



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	5
Chapter 1: Introduction	6
1.1 Problem Definition.....	6
1.2 Research Objectives.....	6
1.3 Research Contribution	7
1.4 Thesis Organization.....	7
Figure 1.1 Structure of the thesis showing the main chapters and conclusion.....	7
Chapter 2: Overview.....	8
2.1 Fake News Concept.....	8
2.2 Natural Language Processing in Fake News Detection	8
2.3 Machine Learning Approaches	8
2.4 Deep Learning Approaches	8
Chapter 3: Related Work.....	10
3.1 Reviewed Studies	10
3.2 Synthesis Matrix of Reviewed Studies	11
Table 3.1: Comparison Matrix of Fake News Detection Studies.....	11
3.3 Dataset-Oriented Comparison Matrix	11
Table 3.2: Dataset Comparison Matrix	12
3.4 Observations from the Matrix Analysis	12
3.5 Research Gaps and Opportunities	12
Conclusion	14
References	15

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 recently evolved from the status of an important source of information to this of modern society. These platforms make news content available near real time and allow information to be shared continuously. But the very same environment has also led to the proliferation of misleading and made-up news. Such fake news continues to pose a formidable challenge by skewing facts, swaying public perception, and diminishing confidence in reliable media outlets. The problem of fake news has attracted increased attention with the rise of digital media and social media. Manual verification is not achievable at the same pace for the information which travels rapidly online as there are hundreds of news articles that are getting published every single day. Consequently, bad news or misleading content can spread fast through a large audience before detection or correction. This challenge could be more noticeable in the Arabic-language news. But as the Arabic language is rich in morphology with a wide variety of writing styles as well as formal versus informal usage. It's these features that make automated text analysis harder, and make existing fake news detection more difficult. Moreover, the supply of large, reliable, and diverse Arabic fake news datasets is relatively scarce.

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 on 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

The results of this study add to the academic literature in fake news detection research on Arabic-language material, and serve as a basis for subsequent research. By categorizing current research studies into methodological groups, it can provide a clear outline of the developments in detection techniques. The work also illustrates the strengths and weaknesses of currently available methods without drawing detailed experimental comparisons. This phase will contribute in identifying the research gaps at the level of Arabic news analysis such as data availability, linguistic complexity, and assessment. The results reinforce the design of an automated Arabic fake news detection mechanism which will be incorporated in the forthcoming project phase.

1.4 Thesis Organization

This report is divided into the following three chapters. The research problem, objectives, and contribution are explained in Chapter 1. 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.

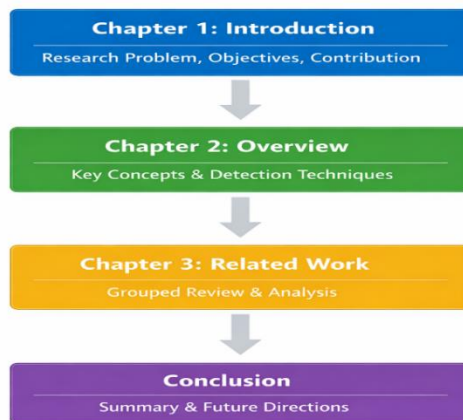


Figure 1.1 Structure of the thesis showing the main chapters and conclusion

Chapter 2: Overview

In this chapter, we review the main concepts and methods used for fake news detection. The emphasis is on the definition of central terms, major computational frameworks, and methods for Arabic language news analysis.

2.1 Fake News Concept

This is the idea that news content is deliberately untrue. Such information is typically written to appear credible journalism, and detection is difficult by nature. Previous studies suggest that fake news may not be the same as real news in tone, emotional impact, and story coherence, though these distinctions are generally not clearly reported. Multiple studies have stated that the difference between fake news and credible news lies in linguistic features, emotional tone and consistency of content (Vosoughi et al., 2018). These features have been commonly exploited as features of automated detection systems.

2.2 Natural Language Processing in Fake News Detection

Natural Language Processing is concerned with the ability of computers to interpret and understand human text. In the area of fake news detection, NLP techniques are primarily utilized to prepare text for automated analysis and extract patterns that aid in classification. Since raw news text is rarely in the form of raw data for direct processing, operations such as tokenization and normalization are often carried out prior to performing model training. Typical NLP processes involve text normalization, tokenization, and invalid content removal. These steps enable us to convert raw text to a structure relevant to machine learning or deep learning models (Aggarwal, 2018). NLP facilitates automated analysis of large-scale news data significantly.

2.3 Machine Learning Approaches

Early fake news detection systems utilized traditional machine learning techniques along with handcrafted textual features (Shu et al., 2019). These methodologies rely on classifiers like Support Vector Machines, Naïve Bayes and Random Forest to differentiate between fake or real news. Although the performance of the machine learning model shows fair performance in the controlled environment, the performance generally depends on the quality of feature engineering and dataset (Zhou & Zafarani, 2020). These limitations drove the utilization of more advanced learning techniques.

2.4 Deep Learning Approaches

Deep learning methods enhanced fake news detection by allowing automatic learning of textual representations directly from data (LeCun et al., 2015). Neural network-based architectures including Convolutional Neural Networks and Recurrent Neural Networks have been utilized extensively in order to represent semantic and sequence pattern in news content. Literature is clear that deep learning models surpass traditional machine learning methods in the performance of text

classification, including fake news detection tasks (Wang, 2017). Nevertheless, these models continue to suffer from problems that pertain to long text sequences and domain generalization.

2.5 Transformer-Based Models

Transformer-based models represent a huge step forward in the state of art in natural language processing. More effective than previous architectures, these are based on attention mechanisms that learn contextual linkages inside text (Vaswani et al., 2017). Transformer models like BERT and their variants have achieved high performance for fake news detection, because they model long-range dependencies with good accuracy (Devlin et al., 2019). Arabic transformer models obtain better results by using large-scale Arabic corpora, which can be used as news Arabic (Antoun et al., 2020).

2.6 Arabic News Datasets

The performance of fake news detection systems depends on the availability of reliable datasets. The dataset of fake news in Arabic is generally small and non-diverse as compared to all-languages datasets (Himdi et al., 2024). Previous Arabic datasets are limited by being focused on specific subjects or platforms, limiting overall evaluation and model generalizing (Shehata et al., 2024). This limitation serves as an avenue for future datasets construction and validation efforts

Chapter 3: Related Work

3.1 Reviewed Studies

In the method of detection employed, the reviewed studies were grouped into three methodological categories. This classification is an essential comparison of different methodological trends on the detection of fake news across research lines.

Predominantly traditional machine learning algorithms vs. pre-trained neural networks. This includes research efforts based on classical machine learning classifiers, or early-stage neural models with hand-engineered features. Each of them shows us acceptable performance with little contextualization. Park & Chai (2023) Singhania et al. (2023) and Zhou et al. (2023)

Classifying Fake News by Arabic: In Arabic You Can also Learn from Deep Learning Models. The group emphasizes deep learning methods for Arabic news using CNN, LSTM and GRU architectures. Though providing advancements to contextual modeling, the size of datasets and the domain-specific nature of these studies limit their application. Himdi et al. (2025) Shehata et al. (2024) and Khan et al. (2024)

Transformer-Based and Intelligent Architectures. This group comprises work that uses transformer-based models and advanced language architectures as ways to improve semantical interpretation and the accuracy of detection. Bashaddadh et al. (2025) Matviychuk et al. (2024) Alam et al. (2024) and 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 provide several interesting points. The classical methodologies for machine learning rely on standardized fact-checking datasets and the deep learning ones are based on more comprehensive and more diverse data sources. Transformer-based analysis performs well in multiple domains of research but is computationally expensive. The limitations of datasets still pose a significant challenge in the context of Arabic fake news detection, and thus there is a continued need for significantly larger and more diverse Arabic news datasets.

3.5 Research Gaps and Opportunities

Despite the growing number of papers addressing fake news detection, we see some gaps in the existing studies, especially on Arabic-language news content. The literature highlights a significant limitation in its use as the scarcity of large, heterogeneous datasets of Arabic fake news; data are primarily available. Due to most of the prior studies relying on small or domain-specific datasets which are not large enough, the methods are difficult to assess comprehensively and do not provide a suitable representation for applied models for real-world circumstances. There is also a gap because of the lack of emphasis given on cross-domain evaluation and real-world validation. The majority of the studies included in this review examine models in experimental conditions, and no real-world deployment is yet proposed for active online news sites. This restriction makes the existing methods less applicable to the new environments as fast growing and changing news. Furthermore, while transformer-based models do a very good job on fake news detection, they continue to fall short on systematic comparisons between Arabic transformer models. However, a number of studies choose a single model without reference to different architectures, so it may be challenging to investigate the respective strengths and weaknesses in the context of Arabic. Methodologically, there has been little attention focused on interpretability and explainability of models. Knowing what type of news content a model considers to be fake news is important for building trust in automated detection systems, especially when sensitive news content is involved. These research gaps give rise to a number of future avenues of

investigation. Moving on, more scalable and more varied Arabic fake news databases could be produced using verified news sites. Moreover there is an opportunity to develop and test transformer-based models under real-world conditions, with a focus on cross-domain robustness and practical application. Furthermore, it is hoped that if explainable AI techniques can be incorporated, the transparency and practicality of Arabic fake news detection systems may be increased.

Conclusion

This study focused on the fake news detection literature particularly on Arabic news information published online. The study found that the increase of digital media had created more demand for automated detection methods able to cope with large flow of information. Conversely, linguistic features of Arabic and the restricted datasets remain issues that were not fully explored in previous studies. The review of literature showed a progression in methods of detection ranging from classical ML methods, to deep learning models and recently adopting transformer-based architectures. Conventional approaches are simple and easy to implement, but they have less capacity to represent semantic and contextual information. Deep learning frameworks enhance the context-specific context but typically require larger data and fine-tuning. Although transformer-based models have higher performance by representing long-range dependencies, they are limited in operation due to the computational resource, coupled with the absence of common Arabic benchmarks. While some of the literature synthesis and matrix comparisons of the reviewed texts highlighted notable research gaps related to dataset diversity, cross-domain evaluation, and detection system deployment in real-world applications, the majority of existing Arabic fake news studies were small or domain-specific, limiting both generalizability and practical usability. In summary, this phase set conceptual and analytical base for the project's implementation phase. Such study is important and the results indicate the requirement for an automated Arabic fake news detection system to eliminate the existing limitations to the existing works on the detection of fictitious fake news by using suitable datasets and sophisticated methods of modeling. We will continue working toward the implementation of our system, model evaluation and experimental testing in real online news-based applications.

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>