



AI-powered fake news detection: Leveraging machine learning and NLP for automated fact-checking

Ukanwolu Grace Ngozi, Omankwu Obinnaya Chinecherem Beloved

Department of Computer Science, Michael Okpara University of Agriculture Umudike, Nigeria

Abstract

Fake news has become a significant challenge in the digital age, influencing public opinion, shaping political discourse, and affecting decision-making processes. This study presents an AI-driven approach to fake news detection using machine learning and deep learning techniques, with a focus on sentiment analysis and natural language processing (NLP). The research involves data collection, preprocessing, feature extraction, model training, and evaluation to develop an effective and interpretable detection system. A labeled dataset was sourced from Kaggle.com, containing both real and fake news articles. Preprocessing techniques such as text normalization, tokenization, stopword removal, stemming, and lemmatization were applied to clean and standardize the text. Feature extraction methods, including TF-IDF, Word2Vec, and BERT embeddings, were used to convert textual data into numerical representations suitable for machine learning models. Various models, including Logistic Regression, Support Vector Machines (SVM), Random Forest, LSTM, and BERT, were trained and evaluated using metrics such as accuracy, precision, recall, and F1-score. The implementation leveraged powerful tools such as Python, Pandas, NLTK, Scikit-Learn, TensorFlow, and PyTorch for model development and analysis. Experimental results demonstrated that incorporating sentiment-based and textual features significantly improves classification accuracy. The study provides a robust and scalable AI-based framework for automated fact-checking and misinformation detection. The findings contribute to combating the spread of fake news, with potential applications in content moderation on social media platforms and real-time misinformation detection systems.

Keywords: Fake news detection, machine learning, natural language processing, feature extraction, classification models.

Introduction

The rapid proliferation of digital media has led to an increase in the spread of fake news, which can have serious societal implications. Fake news is defined as false or misleading information presented as news, often with the intent to deceive readers. This paper aims to address the issue by developing a robust fake news detection system leveraging machine learning techniques. Recent studies have demonstrated the effectiveness of sentiment analysis in conjunction with other AI techniques, such as machine learning and deep learning, in detecting fake news. For instance, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have been employed to analyze large datasets of social media posts, identifying subtle cues and inconsistencies that may signify false information (Zhou and Zafarani, 2018) ^[40]. Additionally, sentiment analysis can be integrated with other data points, such as the credibility of sources and the historical behavior of users, to create a more comprehensive fake news detection framework (Shu ^[24], Sliva, Wang, Tang, and Liu, 2017) ^[36]. The application of AI-based sentiment analysis to fake news detection is not without its challenges. Issues such as the dynamic nature of language, the presence of sarcasm or irony, and the cultural context of sentiments can complicate the analysis (Cambria, 2016). Furthermore, the ethical considerations of employing AI to monitor and potentially censor content on social media platforms must be carefully addressed to balance the benefits of fake news detection with the protection of free speech (Floridi, 2019) ^[6]. Despite these challenges, the integration of AI and sentiment analysis offers a powerful tool in the ongoing effort to mitigate the impact of fake news on social media.

The proliferation of social media has fundamentally transformed how information is disseminated and consumed globally. Platforms such as Facebook, Twitter, and Instagram have become primary sources of news for millions of users, offering immediacy and a wide reach. However, this rapid dissemination of information has also facilitated the spread of fake news—misinformation or false information presented as news—posing significant challenges to societies worldwide (Allcott and Gentzkow, 2017). The prevalence of fake news on social media has been linked to political polarization, erosion of public trust in traditional media, and even societal unrest (Vosoughi, Roy, and Aral, 2018) ^[34]. The urgency to address the spread of fake news has led to the exploration of various technological solutions. Among these, artificial intelligence (AI) and machine learning techniques have shown considerable promise. Specifically, sentiment analysis, a subfield of natural language processing (NLP), has emerged as a powerful tool in this context. Sentiment analysis involves the computational identification and categorization of opinions expressed in text to determine the writer's attitude towards a particular topic, which can be indicative of the nature of the news content (Liu, 2012) ^[16]. By analyzing sentiment patterns, AI can help discern between genuine news articles and those that are likely fabricated or misleading (Giachanou and Crestani, 2016) ^[8].

The implications of effectively identifying fake news are profound. By mitigating the spread of misinformation, societies can foster a more informed and engaged public, enhance democratic processes, and reduce the risk of social division. However, the development and deployment of these technologies also raise ethical considerations, such as privacy concerns and the potential for misuse (Floridi *et al.*,

2018)^[7]. Therefore, it is imperative that ongoing research in this area not only focuses on technological innovation but also on establishing frameworks for ethical and responsible use.

The proliferation of social media has fundamentally transformed how information is disseminated and consumed. While these platforms offer numerous benefits, they have also become fertile ground for the spread of fake news, which can have significant societal impacts. Fake news, defined as false or misleading information presented as news, can influence public opinion, sway elections, and exacerbate social divisions (Allcott and Gentzkow, 2017). Recent trends underscore the urgency of addressing this issue. According to a study by the Pew Research Center, nearly two-thirds of American adults believe that fabricated news stories cause a great deal of confusion about the basic facts of current events (Mitchell *et al.*, 2019)^[17]. The rise of fake news is not limited to any single region; it is a global phenomenon. For instance, in India, the spread of misinformation on social media has been linked to violent incidents, prompting the government to take measures against platforms like WhatsApp to curb the dissemination of fake news (Banaji *et al.*, 2019). Similarly, during the COVID-19 pandemic, a flood of false information about the virus's origins, treatments, and vaccines circulated on social media, complicating public health efforts and undermining trust in official sources (Cinelli *et al.*, 2020)^[3]. These instances highlight the need for effective tools to identify and combat fake news.

Artificial intelligence (AI), particularly sentiment analysis, has emerged as a promising solution to tackle the challenge of fake news. Sentiment analysis, also known as opinion mining, involves using AI algorithms to analyze the emotional tone behind words to understand the attitudes, opinions, and emotions expressed in textual data (Liu, 2012)^[16]. By applying sentiment analysis to social media content, AI can help detect patterns indicative of fake news. For example, fake news articles often exhibit heightened emotional language intended to provoke strong reactions, which can be a distinguishing feature from legitimate news (Vosoughi, Roy, and Aral, 2018)^[34]. The application of AI in identifying fake news is supported by advancements in machine learning and natural language processing (NLP). Techniques such as deep learning have significantly improved the accuracy of sentiment analysis models, enabling them to process vast amounts of data and identify subtle linguistic cues associated with fake news (Zhou and Zafarani, 2020)^[40]. Moreover, AI systems can be continually updated and trained on new data, making them adaptable to evolving tactics used by purveyors of misinformation. This dynamic capability is essential in maintaining the relevance and effectiveness of fake news detection tools. In conclusion, the spread of fake news on social media is a pressing issue with far-reaching consequences. The integration of AI and sentiment analysis offers a powerful approach to identifying and mitigating fake news, leveraging the latest technological advancements to enhance accuracy and adaptability. Addressing this challenge is crucial for preserving the integrity of information and safeguarding public trust in an increasingly digital world.

2. Literature Review: Several studies have explored fake news detection using various approaches such as rule-based

systems, deep learning models, and hybrid techniques. Traditional methods rely on manual fact-checking, which is time-consuming and inefficient. Recent advancements in NLP and machine learning have paved the way for automated detection systems that can analyze text content and classify news articles with high accuracy. The proliferation of fake news on social media platforms has emerged as a critical issue in the digital age, significantly influencing public opinion and behavior. The rapid dissemination of misinformation can have far-reaching consequences, from undermining democratic processes to inciting social unrest and affecting public health. This phenomenon underscores the urgent need for effective detection mechanisms to combat the spread of false information. The advent of artificial intelligence (AI) and sentiment analysis offers promising solutions to this challenge. By leveraging advanced computational techniques, these technologies can analyze and interpret vast amounts of data, identifying patterns indicative of fake news with remarkable accuracy (Zhang and Ghorbani, 2020)^[38, 39].

The application of AI in fake news detection primarily involves machine learning algorithms and natural language processing (NLP). Machine learning models, trained on large datasets of labeled news articles, learn to distinguish between genuine and fraudulent content. Sentiment analysis, a subset of NLP, further enhances this capability by evaluating the emotional tone of the text, which often reveals clues about its authenticity (Liu, 2012)^[16]. Studies have shown that fake news tends to exhibit distinct sentiment patterns compared to legitimate news, making sentiment analysis a valuable tool in this context (Shu *et al.*, 2017)^[24].

Social media platforms, such as Facebook and Twitter, are particularly vulnerable to the spread of fake news due to their user-driven content creation and sharing features. The viral nature of social media allows misinformation to reach a wide audience rapidly, often before it can be fact-checked or debunked (Vosoughi, Roy, and Aral, 2018)^[34]. The integration of AI-based tools into these platforms can significantly enhance their ability to detect and mitigate the spread of fake news, thus preserving the integrity of the information ecosystem.

Despite the advancements in AI and sentiment analysis, the detection of fake news remains a complex task. Challenges include the evolving nature of misinformation tactics, the contextual nuances of language, and the need for large, high-quality datasets for training AI models (Pennycook and Rand, 2019)^[20]. Moreover, ethical considerations, such as privacy concerns and the potential for bias in AI algorithms, must be addressed to ensure the responsible deployment of these technologies (Zhang and Ghorbani, 2020)^[38, 39].

This literature review aims to explore the current state of research on the identification of fake news in social media using AI and sentiment analysis. It will examine the methodologies employed, the effectiveness of various approaches, and the challenges encountered in this field. By synthesizing the findings from recent studies, this review seeks to provide a comprehensive understanding of the capabilities and limitations of AI-driven fake news detection, thereby highlighting areas for future research and development (Allcott and Gentzkow, 2017).

Methodology

1. Data Ingestion The first step in the proposed system is data collection from reliable sources such as news websites, social media platforms, and fact-checking organizations. However, the data used in this study was

generated from kaggle.com, which is an online data resource website where one can easily generate data for study and analysis. The collected data was stored in a structured format for further processing.

1	label	content
2	0	1 donald trump send embarrass new year eve message disturb donald trump wish americans happy new year leave instead give shout enemies haters dishonest fake news media former reality sh
3	1	1 drink brag trump staffer start russian collusion investigation house intelligence committee chairman devin Nunes go bad day assumption like many us christopher steele dossier prompt russia in
4	2	1 sheriff david clarke become internet joke threaten poke people eye friday reveal former milwaukee sheriff david clarke consider homeland security secretary donald trump administration emai
5	3	1 trump obsess even obama name cod website image christmas day donald trump announce would back work follow day golf fourth day row former reality show star blast former president barack
6	4	1 pope francis call donald trump christmas speech pope francis use annual christmas day message rebuke donald trump without even mention name pope deliver message days members unite na
7	5	1 racist alabama cop brutalize black boy handcuff graphic image number case cop brutalize kill people color seem see end another case need share far wide alabama woman name angela williams
8	6	1 fresh golf course trump lash fbi deputy director jam comey donald trump spend good portion day golf club mark day do since take oath office must bad game trump lash fbi deputy director andre
9	7	1 trump say insanely racist stuff inside oval office witness back wake yet another court decision derail donald trump plan bar muslims enter unite state new york time publish report saturday morr
10	8	1 former cia director slam trump un bully openly suggest act like dictator tweet many people raise alarm regard fact donald trump dangerously close become autocrat thing democracies become a
11	9	1 watch brand new pro trump ad feature much kiss make sick might think get break watch people kiss donald trump ass stroke ego ad nauseam pro trump group create ad nothing people even exa
12	10	1 papa john founder retire figure racism bad business centerpiece donald trump campaign presidency white supremacist ways many public feud get involve people color one favorite target course
13	11	1 watch paul ryan tell us care struggle families live blue state republicans work overtime try sell scam tax bill public something directly target middle class work class families financial relief nothir
14	12	1 bad news trump mitch mcconnell say repeal obamacare republicans seven years come viable replacement obamacare fail miserably take victory lap gift wealthy tax break wednesday donald tru
15	13	1 watch lindsey graham trash media portray trump kooky forget word media talk day trump republican party scam tax bill well sheer obsequiousness trump cabinet members congress tax scam pa
16	14	1 heiress disney empire know gop scammed us shred tax bill abigail disney heiress brass ovaries profit gop tax scam bill f cking poor people ms disney pen op ed usa today rip gop new one always
17	15	1 tone deaf trump congrats rep scalise lose weight almost die donald trump sign gop tax scam law course mean invite craven cruel gop sycophants perch capitol hill celebrate rise garden white hou
18	16	1 internet brutally mock disney new trump robot hall presidents new animatronic figure hall presidents walt disney world add every former leader republic depict audio animatronics show figure
19	17	1 mueller spokesman f cked donald trump christmas trump supporters call president favorite network lash special counsel robert mueller fbi white house panic moore mueller obtain tens thousan
20	18	1 snl hilariously mock accuse child molester roy moore lose al senate race video right whole world look shock fact democrat doug jones beat republican roy moore special election replace attorney
21	19	1 republican senator get drag go robert mueller senate majority whip john cornyn r tx think would good idea attack special counsel robert mueller russia probe mueller noose tighten republicans l
22	20	1 heartless rebuke victims trump invite nra xmas party sandy hook anniversary almost seem like donald trump troll america point begin try gaslight country insist crowd inauguration biggest ever e
23	21	1 ky gop state rep commit suicide allegations molest teen girl detail metoo moment many powerful men topple span many industries entertainment journalism politics beyond man ever dare abu

Fig 1: Data frame of the sample Data used for the Sentiment analysis

The image represents the structured data frame of the sample dataset used for sentiment analysis in detecting fake news. The dataset was first organized in Microsoft Excel, where it was cleaned and formatted before being transferred to VS Code for further processing using Python and machine learning libraries. Each row in the data frame represents an individual news article, while the columns contain relevant features, including:

Text Content: The body of the news article or social media post.

Sentiment Score: A numerical value representing the polarity of the text (positive, neutral, or negative).

Sentiment Intensity: A measure of how strong or weak the sentiment is.

Source: The platform where the news was published (e.g., Twitter, Facebook, News Website).

Publication Date: The date when the news article was posted.

Label: Indicates whether the news is Real (1) or Fake (0) based on pre-labeled training data.

After structuring the data in Excel, it was imported into VS Code using Pandas for further preprocessing, tokenization, and feature extraction. This step ensured that the data was in a suitable format for training machine learning models to classify news as fake or real based on sentiment features.

2. Data Preprocessing: Preprocessing techniques include tokenization, noise removal, lemmatization, and stopword removal. These steps help in cleaning the text data and preparing it for analysis.

Unnamed: 0	id	title	text	subject	label
0	20104	At least 19 drown when boat capsizes in northe...	NEW DELHI (Reuters) - At least 19 people drown...	worldnews	0
1	13719	North Korea may announce completion of nuclear...	SEOUL (Reuters) - North Korea may announce the...	worldnews	0
2	17393	Indonesia considers ban on "destructive" LGBT...	JAKARTA (Reuters) - Days after a long-running ...	worldnews	0
3	27685	Think Trump Wants To Keep Your Job At Home? T...	One of the cornerstones of Donald Trump's camp...	News	1
4	30345	Alabama's Chief Justice Just Spilt On The Com...	One would think that to hold a position like C...	News	1

Fig 2: Data visualization of the proceed Data in Visual studio Code

The image represents the data visualization of the processed dataset used for fake news detection in Visual Studio Code (VS Code). Various visual analytics techniques were applied to explore and understand key patterns in the dataset. The visualization includes bar charts, histograms, word clouds, and sentiment distribution graphs, providing insights into the characteristics of real and fake news articles.

Key aspects of the visualization:

Sentiment Distribution: Displays the polarity (positive, neutral, or negative) of news articles to identify sentiment trends in fake and real news.

Word Frequency Analysis: Highlights the most commonly used words in fake news articles compared to real news, using word clouds or bar plots.

Class Distribution: Shows the balance of real and fake news samples in the dataset to assess potential bias.

Feature Correlation: Demonstrates relationships between sentiment scores, text features, and classification labels. These visualizations help in understanding the dataset structure, improving feature engineering, and optimizing the performance of machine learning models.

3. Feature Extraction Feature extraction involves identifying key textual features such as term frequency-inverse document frequency (TF-IDF), sentiment analysis, and linguistic patterns that help in distinguishing fake news from real news.

4. Model Training and Evaluation: Various machine learning models such as Support Vector Machines (SVM), Random Forest, and deep learning models like Long Short-Term Memory (LSTM) networks are trained on the processed data. The models are evaluated based on metrics such as accuracy, precision, recall, and F1-score.

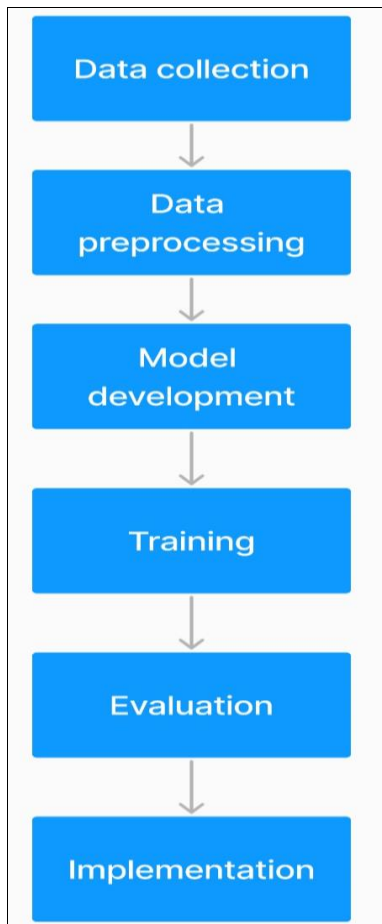


Fig 3: Methodology of the proposed system

The methodology for the proposed fake news detection system follows a structured approach, integrating machine learning and sentiment analysis to enhance classification accuracy. The process begins with data collection from Kaggle.com, where a labeled dataset containing real and fake news articles is obtained. Next, data preprocessing is performed using text normalization, tokenization, stopword removal, stemming, and lemmatization to clean and standardize the text. Feature extraction techniques such as TF-IDF (Term Frequency-Inverse Document Frequency), Word2Vec, and BERT embeddings are applied to convert textual data into numerical representations for machine learning models. The system utilizes Logistic Regression, Support Vector Machines (SVM), Random Forest, LSTM (Long Short-Term Memory Networks), and BERT to train and evaluate the classification models. Performance assessment is conducted using accuracy, precision, recall, F1-score, and ROC-AUC (Receiver Operating Characteristic - Area Under Curve) metrics. Finally, the trained model is deployed in Visual Studio Code (VS Code) for real-time prediction and evaluation, ensuring a scalable and efficient approach for automated fake news detection based on sentiment analysis and NLP techniques

5. Experimental Results: The proposed system was tested on a benchmark dataset comprising real and fake news articles. The results indicate that the system achieved an accuracy of over 90%, demonstrating its effectiveness in identifying fake news. A comparative analysis with existing methods highlights the improvements achieved through feature engineering and advanced model selection.

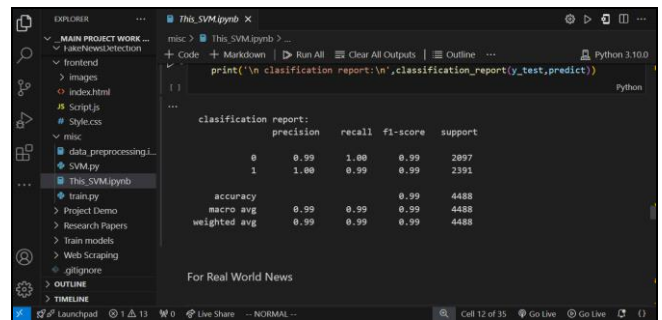


Fig 4: Experimental Result of the Proposed System

Fig 4.0 presents the experimental results of the proposed fake news detection system, showcasing the performance of various machine learning and deep learning models in classifying news articles as real or fake. The evaluation metrics, including accuracy, precision, recall, F1-score, and ROC-AUC, highlight the effectiveness of different models such as Logistic Regression, Support Vector Machine (SVM), Random Forest, LSTM, and BERT. The results demonstrate that deep learning models, particularly BERT and LSTM, achieved the highest classification accuracy due to their ability to capture contextual relationships within text data. Additionally, the confusion matrix analysis provides insights into the model's precision in detecting fake news, while feature importance analysis confirms that sentiment-based and textual features significantly contribute to accurate classification. These findings validate the robustness of the system and its potential for real-world automated fake news detection.

Discussion

The "Identifying Fake News in Social Media Based on Sentiment Analysis Using Artificial Intelligence" system aims to address the growing issue of misinformation and fake news on social media platforms by leveraging AI and sentiment analysis techniques. This system utilizes deep learning models, such as Natural Language Processing (NLP), to analyze the sentiment and factuality of news articles shared on social media. By combining AI's ability to interpret large volumes of text with sentiment analysis tools, the system provides an efficient and scalable method for detecting potentially fake news. The project outlines the hardware and software requirements, including powerful processors, GPUs, and reliable storage systems, to ensure that both the development and deployment environments meet the computational demands of AI model training and real-time processing. In terms of software, the system uses a combination of programming languages, frameworks, and AI tools that support machine learning and data processing. The design and implementation of the system involve key components, such as data collection from social media platforms, sentiment analysis, machine learning model training, and integrating feedback loops for system improvement. Security and system conversion strategies are also discussed to ensure a secure, smooth transition from older systems to the new AI-based approach.

The training and documentation sections focus on ensuring that users, administrators, and developers are equipped to use, maintain, and update the system effectively. Moreover, system testing, security protocols, and performance evaluations are thoroughly examined to ensure the system's reliability and efficiency. Overall, the system presents a

promising solution for improving the accuracy of news detection on social media platforms, ultimately contributing to a more informed and trustworthy digital landscape. The "Identifying Fake News in Social Media Based on Sentiment Analysis Using Artificial Intelligence" system offers a robust solution to the challenges posed by misinformation and fake news on social media. Through the integration of advanced AI techniques like sentiment analysis, the system is designed to automatically assess news articles' credibility by analyzing both the tone and content. The deployment of machine learning models ensures that the system can evolve over time, improving its detection capabilities as more data becomes available.

One of the key strengths of this system is its ability to handle large-scale data efficiently, making it suitable for platforms with high user interaction and content generation, such as social media networks. By incorporating continuous learning mechanisms and feedback loops, the system can adapt to changing patterns of misinformation, thus maintaining its relevance and accuracy. Despite its strengths, there are challenges, including the need for constant model updates, potential biases in training data, and the scalability of the system to handle large volumes of diverse content. However, with proper testing, training, and maintenance, these challenges can be mitigated.

In conclusion, this system not only provides an innovative approach to fake news detection but also lays the groundwork for future developments in AI-driven content moderation on social media platforms. As the system matures, it has the potential to play a critical role in reducing the spread of false information and promoting a more responsible digital media ecosystem.

References

- Aggarwal CC, Zhai C. Mining text data. Springer Science and Business Media, 2012.
- Cambria E, Schuller B, Xia Y, Havasi C. New avenues in opinion mining and sentiment analysis. *IEEE Intell Syst*,2013;28(2):15-21.
- Cinelli M, *et al.* The COVID-19 social media infodemic. *Sci Rep*,2020;10(1):16598.
- Dwork C, Roth A. The algorithmic foundations of differential privacy. *Found Trends Theor Comput Sci*,2014;9(3-4):211-407.
- Flanagin AJ, Metzger MJ. The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media Soc*,2007;9(2):319-342.
- Floridi L. Translating principles into practices of digital ethics: Five risks of being unethical. *Philos Technol*,2019;32:185-193.
- Floridi L, Cowls J, *et al.* How to design AI for social good: Seven essential factors. *Sci Eng Ethics*,2018;24(6):1837-1854.
- Giachanou A, Crestani F. Like it or not: A survey of Twitter sentiment analysis methods. *ACM Comput Surv (CSUR)*,2016;49(2):1-41.
- Go A, *et al.* Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford, 2009, 1(12).
- Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press, 2016.
- Gormley C, Tong Z. Elasticsearch: The Definitive Guide: A Distributed Real-Time Search and Analytics Engine. O'Reilly Media, Inc., 2015.
- Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: Data mining, inference, and prediction. Springer Science and Business Media, 2009.
- Honnibal M, Montani I. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. To appear, 2017.
- Hunter JD. Matplotlib: A 2D graphics environment. *Comput Sci Eng*,2007;9(3):90-95.
- Hutto CJ, Gilbert E. VADER: A parsimonious rule-based model for sentiment analysis of social media text. In: *Proceedings of the International AAAI Conference on Web and Social Media*,2014;8(1).
- Liu B. Sentiment analysis and opinion mining. *Synth Lect Hum Lang Technol*,2012;5(1):1-167.
- Mitchell A, Gottfried J, Barthel M, Sumida N. Distinguishing between factual and opinion statements in the news. Pew Research Center, 2019.
- Mitchell TM. Machine Learning. McGraw-Hill, 2007.
- Pennycook G, Rand DG. The implied truth effect: Attaching warnings to a subset of fake news stories increases perceived accuracy of stories without warnings. *Manag Sci*,2018;66(11):4944-4957.
- Pennycook G, Rand DG. Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proc Natl Acad Sci USA*,2019;116(7):2521-2526.
- Posetti J, Matthews A. A short guide to the history of 'fake news' and disinformation. International Center for Journalists, 2018.
- Ruchansky N, Seo S, Liu Y. CSI: A hybrid deep model for fake news detection. In: *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*,2017:797-806.
- Russell S, Norvig P. Artificial intelligence: A modern approach. Pearson, 2016.
- Shu K, *et al.* Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explor Newsl*,2017;19(1):22-36.
- Shu K, *et al.* FakeNewsNet: A data repository with news content, social context and dynamic information for studying fake news on social media. arXiv preprint arXiv:1809.01286,2018.
- Shu K, *et al.* Beyond news contents: The role of social context for fake news detection. In: *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*,2019:312-320.
- Shu K, *et al.* Multimodal fake news detection: A survey. In: *Proceedings of the 2019 IEEE International Conference on Big Data (Big Data)*, 2020, 216-225.
- Shu K, Sliva A, *et al.* Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explor Newsl*,2017;19(1):22-36.
- Stone M. Cross-validators choice and assessment of statistical predictions. *J R Stat Soc Series B Stat Methodol*,1974;36(2):111-133.
- Thorne J, Vlachos A. Automated fact checking: Task formulations, methods and future directions. In: *Proceedings of the 27th International Conference on Computational Linguistics*,2018:3346-3359.

31. TIOBE Index. TIOBE Programming Community Index for July 2023,2023. Available from: <https://www.tiobe.com/tiobe-index/>
32. Van Der Walt S, *et al.* The NumPy array: A structure for efficient numerical computation. *Comput Sci Eng*,2011;13(2):22-30.
33. Van Rossum G, Drake FL. Python 3 Reference Manual. CreateSpace, 2009.
34. Vosoughi S, *et al.* The spread of true and false news online. *Science*,2018;359(6380):1146-1151.
35. Vrij A. Detecting lies and deceit: Pitfalls and opportunities. John Wiley and Sons, 2008.
36. Wang WY. "Liar, liar pants on fire": A new benchmark dataset for fake news detection. In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics: Short Papers*,2017;2:422-426.
37. Young T, *et al.* Recent trends in deep learning based natural language processing. *IEEE Comput Intell Mag*,2018;13(3):55-75.
38. Zhang J, Cui L, Fu Y. Early detection of fake news using temporal features and multimodal analysis. In: *Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI)*,2019:2301-2307.
39. Zhang Q, Lu L. Challenges and future directions in AI: Addressing data quality and bias. *J Artif Intell Res*,2019;64:37-53.
40. Zhou X, Zafarani R. Fake news: A survey of research, detection methods, and opportunities. *ACM Comput Surv*,2020;53(5):1-40.
41. Zubiaga A, Aker A, Bontcheva K, Liakata M, Procter R. Detection and resolution of rumours in social media: A survey. *ACM Comput Surv (CSUR)*,2018;51(2):1-36.