

Course Title: Research Methodology in Computer Science

Course Code: AI800

Weighting: 20%

Lecturer: Dr. Mohammed Anbar

Hand-out Date: 21 Feb 2026

Hand-in Date: 14 March 2026

Hand-in Method: Online Submission via E-Learning.

Assignment Report: literature Review Fake News Detection Using Machine Learning

Table of contents:

- Summary research paper's
- Critical Analysis
- Connections
- Conflicts between the findings
- Synthesis Matrix for Literature Review
- References

Summary research paper's:

1- Fake News Detection on Social Media: A Data Mining Perspective.

Fake news on social media is a growing problem, but on this study offer to light on how to combat it. What they discovered is that social media platforms have accelerated the spread of misinformation like wildfire. To combat this, they say, we need to examine both the substance of the news as well as the social setting in which it's disseminated. This includes understanding the language, tone and patterns of text used in the news, as well as user engagement with that news likes, shares, comments and spread of the news through networks. Their framework integrates machine learning and social network analysis to mitigate fake However, they also highlight some key challenges, such as catching misinformation early, dealing with limited data, and staying one step ahead of fake news creators who are constantly adapting. emphasize the need for a multidisciplinary approach that brings together data mining, natural language processing, and social network analysis. It's important for future research in fake news detection, and its influence can be seen in many subsequent studies. By understanding how fake news spreads and how to detect it.

2- “Liar, Liar Pants on Fire”: A New Benchmark Dataset for Fake News Detection:

In this paper work on fake news and its data base. The dataset includes over twelve thousand short political statements sourced from real-world fact-checking repositories. Each statement was marked up with six levels of truthfulness, from “true” to “pants on fire,” so researchers could conduct fine-grained classification tasks. The main goal of this work was to yield a consistent and benchmark dataset for assessing fake news detection models. We also analyzed the results of various machine learning algorithms including Logistic Regression, Support Vector Machines (SVM), and neural network models. With experimental results showing that machine learning techniques generally achieve high accuracy to classify statements by their truthfulness levels. Apart from the actual data, the LIAR dataset contains metadata like speaker, party and context that can be utilised for classification enhancing. This dataset has been pivotal to help advance research surrounding misinformation detection. Thus, this dataset has been used by many recent studies as a benchmark for comparing and assessing various machine learning and deep learning models.

3- Detecting Fake News using Machine Learning: A Systematic Literature Review:

this research using conventional machine learning approaches to identify internet falsehoods Researchers were focused on the text of news articles in order to search for generic misinformation. Using n-gram analysis techniques, relevant and unique patterns were extracted from large datasets of news. These features were input to various machine learning classifiers (for example Naive Bayes, SVM or Decision Trees). These findings show that if provided with the right features of the text, traditional machine learning can classify fake and real news accurately. They assess these algorithms' performance based on experiment tests on large generated datasets. The results showed that the models achieved very high accuracy over the misleading content detection. Foehner et al. also highlighted the significance of feature selection and preprocessing for achieving good classification performance. The research concluded that techniques, which were inspired by classical machine learning algorithms, can be useful when attempting to detect fake news.

4- Fake News Detection on Social Media using Geometric Deep Learning:

When it comes to spotting fake news, researchers have been looking at new methods. A specific approach to solve this problem was proposed in 2017 by Federico Monti and his collaborators through graph-based deep learning methods. That is, their study is not limited to the textual analysis of news/article but also focus on how news propagates through social networks and users interact with each other. The researchers applied geometric deep learning to study the relationships between users, posts, and news sources to understand how information is propagated in this type of network. They achieved this by making graphs of social media interactions, which captured the manner in which information spreads. The team then used graph neural networks to study these patterns and detect misinformation. Their experiments demonstrated that fake news prediction task is easily influenced by surrounding network structure information. User behavior and community interactions hold important clues about misinformation, which are lost in common content-level analysis, the study said. It improves our understanding of how fake news is propagated online. In conclusion, the research shows that using social network analysis in combination with deep learning approaches can be an effective approach to detect fake news. Focusing on how news propagates through social networks can reveal the problem's richness and suggest solutions that work better.

5- FNDNet: A Deep Convolutional Neural Network for Fake News Detection:

They proposed FNDNet, a deep learning model tailored to fake news detection. In this approach, a model based on Convolutional Neural Networks (CNN) is used to extract semantic and contextual information from normalized text data. Data up to October 2023 [27]. The convolutional layers learn meaningful patterns and features from the input text, and they are passed to a classification layer that predicts whether the news is real or fake. The performance of the model was assessed by using benchmark datasets which are widely used in fake news detection literature. The experimental results demonstrated high accuracy of FNDNet and compared with several traditional machine learning methods, the performance was better. This paper shows that the deep learning models are capable of capturing complicated patterns in text. the contribution of CNN architectures for automatic detection of misinformation is also observed in the study. the proposed model is successful in enhancing the performance of fake news detection systems.

6- A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities:

This survey covers the general detection methods as well as the comprehensive overview of fake news research. In this overview, the authors discuss hundreds of articles that cover bot detection and related topics like fake news detection on social media platforms. They classify the approaches to fake news detection into content based methods and social context based methods, with a third class of hybrid approaches. This can be a linguistic feature based or writing pattern analysis in news articles. The approaches under social context are based on user interactions, sharing behavior, and the propagation of network patterns. Hybrid methods employ both kinds of features together to enhance detection accuracy. The survey also touches upon some of the core challenges facing this area, including changing tactics in misinformation campaigns, a lack of labeled datasets, and issues with detecting satire or sarcasm. In addition, the authors point out that interdisciplinary research combining machine learning, social science, and network analysis is essential. This study also highlights the area of potential future research for better fake news detection systems. This survey provides a reference which summarizes the existing knowledge, and offers guidance for future efforts in misinformation detection.

Critical Analysis

Random forests, decision trees, SVM and a variety of deep learning techniques are identified as the most effective machine learning methods for fake news detection in the studies analysed. Every one of these studies advances our understanding, but all have limitations that reflect gaps that need to be filled with additional research.

A framework relies on combining content based features and social context data to detect fake news shared in social media platforms. This work is stronger for its holistic approach that considers both textual features and user behavior. By studying the spread of information through networks an understanding of how misinformation subsequently spreads emerges. But this only offers a conceptual proposal and does not describe the implementation of an actual system, making it impossible to verify its efficacy in practice.

the introduction of a new dataset: LIAR, which quickly became the defector benchmark dataset in fake news detection research. The primary benefit of this work is that it provides a common dataset that enables researchers to assess and contrast different machine learning models. the dataset mainly consists of brief political sound bites, which could limit its relevance to other types of news text or domains such as health care, business, and entertainment.

Detecting fake news is an extensive area in itself and through n-gram analysis and the application of textual features, traditional machine learning techniques have shown effective results as per the research by Hadeer Ahmed, Issa Traore and Sherif Saad. The benefit of this methodology is that it is simple and computationally cheap. So far, conventional machine learning models are based on structured features and often lack the ability to capture high-level semantic relationships within text adequately in contrast to latest deep learning models.

The proposed approach by Federico Monti et al. is a remarkable step forward in that it leverages graph-based deep learning techniques for the task of fake news detection. Their model examines the features of social networks and diffusion on news content. So, we need to use relational data containing the relationship between users and information sources so that it would boost detection accuracy. But the model relies on large-scale social network data, which is not always available due to privacy issues or platform restrictions.

The deep learning model FNDNet by Rohit Kaliyar et al shows how CNNs are able to learn semantic patterns out of textual information. The authors report improved classification accuracy over several traditional machine learning techniques. These attributes make deep learning models more advantageous, but they still depend on large labeled datasets and

consume increased computational resources for training compared to other machine learning methods, which can hinder their practicability in certain cases.

Finally, in the survey by Xinyi Zhou and Reza Zafarani they offer a comprehensive overview of the research in detection of fake news grouping methods into three classes: content-based, social context based, and hybrid approaches. Although the study provides useful theoretical contributions and important obstacles within this field, it fails to offer a new model or experimental validation. Rather, its primary contribution is as a synthesis of current research and setting forth possible avenues for future studies.

Overall, these studies show together that fake news

Connections

The reviewed studies share several common themes and methodological connections. Most of the research emphasizes the importance of machine learning techniques for automating fake news detection. Early works such as those by Hadeer Ahmed and colleagues focus primarily on traditional machine learning algorithms using textual features extracted from news articles. These studies demonstrate that linguistic patterns can provide valuable indicators for distinguishing fake news from legitimate content.

Later research builds upon these foundations by incorporating more advanced approaches. For example, the work of Rohit Kaliyar and colleagues demonstrates how deep learning models, particularly CNN architectures, can automatically learn complex semantic representations from textual data. This progression reflects a broader shift in the field toward deep learning techniques for improved classification performance.

Another important connection among the studies is the recognition that textual content alone may not be sufficient for reliable fake news detection. The framework proposed by Kai Shu and colleagues highlights the importance of incorporating social context features, such as user interactions and information propagation patterns. Similarly, the work of Federico Monti further expands this idea by using graph-based deep learning models to analyze the spread of information across social networks.

Conflicts Between the Findings

The reviewed studies share several common themes and methodological connections. Most of the research emphasizes the importance of machine learning techniques for automating fake news detection. Early works such as those by Hadeer Ahmed and colleagues focus primarily on traditional machine learning algorithms using textual features extracted from news articles. These studies demonstrate that linguistic patterns can provide valuable indicators for distinguishing fake news from legitimate content.

Later research builds upon these foundations by incorporating more advanced approaches. For example, the work of Rohit Kaliyar and colleagues demonstrates how deep learning models, particularly CNN architectures, can automatically learn complex semantic representations from textual data. This progression reflects a broader shift in the field toward deep learning techniques for improved classification performance.

Another important connection among the studies is the recognition that textual content alone may not be sufficient for reliable fake news detection. The framework proposed by Kai Shu and colleagues highlights the importance of incorporating social context features, such as user interactions and information propagation patterns. Similarly, the work of Federico Monti further expands this idea by using graph-based deep learning models to analyze the spread of information across social networks.

In summary, while the reviewed studies adopt different methodologies and datasets, they collectively demonstrate that hybrid approaches integrating textual analysis, social context, and advanced machine learning techniques are likely to provide the most effective solutions for fake news detection.

Synthesis Matrix for Literature Review:

Study	Authors	Year	Research Design	Key Findings/Results	Methodology	Limitations
Fake News Detection on Social Media: A Data Mining Perspective	Shu et al.	2017	ML + Social Context	Proposes a data mining framework combining content-based features and social context information. Analyzes textual data along with user interactions such as likes, shares, and propagation patterns on social media.	Combining textual analysis with social network context improves fake news detection performance and provides a deeper understanding of how misinformation spreads online.	Mainly conceptual framework with limited experimental implementation; relies on social media interaction data that may not always be accessible.
'Liar, Liar Pants on Fire': A New Benchmark Dataset for Fake News Detection	Wang	2017	LR, SVM, Neural Networks	Introduces the LIAR dataset containing thousands of labeled political statements with six levels of truthfulness. Evaluates models such as Logistic Regression, Support Vector Machines, and Neural Networks.	Provides a widely used benchmark dataset that enables consistent evaluation of fake news detection models and comparison between algorithms.	Dataset focuses mainly on short political statements, limiting its generalization to other domains or long-form news articles.
Detection of Online Fake News Using N-gram Analysis and Machine Learning Techniques	Ahmed et al.	2017	Naive Bayes, SVM	Applies traditional machine learning methods including Naive Bayes, Support Vector Machines, and Decision Trees. Uses n-gram text analysis to extract linguistic features from news articles.	Shows that classical machine learning techniques combined with textual features can effectively classify fake and real news.	Relies heavily on handcrafted features and may struggle with deeper semantic or contextual meaning in complex texts.
Fake News Detection on Social Media Using	Monti et al.	2019	Graph Neural Networks	Uses graph-based deep learning with geometric deep learning and graph neural networks to model	Demonstrates that analyzing information propagation and social network structure	Requires large-scale social network interaction data which may be difficult to

Geometric Deep Learning				relationships between users, posts, and news sources.	improves detection accuracy compared with content-only approaches.	obtain due to privacy and access limitations.
A Deep Convolutional Neural Network for Fake News Detection	Kaliyar et al.	2021	CNN (FNDNet)	Introduces FNDNet based on Convolutional Neural Networks (CNN) with word embeddings to capture semantic patterns in textual data.	Achieves high accuracy on benchmark datasets and shows that deep learning approaches outperform many traditional machine learning models.	Requires large labeled datasets and high computational resources; model interpretability can be limited.
A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities	Zhou & Zafarani	2020	Survey Study	Comprehensive survey categorizing fake news detection approaches into content-based, social context-based, and hybrid methods.	Provides a structured overview of fake news detection techniques and highlights major research challenges and future opportunities.	Does not propose a new detection model or experimental results since it focuses on reviewing existing literature.

References

- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media: A Data Mining Perspective. ACM SIGKDD Explorations. <chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://arxiv.org/pdf/1708.01967>
- Wang, W. Y. (2017). 'Liar, Liar Pants on Fire': A New Benchmark Dataset for Fake News Detection. ACL. <chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://arxiv.org/pdf/1705.00648>
- Ahmed, H., Traore, I., & Saad, S. (2017). Detection of Online Fake News Using N-gram Analysis and Machine Learning Techniques. <chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://arxiv.org/pdf/2102.04458>
- Monti, F., Frasca, F., Eynard, D., Mannion, D., & Bronstein, M. (2019). Fake News Detection on Social Media using Geometric Deep Learning. <chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://arxiv.org/pdf/1902.06673>
- Kaliyar, R., Goswami, A., & Narang, P. (2021). FNDNet: A Deep Convolutional Neural Network for Fake News Detection. <https://www.sciencedirect.com/science/article/abs/pii/S1389041720300085>
- Zhou, X., & Zafarani, R. (2020). A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. <chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://arxiv.org/pdf/1812.00315>