02):193-216.
18. Beauvais C. Fake news: why do we believe it? Joint Bone Spine. 2022;89(4):105371.
19. Bhattacharya A. Applied machine learning explainability techniques: make ML models explainable and trustworthy for practical applications using LIME, SHAP, and more. Birmingham: Packt Publishing Ltd; 2022.
20. Birchall C, Knight P. Conspiracy theories in the time of COVID-19. Abingdon: Taylor & Francis; 2022.
21. Bondielli A, Marcelloni F. A survey on fake news and rumour detection techniques. Inf Sci. 2019;497:38-55.
22. Bowers AJ, Zhou X. Receiver operating characteristic (ROC) area under the curve (AUC): a diagnostic measure for evaluating the accuracy of predictors of education outcomes. J Educ Stud Placed Risk. 2019;24(1):20-46.
23. Carrington AM, Manuel DG, Fieguth PW, Ramsay T, Osmani V, Wernly B, *et al*. Deep ROC analysis and AUC as balanced average accuracy, for improved classifier selection, audit and explanation. IEEE Trans Pattern Anal Mach Intell. 2022;45(1):329-41.
24. Casaleiro R. Protection and control of personal identifiable information: the PoSeID-on approach. J Data Prot Priv. 2020;3(2):199-228.
25. Cavaliere P. From journalistic ethics to fact-checking practices: defining the standards of content governance in the fight against disinformation. J Media Law. 2020;12(2):133-65.
26. Chadwick A, Stanyer J. Deception as a bridging concept in the study of disinformation, misinformation, and misperceptions: toward a holistic framework. Commun Theory. 2022;32(1):1-24.
27. Chen W, Milosevic Z, Rabhi FA, Berry A. Real-time analytics: concepts, architectures, and ML/AI considerations. IEEE Access. 2023;11:71634-57.
28. Cherish EI, Aniebonam EE, Cong TT, Cong TT. Impact

of AI on financial performance of enterprises. Cuestiones Fisioterapia. 2025;54(5):385-96.
29. Chitta K, Álvarez JM, Haussmann E, Farabet C. Training data subset search with ensemble active learning. IEEE Trans Intell Transp Syst. 2021;23(9):14741-52.
30. Cover R, Haw A, Thompson JD. Fake news and conspiracy theories. In: Fake news in digital cultures: technology, populism and digital misinformation. Bingley: Emerald Publishing Limited; 2022. p. 77-91.
31. Damstra A, Boomgaarden HG, Broda E, Lindgren E, Strömbäck J, Tsfati Y, *et al*. What does fake look like? A review of the literature on intentional deception in the news and on social media. Journal Stud. 2021;22(14):1947-63.
32. Demartini G, Mizzaro S, Spina D. Human-in-the-loop artificial intelligence for fighting online misinformation: challenges and opportunities. IEEE Data Eng Bull. 2020;43(3):65-74.
33. Dubuc T, Stahl F, Roesch EB. Mapping the big data landscape: technologies, platforms and paradigms for real-time analytics of data streams. IEEE Access. 2020;9:15351-74.
34. Ebepu OO, Aniebonam EE, Waheed OO, Asamoah F. Advanced market analysis and United States business growth: identifying emerging opportunities for sustainable profitability. [place unknown: publisher unknown]; 2025.
35. Ebepu OO, Okpeseyi SBA, John-Ogbe J, Emmanuel E. Harnessing data-driven strategies for sustained United States business growth: a comparative analysis of market leaders. [place unknown: publisher unknown]; 2024.
36. Erikson RS, Tedin KL. American public opinion: its origins, content, and impact. Abingdon: Routledge; 2019.
37. Evans H, Agoro H. Integrating generative AI with IoT devices for enhanced user experience. [place unknown: publisher unknown]; 2023.
38. Folorunso A, Nwatu CE, Adedoyin A, Ogundipe F. Policy framework for cloud computing: AI, governance, compliance, and management. Glob J Eng Technol Adv. 2024;21(02):114-26.
39. Goldman AI, Baker D. Free speech, fake news, and democracy. First Amend L Rev. 2019;18:66.
40. Guo B, Ding Y, Yao L, Liang Y, Yu Z. The future of false information detection on social media: new perspectives and trends. ACM Comput Surv. 2020;53(4):1-36.
41. Halliday NN. Assessment of major air pollutants, impact on air quality and health impacts on residents: case study of cardiovascular diseases [master's thesis]. Cincinnati (OH): University of Cincinnati; 2021.
42. Hamed SK, Ab Aziz MJ, Yaakub MR. A review of fake news detection approaches: a critical analysis of relevant studies and highlighting key challenges associated with the dataset, feature representation, and data fusion. Heliyon. 2023;9(10):e20674.
43. Hendrix J, Morozoff D. Media forensics in the age of disinformation. In: Multimedia forensics. Singapore: Springer; 2022. p. 7-40.
44. Hersh E, Krupnikov Y. Introduction to freedom of expression in an age of social media, misinformation, and political polarization. PS Polit Sci Polit. 2023;56(2):219-21.
45. Joeaneke P, Kolade TM, Obioha Val O, Olisa AO, Joseph S, Olaniyi OO. Enhancing security and traceability in aerospace supply chains through blockchain technology. SSRN [Preprint]. 2024 [cited 2025 Aug 8]. Available from: https://ssrn.com/abstract=4995935.
46. Joeaneke P, Obioha Val O, Olaniyi OO, Ogungbemi OS, Olisa AO, Akinola OI. Protecting autonomous UAVs from GPS spoofing and jamming: a comparative analysis of detection and mitigation techniques. SSRN [Preprint]. 2024 [cited 2025 Aug 8]. Available from: https://ssrn.com/abstract=XXXXXXX.
47. John AO, Oyeyemi BB. The role of AI in oil and gas supply chain optimization. Int J Multidiscip Res Growth Eval. 2022;3(1):1075-86.
48. Juneja P, Mitra T. Human and technological infrastructures of fact-checking. Proc ACM Hum-Comput Interact. 2022;6(CSCW2):1-36.
49. Karegar F, Pettersson JS, Fischer-Hübner S. The dilemma of user engagement in privacy notices: effects of interaction modes and habituation on user attention. ACM Trans Priv Secur. 2020;23(1):1-38.
50. Kolajo T, Daramola O, Adebiyi A. Big data stream analysis: a systematic literature review. J Big Data. 2019;6(1):47.
51. Ladani DJ, Desai NP. Stopword identification and removal techniques on tc and ir applications: a survey. In: 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS); 2020 Mar 6-7; [place unknown]. Piscataway (NJ): IEEE; 2020. p. 466-72.
52. Larraz I. Towards automated fact-checking in Africa: the experience with artificial intelligence at Africa Check. In: Digitisation, AI and algorithms in African journalism and media contexts: practice, policy and critical literacies. Bingley: Emerald Publishing Limited; 2023. p. 3-16.
53. Lavin A, Gilligan-Lee CM, Visnjic A, Ganju S, Newman D, Ganguly S, *et al*. Technology readiness levels for machine learning systems. Nat Commun. 2022;13(1):6039.
54. Lei L, Li Y, Luo Y. Production and dissemination of corporate information in social media: a review. J Account Lit. 2019;42(1):29-43.
55. Ling C, Zhao X, Lu J, Deng C, Zheng C, Wang J, *et al*. Domain specialization as the key to make large language models disruptive: a comprehensive survey. arXiv [Preprint]. 2023 [cited 2025 Aug 8]. Available from: https://arxiv.org/abs/2305.18703.
56. Lugea J. Linguistic approaches to fake news detection. In: Data science for fake news: surveys and perspectives. Cham: Springer; 2021. p. 287-302.
57. Mahbooba B, Timilsina M, Sahal R, Serrano M. Explainable artificial intelligence (XAI) to enhance trust management in intrusion detection systems using decision tree model. Complexity. 2021;2021:6634811.
58. Manghi P, Atzori C, De Bonis M, Bardi A. Entity deduplication in big data graphs for scholarly communication. Data Technol Appl. 2020;54(4):409-35.
59. Marchant GE, Gutierrez CI. Soft law 2.0: an agile and effective governance approach for artificial intelligence. Minn J L Sci Tech. 2022;24:375.
60. Mavrogiorgou A, Kiourtis A, Manias G, Symvoulidis C, Kyriazis D. Batch and streaming data ingestion towards creating holistic health records. Emerg Sci J.

2023;7(2):339-53.
61. McDaniel P, Koushanfar F. Secure and trustworthy computing 2.0 vision statement. arXiv [Preprint]. 2023 [cited 2025 Aug 8]. Available from: https://arxiv.org/abs/2308.00623.
62. Mehmood E, Anees T. Challenges and solutions for processing real-time big data stream: a systematic literature review. IEEE Access. 2020;8:119123-43.
63. Mesinovic M, Watkinson P, Zhu T. Explainable AI for clinical risk prediction: a survey of concepts, methods, and modalities. arXiv [Preprint]. 2023 [cited 2025 Aug 8]. Available from: https://arxiv.org/abs/2308.08407.
64. Mishra S, Shukla P, Agarwal R. Analyzing machine learning enabled fake news detection techniques for diversified datasets. Wirel Commun Mob Comput. 2022;2022:1575365.
65. Murmann P, Karegar F. From design requirements to effective privacy notifications: empowering users of online services to make informed decisions. Int J Hum-Comput Interact. 2021;37(19):1823-48.
66. Nailwal S, Singhal S, Singh NT, Raza A. Deepfake detection: a multi-algorithmic and multi-modal approach for robust detection and analysis. In: 2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE); 2023 Nov; [place unknown]. Piscataway (NJ): IEEE; 2023. p. 1-8.
67. Nemes L, Kiss A. Information extraction and named entity recognition supported social media sentiment analysis during the COVID-19 pandemic. Appl Sci. 2021;11(22):11017.
68. Ngoc NM, Van Thoi B, Tien NH, Aniebonam EE. Una evaluación de los precios de los activos para los rendimientos de las carteras de la industria estadounidense. Cuestiones Fisioterapia. 2025;54(3):2107-16.
69. Nguyen VL, Shaker MH, Hüllermeier E. How to measure uncertainty in uncertainty sampling for active learning. Mach Learn. 2022;111(1):89-122.
70. Nowroozi E, Mohammadi M, Savaş E, Mekdad Y, Conti M. Employing deep ensemble learning for improving the security of computer networks against adversarial attacks. IEEE Trans Netw Serv Manag. 2023;20(2):2096-105.
71. Nwabekee US, Aniebonam EE, Elumilade OO, Ogunsola OY. Integrating digital marketing strategies with financial performance metrics to drive profitability across competitive market sectors. J Mark Financ Perform. 2021;5(2):76-91.
72. Nwabekee US, Aniebonam EE, Elumilade OO, Ogunsola OY. Predictive model for enhancing long-term customer relationships and profitability in retail and service-based. [place unknown: publisher unknown]; 2021.
73. Nwakanma CI, Ahakonye LAC, Njoku JN, Odirichukwu JC, Okolie SA, Uzondu C, *et al*. Explainable artificial intelligence (XAI) for intrusion detection and mitigation in intelligent connected vehicles: a review. Appl Sci. 2023;13(3):1252.
74. Nwankwo EE, Aniebonam EE, Chikodiri US, Rita CN. Entrepreneurial environment and cross-cultural management: a road map for sustaining Nigerian development. [place unknown: publisher unknown]; 2025.
75. Nwankwo EE, Ewim DRE, Aniebonam EE, Chikodiri US, Rita CN. Skills for engineers to become entrepreneurs for sustainable development in Anambra State. World. 2025;2579:0544.
76. Nwankwo EE, Okafor GI, Aniebonam EE, Chikodiri US. Relationship between gender discrimination, work-life balance and mental health: the Nigerian female educators' perspective. Int Inst Acad Res Dev. 2025;11(1):296-307.
77. Obioha Val O, Lawal T, Olaniyi OO, Gbadebo MO, Olisa AO. Investigating the feasibility and risks of leveraging artificial intelligence and open source intelligence to manage predictive cyber threat models. SSRN [Preprint]. 2025 [cited 2025 Aug 8]. Available from: https://ssrn.com/abstract=XXXXXXX.
78. Obioha Val O, Olaniyi OO, Gbadebo MO, Balogun AY, Olisa AO. Cyber espionage in the age of artificial intelligence: a comparative study of state-sponsored campaign. SSRN [Preprint]. 2025 [cited 2025 Aug 8]. Available from: https://ssrn.com/abstract=XXXXXXX.
79. Ogundipe F, Bakare OI, Sampson E, Folorunso A. Harnessing digital transformation for Africa's growth: opportunities and challenges in the technological era. [place unknown: publisher unknown]; 2023.
80. Ogundipe F, Sampson E, Bakare OI, Oketola O, Folorunso A. Digital transformation and its role in advancing the sustainable development goals (SDGs). Transformation. 2019;19:48.
81. Ogunjobi O, Aniebonam EE, Faisal R, Durojaiye AT. Expanding access to capital: role of structured finance in stimulating SME growth and economic resilience in the United States. [place unknown: publisher unknown]; 2024.
82. Okon SU, Olateju O, Ogungbemi OS, Joseph S, Olisa AO, Olaniyi OO. Incorporating privacy by design principles in the modification of AI systems in preventing breaches across multiple environments, including public cloud, private cloud, and on-prem. SSRN [Preprint]. 2024 [cited 2025 Aug 8]. Available from: https://ssrn.com/abstract=XXXXXXX.
83. Olisa AO. Quantum-resistant blockchain architectures for securing financial data governance against next-generation cyber threats. J Eng Res Rep. 2025;27(4):189-211.
84. Oni O, Iloeje KF. Optimized Fast R-CNN for automated parking space detection: evaluating efficiency with MiniFasterRCNN. [place unknown: publisher unknown]; 2025.
85. Oni O. Memory-enhanced conversational AI: a generative approach for context-aware and personalized chatbots. [place unknown: publisher unknown]; 2025.
86. Onibokun T, Ejibenam A, Ekeocha PC, Onayemi HA, Halliday N. The use of AI to improve CX in SAAS environment. [place unknown: publisher unknown]; 2022.
87. Orenuga A, Oyeyemi BB, Olufemi John A. AI and sustainable supply chain practices: ESG goals in the US and Nigeria. [place unknown: publisher unknown]; 2024.
88. Oyeyemi BB. Artificial intelligence in agricultural supply chains: lessons from the US for Nigeria. [place unknown: publisher unknown]; 2022.
89. Oyeyemi BB. Data-driven decisions: leveraging

predictive analytics in procurement software for smarter supply chain management in the United States. [place unknown: publisher unknown]; 2023.
90. Oyeyemi BB, Orenuga A, Adelakun BO. Blockchain and AI synergies in enhancing supply chain transparency. [place unknown: publisher unknown]; 2024.
91. Peng H, Li J, Song Y, Yang R, Ranjan R, Yu PS, *et al*. Streaming social event detection and evolution discovery in heterogeneous information networks. ACM Trans Knowl Discov Data. 2021;15(5):1-33.
92. Prochaska S, Duskin K, Kharazian Z, Minow C, Blucker S, Venuto S, *et al*. Mobilizing manufactured reality: how participatory disinformation shaped deep stories to catalyze action during the 2020 US presidential election. Proc ACM Hum-Comput Interact. 2023;7(CSCW1):1-39.
93. Radanliev P, De Roure D, Santos O. Red teaming generative AI/NLP, the BB84 quantum cryptography protocol and the NIST-approved quantum-resistant cryptographic algorithms. arXiv [Preprint]. 2023 [cited 2025 Aug 8]. Available from: https://arxiv.org/abs/2310.04425.
94. Rajest SS, Singh B, Obaid AJ, Regin R, Chinnusamy K, editors. Advances in artificial and human intelligence in the modern era. Hershey (PA): IGI Global; 2023.
95. Rath B, Salecha A, Srivastava J. Fake news spreader detection using trust-based strategies in social networks with bot filtration. Soc Netw Anal Min. 2022;12(1):66.
96. Rathore SPS, Farhaoui Y, Aniebonam EE, Nagpal T, Kaushik P. AI-driven traffic congestion management: a predictive analytics approach for smart cities. In: 2025 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI); 2025 Mar; [place unknown]. Piscataway (NJ): IEEE; 2025. p. 1-6.
97. Reglitz M. Fake news and democracy. J Ethics Soc Phil. 2022;22:162.
98. Safder I, Hassan SU, Visvizi A, Noraset T, Nawaz R, Tuarob S. Deep learning-based extraction of algorithmic metadata in full-text scholarly documents. Inf Process Manag. 2020;57(6):102269.
99. Santos FCC. Artificial intelligence in automated detection of disinformation: a thematic analysis. Journal Media. 2023;4(2):679-87.
100. Selesi-Aina O, Obot NE, Olisa AO, Gbadebo MO, Olateju O, Olaniyi OO. The future of work: a human-centric approach to AI, robotics, and cloud computing. J Eng Res Rep. 2024;26(11):10-9734.
101. Singh J, Sahu DP, Gupta T, Singhal D, Lal B, Turukmane AV. Fake news detection using transfer learning. In: International Conference on Recent Trends in Image Processing and Pattern Recognition; 2023 Dec; Cham: Springer Nature Switzerland; 2023. p. 14-24.
102. Sinha D. Practical machine learning on Databricks: seamlessly transition ML models and MLOps on Databricks. Birmingham: Packt Publishing Ltd; 2023.
103. Sinha K, Ghosh N, Sil PC. A review on the recent applications of deep learning in predictive drug toxicological studies. Chem Res Toxicol. 2023;36(8):1174-205.
104. Tufchi S, Yadav A, Ahmed T. A comprehensive survey of multimodal fake news detection techniques: advances, challenges, and opportunities. Int J Multimed Inf Retr. 2023;12(2):28.
105. Ünver A. Emerging technologies and automated fact-checking: tools, techniques and algorithms. [place unknown: publisher unknown]; 2023.
106. Verma PK, Agrawal P, Amorim I, Prodan R. WELFake: word embedding over linguistic features for fake news detection. IEEE Trans Comput Soc Syst. 2021;8(4):881-93.
107. Wang D, Zhang W, Wu W, Guo X. Soft-label for multi-domain fake news detection. IEEE Access. 2023;11:98596-606.
108. Weidinger L, Mellor J, Rauh M, Griffin C, Uesato J, Huang PS, *et al*. Ethical and social risks of harm from language models. arXiv [Preprint]. 2021 [cited 2025 Aug 8]. Available from: https://arxiv.org/abs/2112.04359.
109. Weigl L, Amard A, Marxen H, Roth T, Zavolokina L. User-centricity and public values in eGovernment: friend or foe? In: Proceedings of the 30th European Conference on Information Systems (ECIS). Atlanta (GA): Association for Information Systems; 2022.
110. Zafar I, Anwar S, Yousaf W, Nisa FU, Kausar T, ul Ain Q, *et al*. Reviewing methods of deep learning for intelligent healthcare systems in genomics and biomedicine. Biomed Signal Process Control. 2023;86:105263.
111. Zhao W, Alwidian S, Mahmoud QH. Adversarial training methods for deep learning: a systematic review. Algorithms. 2022;15(8):283.
112. Zhao X, Lu J, Deng C, Zheng C, Wang J, Chowdhury T, *et al*. Domain specialization as the key to make large language models disruptive: a comprehensive survey. arXiv [Preprint]. 2023 [cited 2025 Aug 8]. Available from: https://arxiv.org/abs/2305.18703.
113. Zhou X, Zafarani R. Network-based fake news detection: a pattern-driven approach. ACM SIGKDD Explor Newsl. 2019;21(2):48-60.