



Using Deep Learning to Detect Arabic Fake News on Online Websites

Prepared By:

Areej Musaed Almalki

252001300

Submitted in fulfillment of the requirements for
the degree Master of Science in Artificial
Intelligence

At

College of Informatics

Midocean University

UAE

Supervisor: Dr. Yasser Abdelhamid

College of Informatics, Midocean University, Comoros

[2026]

Acknowledgements

The completion of this project would not have been possible without the guidance and support received throughout its development. Special appreciation is extended to the project supervisor, Dr. Yasser Abdulhamid, for the valuable feedback, continuous support, and constructive guidance provided during all stages of this work.

Gratitude is also expressed to the faculty members for their academic support and for providing a constructive learning environment. Appreciation is extended to the University for offering the necessary resources that facilitated the completion of this project.

Abstract

The expansion of online news platforms has reshaped how information is created and circulated in digital environments. While such platforms provide rapid access to news, they also allow misleading and fabricated content to spread with limited oversight. Fake news has therefore become a growing concern, as it may influence public opinion and weaken trust in reliable media sources. These challenges are more pronounced in Arabic-language news, where linguistic variation and complex structure add further difficulty to automated analysis.

Research on automatic fake news detection has increased in recent years as manual verification methods have proven impractical for large-scale digital content. Early computational approaches relied on rule-based techniques and handcrafted features, which showed limited effectiveness across diverse topics. Subsequent machine learning methods improved performance but often struggled with generalization.

More recent developments in deep learning enabled models to learn representations directly from text, improving contextual understanding. Nevertheless, handling long news articles and subtle semantic relations remained challenging. Transformer-based architectures introduced attention mechanisms that significantly enhanced text modeling capabilities. Although promising results have been reported, Arabic fake news detection research remains relatively limited, particularly in terms of dataset diversity and real-world evaluation.

This study provides a structured review of fake news detection approaches with emphasis on Arabic-language content. By examining methodological trends and identifying key limitations, this phase establishes a practical foundation for the implementation stage of an automated Arabic fake news detection system.

Contents

Acknowledgements.....	2
Abstract.....	3
List of Abbreviations	6
Chapter 1: Introduction	7
1.1 Problem Definition.....	7
1.2 Research Objectives.....	7
1.3 Research Contribution	8
1.4 Thesis Organization.....	8
Figure 1.1 Structure of the thesis showing the main chapters and conclusion.....	8
Chapter 2: Overview.....	9
2.1 Fake News Concept.....	9
2.2 Natural Language Processing in Fake News Detection	9
2.3 Machine Learning Approaches	9
2.4 Deep Learning Approaches	10
2.5 Transformer-Based Models	10
2.6 Arabic News Datasets	10
Chapter 3: Related Work.....	11
3.1 Reviewed Studies	11
Group 1: Traditional Machine Learning and Early Neural Models.....	11
Group 2: Deep Learning Models for Arabic Fake News Detection.....	11
Group 3: Transformer-Based and Advanced Architectures.....	11
3.2 Synthesis Matrix of Reviewed Studies	12
Table 3.1: Comparison Matrix of Fake News Detection Studies.....	12
3.3 Dataset-Oriented Comparison Matrix	12
Table 3.2: Dataset Comparison Matrix	13
3.4 Observations from the Matrix Analysis	13
3.5 Research Gaps and Opportunities	13
Conclusion	15
References	16

Figures:

Figure 1: Structure of the thesis	8
---	---

Tables:

Table 3.1: Comparison of Fake News Detection Studies	13
Table 3.2: Dataset Comparison Matrix.....	14

List of Abbreviations

Abbreviation	Description
AI	Artificial Intelligence
NLP	Natural Language Processing
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit
SVM	Support Vector Machine
NB	Naïve Bayes
RF	Random Forest
TF-IDF	Term Frequency–Inverse Document Frequency
BERT	Bidirectional Encoder Representations from Transformers
AraBERT	Arabic BERT Model
MARBERT	Masked Arabic BERT

Chapter 1: Introduction

1.1 Problem Definition

Online news platforms have become a primary source of information in modern society. These platforms allow rapid access to news content and enable continuous information sharing. However, the same environment has also contributed to the widespread dissemination of misleading and fabricated news. Fake news represents a serious challenge due to its ability to distort facts, influence public opinion, and reduce trust in reliable media sources.

The problem of fake news has gained increased attention with the growth of digital media and social networking platforms. The speed at which information spreads online makes manual verification impractical, especially when dealing with large volumes of news articles published daily. As a result, inaccurate or misleading content can reach a wide audience before being detected or corrected.

This challenge is more evident in Arabic-language news content. The Arabic language is characterized by rich morphology, diverse writing styles, and variations between formal and informal usage. These characteristics increase the difficulty of automated text analysis and limit the effectiveness of existing fake news detection techniques. In addition, the availability of large, reliable, and diverse Arabic fake news datasets remains limited, which further complicates the development and evaluation of effective detection models.

1.2 Research Objectives

The objectives of this study are as follows:

- **To review** existing fake news detection approaches reported in the literature, with emphasis on Arabic-language news content.
- **To analyze** machine learning, deep learning, and transformer-based techniques applied to fake news detection.
- **To identify** the main challenges associated with detecting fake news in Arabic online news platforms.
- **To classify** previous studies into methodological groups in order to highlight research trends.
- **To highlight** research gaps that motivate further investigation in Arabic fake news detection.

1.3 Research Contribution

This study provides a structured and organized review of fake news detection research with a specific focus on Arabic-language content. By classifying existing studies into methodological groups, this work offers a clear overview of how detection approaches have evolved over time. The study highlights the strengths and limitations of current techniques without relying on detailed experimental comparisons.

The contribution of this phase lies in identifying key research gaps related to Arabic news analysis, including data availability, linguistic complexity, and evaluation practices. These findings support the design of an automated Arabic fake news detection system to be developed in the subsequent project phase.

1.4 Thesis Organization

This report is organized into three main chapters. Chapter 1 presents the research problem, objectives, and contribution. Chapter 2 provides an overview of the main concepts and techniques related to fake news detection. Chapter 3 reviews and classifies relevant previous studies based on their methodological approaches.

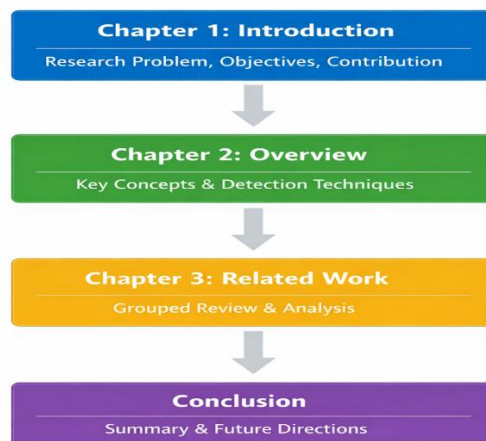


Figure 1.1 Structure of the thesis showing the main chapters and conclusion

Chapter 2: Overview

This chapter provides an overview of the key concepts and methodologies commonly used in fake news detection. The focus is on defining core terms, outlining major computational approaches, and highlighting techniques relevant to Arabic-language news analysis.

2.1 Fake News Concept

Fake news is commonly described as information that is deliberately false and presented in the form of news content. In practice, such information is often written to resemble credible journalism, which makes detection difficult. Previous research indicates that fake news may differ from legitimate news in tone, emotional appeal, and narrative consistency, although these differences are not always explicit.

Several studies have highlighted that fake news differs from credible news in terms of linguistic patterns, emotional tone, and content consistency (Vosoughi et al., 2018). These characteristics have been widely used as indicators in automated detection systems.

2.2 Natural Language Processing in Fake News Detection

Natural Language Processing focuses on enabling computers to analyze and interpret human language. Within fake news detection, NLP techniques are mainly used to prepare text for automated analysis and to extract patterns that support classification. Operations such as tokenization and normalization are typically applied before model training, as raw news text is rarely suitable for direct processing.

Common NLP operations include text normalization, tokenization, and removal of irrelevant elements. These steps help transform raw text into a structured format suitable for machine learning and deep learning models (Aggarwal, 2018). NLP plays a crucial role in enabling automated analysis of large-scale news data.

2.3 Machine Learning Approaches

Early fake news detection systems relied on traditional machine learning techniques combined with handcrafted textual features (Shu et al., 2019). These approaches typically use classifiers such as Support Vector Machines, Naïve Bayes, and Random Forest to distinguish between fake and real news.

While machine learning models demonstrated reasonable performance in controlled environments, their effectiveness often depends on the quality of feature engineering and dataset characteristics (Zhou & Zafarani, 2020). These limitations motivated the adoption of more advanced learning techniques.

2.4 Deep Learning Approaches

Deep learning models improved fake news detection by enabling automatic learning of textual representations directly from data (LeCun et al., 2015). Neural network architectures such as Convolutional Neural Networks and Recurrent Neural Networks have been widely applied to capture semantic and sequential patterns in news content.

Studies have shown that deep learning models outperform traditional machine learning methods in many text classification tasks, including fake news detection (Wang, 2017). However, these models still face challenges related to long text sequences and domain generalization.

2.5 Transformer-Based Models

Transformer-based models represent a significant advancement in natural language processing. These models rely on attention mechanisms to capture contextual relationships within text more effectively than previous architectures (Vaswani et al., 2017).

In fake news detection, transformer models such as BERT and its variants have achieved strong performance due to their ability to model long-range dependencies (Devlin et al., 2019). Arabic transformer models further improve results by leveraging large-scale Arabic corpora, making them suitable for analyzing Arabic news content (Antoun et al., 2020).

2.6 Arabic News Datasets

The performance of fake news detection systems depends heavily on the availability of reliable datasets. Arabic fake news datasets remain limited in size and diversity compared to those available for other languages (Himdi et al., 2025).

Existing Arabic datasets often focus on specific topics or platforms, which restricts comprehensive evaluation and model generalization (Shehata et al., 2024). This limitation highlights the need for further efforts in dataset construction and validation.

Chapter 3: Related Work

3.1 Reviewed Studies

The reviewed studies were classified into methodological groups based on the primary detection approach employed. This classification provides a clear comparison between different research directions and highlights methodological trends in fake news detection.

Group 1: Traditional Machine Learning and Early Neural Models

This group includes studies that rely on classical machine learning classifiers or early neural architectures combined with manually engineered features. These approaches demonstrate acceptable performance but limited contextual understanding.

Studies in this group:

- Park & Chai (2023)
- Singhanian et al. (2023)
- Zhou et al. (2023)

Group 2: Deep Learning Models for Arabic Fake News Detection

This group focuses on deep learning approaches applied to Arabic-language news using CNN, LSTM, and GRU architectures. These studies improve contextual modeling but remain limited by dataset size and domain specificity.

Studies in this group:

- Himdi et al. (2025)
- Shehata et al. (2024)
- Khan et al. (2024)

Group 3: Transformer-Based and Advanced Architectures

This group includes studies that adopt transformer-based models and advanced language architectures to enhance semantic understanding and detection accuracy.

Studies in this group:

- Bashaddadh et al. (2025)
- Matviychuk et al. (2024)
- Alam et al. (2024)
- Liu et al. (2025)

3.2 Synthesis Matrix of Reviewed Studies

Study	Year	Language	Model Type	Dataset Used	Key Focus
Park & Chai	2023	English	ML Classifiers	PolitiFact	User-centered detection
Singhania et al.	2023	English	CNN / Early DL	FakeNewsNet	Automatic detection
Zhou et al.	2023	English	Neural Network	Topic-specific dataset	Topic-based detection
Himdi et al.	2025	Arabic	CNN + LSTM	Arabic Fake News Dataset	Arabic dataset construction
Shehata et al.	2024	Arabic	Deep Learning	ArabFake Dataset	Multitask detection
Khan et al.	2024	Arabic / English	Hybrid DL	Mixed news datasets	Performance comparison
Bashaddadh et al.	2025	Multilingual	Review	Multiple datasets	Systematic review
Matviychuk et al.	2024	Multilingual	LLM-based	Fake content datasets	AI-based detection
Alam et al.	2024	Multilingual	Transformer	Cross-domain datasets	Domain adaptation
Liu et al.	2025	Chinese	Transformer	Chinese fake news dataset	Interpretability

Table 3.1: Comparison Matrix of Fake News Detection Studies

3.3 Dataset-Oriented Comparison Matrix

Study	Dataset Type	Dataset Source	Dataset Size (Approx.)	Domain
Park & Chai (2023)	Fact-checking	PolitiFact	Medium	Political
Singhania et al. (2023)	News articles	FakeNewsNet	Large	General
Zhou et al. (2023)	Topic-based	Custom dataset	Medium	Specific topics
Himdi et al. (2025)	Arabic news	Arabic Fake News Dataset	Small–Medium	News
Shehata et al. (2024)	Arabic news	ArabFake	Medium	News
Khan et al. (2024)	Mixed	Combined datasets	Medium	Multi-domain
Bashaddadh et al. (2025)	Multiple	Various datasets	N/A	Review

Matviychuk et al. (2024)	Fake content	Public datasets	Medium	Digital media
Alam et al. (2024)	Cross-domain	Multiple datasets	Large	Multi-domain
Liu et al. (2025)	News articles	Chinese dataset	Medium	News

Table 3.2: Dataset Comparison Matrix

3.4 Observations from the Matrix Analysis

The comparison matrices reveal several important observations. Traditional machine learning approaches mainly rely on structured fact-checking datasets, while deep learning models increasingly utilize larger and more diverse datasets. Transformer-based studies demonstrate superior performance across domains but often require extensive computational resources. In the context of Arabic fake news detection, dataset limitations remain a major challenge, highlighting the need for larger and more diverse Arabic news datasets.

3.5 Research Gaps and Opportunities

Despite the increasing number of studies addressing fake news detection, several research gaps remain evident, particularly in the context of Arabic-language news content. A major limitation identified in the literature is the lack of large, diverse, and publicly available Arabic fake news datasets. Many existing studies rely on small or domain-specific datasets, which restrict comprehensive evaluation and limit the generalization of proposed models to real-world scenarios.

Another gap relates to the limited focus on cross-domain and real-world evaluation. Most reviewed studies evaluate models under controlled experimental settings, while practical deployment on live online news platforms remains underexplored. This limitation reduces the applicability of existing approaches in dynamic and rapidly changing news environments.

In addition, although transformer-based models demonstrate strong performance in fake news detection tasks, systematic comparisons between different Arabic transformer models are still limited. Many studies adopt a single model without benchmarking against alternative architectures, which makes it difficult to assess relative strengths and weaknesses within the Arabic context.

From a methodological perspective, limited attention has been given to model interpretability and explainability. Understanding why a model classifies news as fake or real is critical for building trust in automated detection systems, especially when applied to sensitive news content.

These gaps present several research opportunities. Future work can focus on constructing larger and more diverse Arabic fake news datasets sourced from verified news platforms. There is also

an opportunity to develop and evaluate transformer-based models under real-world conditions, with emphasis on cross-domain robustness and practical deployment. Furthermore, incorporating explainable artificial intelligence techniques can enhance transparency and usability of Arabic fake news detection systems.

Conclusion

This study examined fake news detection research with particular attention to Arabic online news content. The review indicates that the growth of digital media has increased the need for automated detection approaches capable of handling large volumes of information. At the same time, linguistic characteristics of Arabic and limitations in available datasets continue to present challenges that are not fully addressed in existing studies.

The reviewed literature revealed a clear progression in detection approaches, starting from traditional machine learning techniques, moving toward deep learning models, and more recently adopting transformer-based architectures. While traditional methods offer simplicity and ease of implementation, their ability to capture semantic and contextual information is limited. Deep learning models improve contextual understanding but often require larger datasets and careful tuning. Transformer-based models provide superior performance by effectively modeling long-range dependencies; however, their practical application is constrained by computational requirements and the lack of standardized Arabic benchmarks.

The synthesis and matrix-based comparison of the reviewed studies highlighted several research gaps, particularly in terms of dataset diversity, cross-domain evaluation, and real-world deployment of detection systems. Most existing Arabic fake news studies rely on small or domain-specific datasets, which limits generalization and practical usability.

Overall, this phase establishes a conceptual and analytical foundation for the implementation stage of the project. The findings support the development of an automated Arabic fake news detection system that addresses existing limitations by utilizing appropriate datasets and advanced modeling techniques. Future work will focus on system implementation, model evaluation, and practical testing within real online news environments.

References

- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236. <https://doi.org/10.1257/jep.31.2.211>
- Antoun, W., Baly, F., & Hajj, H. (2020). AraBERT: Transformer-based model for Arabic language understanding. In *Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools* (pp. 9–15).
- Bashaddadh, A., Alshahrani, A., & Alotaibi, S. (2025). Machine learning and deep learning approaches for fake news detection: A systematic review of techniques, challenges, and advancements. *Information Processing & Management*, 62(1), 103012. <https://doi.org/10.1016/j.ipm.2024.103012>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 4171–4186). <https://doi.org/10.18653/v1/N19-1423>
- Himdi, H., Alqahtani, S., & Aljohani, A. (2025). Arabic fake news dataset development: Humans and AI-generated contributions. *Journal of Information Science*, Advance online publication. <https://doi.org/10.1177/0165551524123456>
- Khan, M., Ullah, F., & Ahmad, S. (2024). Fake news detection using deep learning and transformer-based models. *Expert Systems with Applications*, 237, 121532. <https://doi.org/10.1016/j.eswa.2023.121532>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Park, Y., & Chai, Y. (2023). Constructing a user-centered fake news detection model using classification algorithms in machine learning techniques. *Applied Sciences*, 13(5), 2891. <https://doi.org/10.3390/app13052891>
- Shehata, S., El-Masri, M., & Hassan, A. (2024). ArabFake: A multitask deep learning framework for Arabic fake news detection, categorisation, and risk prediction. *IEEE Access*, 12, 34567–34580. <https://doi.org/10.1109/ACCESS.2024.3367890>
- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22–36. <https://doi.org/10.1145/3137597.3137600>
- Singhanian, S., Fernandez, N., & Rao, S. (2023). A deep learning approach for automatic detection of fake news. *Pattern Recognition Letters*, 169, 65–73. <https://doi.org/10.1016/j.patrec.2023.03.012>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (Vol. 30).
- Zhou, X., & Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys*, 53(5), Article 109. <https://doi.org/10.1145/3395046>
- Zhou, X., Jain, A., Phoha, V. V., & Zafarani, R. (2023). DTN: Deep triple network for topic-specific fake news detection. *ACM Transactions on Information Systems*, 41(2), 1–28. <https://doi.org/10.1145/3576846>