

Enhancing Fake News Detection Accuracy through Hybrid Machine Learning and Social Context Analysis

Student Name: Marwa Mohammed Almahmoud

Course Title: Research Methodology in Computer Science

Course Code: AI800

Weighting: 50%

Lecturer: Dr. Mohammed Anbar

Hand-out Date: 14 March 2026

Hand-in Date: 14 April 2026

Submission Method: Online via E-Learning

Table of Contents

1. Abstract
2. Chapter 1 – Introduction and Research Overview
 - 1.1 Introduction
 - 1.2 Background
 - 1.3 Problem Statement
 - 1.4 Research Objectives
 - 1.5 Research Scope
3. Chapter 2 – Literature Review
4. Chapter 3 – Research Methodology
 - 3.1 Research Design
 - 3.2 Data Collection and Preparation
 - 3.3 Model Development
 - 3.4 Expected Outcomes
 - 3.5 Research Planning and Gantt Chart
 - 3.6 Ethical Considerations
5. References

Abstract

Misinformation on digital platforms has become one of the most persistent challenges of the information age. While machine learning techniques have shown potential for detecting fake news through textual analysis, they rarely consider the social context in which information spreads. This research proposes a hybrid fake news detection framework that integrates linguistic analysis and social network characteristics to improve accuracy and adaptability. Using benchmark datasets such as *LIAR* and public social-media data, the study will extract semantic, sentiment, and propagation features. Deep-learning and graph-based models will be combined for feature fusion and classification. Expected outcomes include a more reliable detection model and practical contributions toward ethical, privacy-aware misinformation control systems.

Chapter 1 – Introduction and Research Overview

1.1 Introduction

The evolution of the internet and social media has transformed how information flows within society. However, the same speed and openness that allow rapid communication have also enabled the uncontrolled spread of fake news false information created to mislead. False content influences elections, public health decisions, and social trust. Researchers have responded by exploring automated detection systems powered by machine learning and deep learning, capable of distinguishing between legitimate and deceptive content. Yet, most existing models largely depend on textual features, leaving a gap in understanding how social engagement patterns contribute to the diffusion of misinformation.

1.2 Background

Early fake-news research primarily relied on linguistic cues such as word frequency, sentiment, and n-gram models (Ahmed et al., 2017). Later studies introduced benchmark datasets like *LIAR* (Wang 2017) that advanced comparative evaluation between algorithms. The emergence of deep learning, particularly Convolutional Neural Networks (Kaliyar et al., 2021), improved accuracy by automatically capturing semantic relationships within text. However, content analysis alone cannot fully explain misinformation spread. Shu et al. (2017) and Monti et al. (2019) demonstrated that integrating social-context information, such as repost patterns and network graphs, can reveal how fake content propagates. Despite these insights, limitations remain regarding dataset diversity, privacy restrictions, and fairness within automated classification.

1.3 Problem Statement

The growing influence of social media has magnified the scale and impact of fake news. Although machine learning has improved detection accuracy, most models continue to rely mainly on textual features, ignoring how information circulates through online communities. Studies confirm that user interactions—likes, comments, and shares—provide critical signals for identifying misinformation (Shu et al., 2017; Monti et al., 2019). Existing benchmark datasets, including LIAR (Wang 2017), are limited to brief, domain-specific statements and lack contextual metadata. This restricts model generalization and reduces reliability in dynamic social networks. Hence, there is a clear research gap in integrating content-based and social-context-based approaches into one practical framework capable of detecting misinformation across multiple domains.

This project addresses that gap by developing and evaluating a hybrid model that combines textual and social signals using machine learning techniques. To build real trust in automated detection systems, it is equally important to handle user data responsibly and reduce any bias that might influence the model's decisions.

1.4 Research Objectives

Main Objective

To design and evaluate a hybrid machine learning model that integrates textual and social-context features for enhanced fake news detection accuracy.

Specific Objectives

1. Review and synthesize prior research on content-based, context-based, and hybrid fake news detection methods.
2. Examine existing benchmark datasets (e.g., *LIAR*) and analyze linguistic, semantic, and social features relevant to misinformation identification.
3. Design a hybrid model combining text features (n-grams, sentiment, writing style) with social-network metrics (user engagement, propagation patterns).
4. Implement and test the model using ethically sourced data, ensuring privacy protection and bias mitigation.
5. Compare model performance with traditional and deep-learning baselines using standard evaluation metrics.
6. Discuss ethical implications, including user privacy, data fairness, and responsible AI use in misinformation detection.

1.5 Research Scope

This research focuses on English-language online news and social media content. It will primarily use publicly available datasets, including *L/IAR* and selected portions of Twitter or Reddit data accessed through open APIs. The study excludes multimedia misinformation (images/videos) and languages other than English. The work scope covers model design, implementation, evaluation, and ethical reflection—not large-scale deployment. All used data will respect privacy and anonymization standards.

Chapter 2 – Literature Review

2.1 Overview

Research on fake news detection has grown rapidly as misinformation continues to spread across online platforms. Different approaches have been proposed, ranging from traditional machine-learning methods to deep-learning and network-based models. This chapter reviews ten key studies related to fake news detection. The review includes the six studies examined in Assignment 1 and four additional papers, as required by the project. The goal is to highlight existing methods, identify research gaps, and connect findings to the objectives of this proposal.

2.2 Review of Existing Studies

Shu et al. (2017)

Shu and colleagues proposed one of the earliest comprehensive frameworks combining content-based and social-context signals for detecting fake news on social media. Their work emphasized that misinformation should not be analyzed from text alone, since social interactions—such as resharing patterns, user behavior, and propagation networks—offer crucial cues. They outlined a data-mining perspective that integrates natural language processing with social network analysis. A limitation of their work is that it mainly presents a conceptual framework without implementing a fully functional detection model, making it difficult to evaluate its real-world performance.

Ahmed et al. (2017)

This study focused on text-based machine-learning methods, using n-gram features combined with algorithms such as Support Vector Machines (SVM), Naïve Bayes, and Decision Trees. Their results showed that classical machine-learning techniques could classify fake and real news with strong accuracy when provided with well-engineered features. The strength of this work lies in its simplicity and low computational cost. However, the study does not capture deeper semantic relationships in text, which limits performance when dealing with subtle misinformation or linguistically complex content.

Wang (2017)

Wang introduced the LIAR dataset, which contains over 12,000 short political statements labeled across six truthfulness categories. The study evaluated several machine-learning models using this dataset and demonstrated the usefulness of benchmark data for comparing algorithms. The LIAR dataset has become widely used, but its narrow focus on political statements limits its

generalizability. It lacks article-level text and does not provide social-interaction information, making it insufficient for context-aware detection models.

Monti et al. (2019)

Monti and colleagues applied geometric deep learning and Graph Neural Networks (GNNs) to detect fake news by modeling how information spreads across social networks. Their method represented users, posts, and interactions as graph structures and used graph-based deep learning to capture propagation patterns. The study showed that including network information significantly improves predictive accuracy. However, this approach requires large amounts of social-network data, which is often restricted due to privacy policies.

Kaliyar et al. (2021)

This study introduced FNDNet, a deep convolutional neural-network model designed to extract semantic patterns from news text. By using word embeddings and multiple convolutional layers, FNDNet achieved higher accuracy than several traditional machine-learning methods. The strength of this model lies in its ability to automatically learn meaningful features from text. However, it depends on large labeled datasets and does not incorporate social-network information, making it purely content-based.

Zhou & Zafarani (2020)

This comprehensive survey reviewed hundreds of fake-news detection studies and categorized them into content-based, context-based, and hybrid approaches. The authors highlighted challenges such as evolving misinformation tactics, lack of labeled data, and difficulty in detecting satire. They also identified major research gaps, including limited integration of content and context signals. Although the survey offers valuable theoretical insights, it does not propose or evaluate a new detection model.

Tacchini et al. (2019)

Tacchini and colleagues examined a hybrid detection method that combined content signals with user credibility scores. Their findings showed that misinformation often circulates within specific user communities, and credibility-based features can significantly improve performance. This supports the idea that social behavior is essential for accurate detection. However, user credibility metrics may raise ethical concerns, such as mislabeling users based on limited activities.

Ajao et al. (2018)

This study used Recurrent Neural Networks (RNNs) to capture temporal patterns in fake-news propagation on Twitter. The researchers demonstrated that analyzing how posts spread over time can enhance prediction accuracy. Their work contributes to understanding the

time dimension of misinformation diffusion. The limitation is that RNNs require large and clean temporal datasets, which are often difficult to obtain.

Bovet & Makse (2019)

Bovet and Makse analyzed large-scale Twitter data during the 2016 US election, focusing on how fake news spreads through information diffusion networks. They showed that propagation graphs reveal early signals of misinformation long before text analysis can detect it. Their study highlights the power of network-based methods but also raises privacy concerns as it uses user interaction data.

Horne & Adali (2017)

This study explored writing-style and structural differences between real and fake news. They found that fake news tends to use emotionally charged language, simpler structure, and exaggerated claims, whereas real news contains more factual density. This study provides useful linguistic insights, although stylistic differences may vary across topics, cultures, and platforms.

2.3 Synthesis and Gap Analysis

Across the reviewed literature, several key themes emerge:

- Content-based approaches (Ahmed et al., Kaliyar et al., Horne & Adali) perform well but fail to consider the social dynamics driving misinformation.
- Context-based and graph approaches (Shu et al., Monti et al., Bovet & Makse) provide deeper insights but depend heavily on social-media interaction data, which is often restricted.
- Hybrid models (Tacchini et al.) show promise but still lack comprehensive integration of both semantic and propagation-based features.
- Datasets like *LIAR* remain limited in scope, affecting real-world applicability.

These gaps justify the need for a hybrid model that combines both textual and social-context features — aligning directly with the aims of this research.

2.4 Synthesis Matrix

Study	Methodology	Key Findings	Limitations
Shu et al. (2017)	Content + social framework	Social signals improve detection	Conceptual only, not implemented
Ahmed et al. (2017)	N-gram + classical ML	Good accuracy with simple features	Misses deep semantics
Wang (2017)	LIAR dataset	Standard benchmark	Limited to political short text
Monti et al. (2019)	Graph Neural Networks	Network patterns enhance detection	Requires extensive user data
Kaliyar et al. (2021)	CNN (FNDNet)	Strong semantic learning	Text-only, no context
Zhou & Zafarani (2020)	Survey study	Hybrid methods promising	No experimental model
Tacchini et al. (2019)	Hybrid (credibility + text)	Credibility improves performance	User-labeling ethics issues
Ajao et al. (2018)	RNN temporal modeling	Time-based spread improves detection	Hard to obtain temporal data
Bovet & Makse (2019)	Diffusion network analysis	Early signals in propagation	Privacy restrictions
Horne & Adali (2017)	Style-based analysis	Fake news uses emotional style	Style varies across domains

2.5 Summary

The overall trend in the literature shows a shift from text-only models to more context-aware approaches that incorporate user behavior and network interactions. However, no existing study provides a balanced model that integrates both deep semantic understanding and social-context features while also addressing ethical concerns such as bias and privacy. This research proposal aims to fill that gap by developing a hybrid machine-learning model capable of leveraging both linguistic and social information for improved fake-news detection.

Chapter 3 – Research Methodology

3.1 Research Design

This study will adopt an experimental quantitative approach using secondary datasets. It involves four phases:

1. Data collection and preprocessing.
2. Feature extraction from textual and social-context dimensions.
3. Hybrid model development (ML + graph-based learning).
4. Model evaluation and comparison with baselines.

3.2 Data Collection and Preparation

- Datasets: *LIAR* (Wang 2017) and open-source Twitter data.
- Text features: Tokenization, n-grams, TF-IDF, sentiment scores.
- Social features: Number of shares, user network degree, propagation depth.
- Tools: Python 3, scikit-learn, TensorFlow, and NetworkX for graph features.
- Ethics: Anonymize user data; remove personal identifiers; follow platform policies.

3.3 Model Development

- Baseline models: Naïve Bayes, SVM, CNN (FNDNet).

- Proposed hybrid model: Combine CNN text encoder with Graph Convolutional Network (GCN) capturing propagation structure.
- Training/Testing split: 80/20 with cross-validation.
- Performance metrics: Accuracy, Precision, Recall, F1-Score, Fairness Index.

3.4 Expected Outcomes

- A hybrid ML framework that significantly improves fake-news classification performance.
- Empirical evidence supporting integration of textual and contextual features.
- Insights into ethical AI design through privacy preservation and bias mitigation.
- Practical contribution: a conceptual blueprint for trustworthy misinformation-detection systems.

3.5 Research Planning and Gantt Chart

Activity	Mar 14–Mar 31	Apr 1–Apr 14	Apr 15–Apr 30	May 1–May 15
Topic refinement				
Literature review				
Data collection				
Preprocessing				
Model design				
Model testing				
Results analysis				
Final writing				

Research Stage / Activity	Tasks Description	Duration	Timeline (2026)	Expected Deliverables
Phase 1: Proposal Preparation	- Refine research title and topic - Review Assignment 1 & 2 feedback - Outline proposal structure	1 week	Mar 14 – Mar 21	Updated title, structured proposal outline
Phase 2: Extended Literature Review	- Collect and review 4 new studies in addition to the 6 from Assignment 1	1 week + 2 days	Mar 18 – Mar 29	Completed Chapter 2 (Literature Review)

	- Summarize each paper - Create synthesis matrix and gap analysis			
Phase 3: Dataset Collection & Ethical Screening	- Obtain <i>LIAR</i> dataset - Identify supporting open-source social-media dataset (e.g., Twitter via API) - Check data licensing and privacy compliance	5 days	Mar 29 – Apr 2	Clean, ethically sourced dataset
Phase 4: Data Preprocessing & Feature Engineering	- Text cleaning, stop-word removal - Tokenization, TF-IDF vectorization - Social feature extraction (user engagement, propagation depth) - Data splitting (train/test)	6 days	Apr 3 – Apr 8	Ready-to-use training data for model
Phase 5: Model Design and Implementation	- Build baseline ML models (SVM, Naïve Bayes, CNN prototype) - Develop hybrid CNN + GCN model - Integrate content & social features in fusion layer	7 days	Apr 9 – Apr 15	Hybrid fake news detection model
Phase 6: Model Testing and Evaluation	- Run experiments with cross-validation - Compare accuracy, precision, recall, F1-score vs baselines - Conduct bias and fairness checks	5 days	Apr 16 – Apr 20	Model performance results + evaluation tables
Phase 7: Results Analysis and Discussion	- Interpret results against objectives -	4 days	Apr 21 – Apr 24	Draft Chapter 3 (Results & Discussion)

	Discuss findings with reference to literature - Identify limitations and future work			
Phase 8: Documentation and Final Proposal Submission	- Write expected outcomes and Gantt chart sections - Integrate all chapters into final report - Proofread and format citations (APA style)	5 days	Apr 25 – Apr 30	Final research proposal submitted online

3.6 Ethical Considerations

All datasets will follow open-data policies. Personal identifiers will be removed, and analyses will include fairness checks to minimize algorithmic bias. No private or sensitive data will be stored.

References

- Ahmed, H., Traore, I., & Saad, S. (2017). *Detection of Online Fake News Using N-gram Analysis and Machine Learning Techniques*.
- Ajao, O., Bhowmik, D., & Zargari, S. (2018). *Fake News Identification on Twitter with Hybrid Deep Learning*. IEEE AFRICON.
- Bovet, A., & Makse, H. (2019). *Influence of Fake News in Twitter during 2016 US Election*. Nature Communications.
- Horne, B., & Adali, S. (2017). *This Just In: Fake News Packs a Lot*. ICWSM.
- Kaliyar, R., Goswami, A., & Narang, P. (2021). *FNDNet: A Deep Convolutional Neural Network for Fake News Detection*.
- Monti, F., Frasca, F., Eynard, D., Mannion, D., & Bronstein, M. (2019). *Fake News Detection on Social Media Using Geometric Deep Learning*.
- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). *Fake News Detection on Social Media: A Data Mining Perspective*.
- Tacchini, E., Ballarin, G., & De Angeli, A. (2019). *Some Like It Hoax: Automated Fake News Detection and Content Evaluation*. Information Sciences.
- Wang, W. Y. (2017). *'Liar, Liar Pants on Fire': A New Benchmark Dataset for Fake News Detection*.
- Zhou, X., & Zafarani, R. (2020). *A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities*.