

Course Title: Research Methodology in Computer Science

Course Code: AI800

Weighting: 20%

Lecturer: Dr. Mohammed Anbar

Hand-in Date: 14 Mar 2026

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

Assignment Report: Detecting Fake News Related to the Syrian Conflict on Social Media Using Machine Learning and Deep Learning Techniques

Student: Eman Alhabashi

1 Introduction

With the rapid use of social media platforms in recent years, the way information is created and shared has changed a lot compared with the past. Millions of people now generate and share news every day using platforms such as X (formerly Twitter), Facebook, and other online applications. These platforms make it easier for people to access and exchange information. However, they also allow misleading and false information to spread very quickly among users.

Fake news has become a serious issue in the digital world. False information can influence public opinion and create confusion, especially during important events such as elections or health crises. When misleading information spreads quickly, it becomes difficult for users to determine which sources are trustworthy and which are not.

A clear example of this problem appeared during the COVID-19 pandemic. During that time, social media platforms were filled with posts related to the virus, treatments, and vaccines. Some of this information was correct, but a large portion of it was inaccurate or misleading. This created confusion among users and sometimes affected public health decisions.

Because of these challenges, researchers have started exploring automatic ways to detect fake news using artificial intelligence. Machine learning and natural language processing techniques can analyze large amounts of online text and help identify patterns that may indicate misleading information.

Several studies have proposed different methods to detect fake news. Early research mostly relied on traditional machine learning algorithms that use text features extracted from the content. Later studies began using deep learning models that can automatically learn patterns from large datasets. More recently, transformer-based models such as BERT and AraBERT have shown strong performance in many language processing tasks.

However, detecting fake news in Arabic presents additional challenges. The Arabic language has complex grammar and many dialects that are commonly used in social media communication. Another challenge is the limited availability of large Arabic datasets for misinformation detection. These factors make it more difficult to train accurate models for detecting fake news in Arabic text.

Because of this, researchers are increasingly focusing on developing better datasets and more effective models for Arabic misinformation detection. Reviewing previous research is an important step for understanding how the problem has been studied and what methods have been used.

This report reviews six research papers related to fake news detection using machine learning and deep learning techniques. The aim of this review is to understand the different approaches used in these studies and identify their main findings and limitations.

2 Literature Review

Paper 1

Title: ArCOV19-Rumors: Arabic COVID-19 Twitter Dataset for Misinformation Detection

This study introduced a dataset designed to help researchers detect misinformation in Arabic tweets related to COVID-19. The authors collected verified claims from fact-checking websites and then searched for tweets related to those claims. Each tweet was manually labeled depending on whether it supported the claim, denied it, or was unrelated.

The dataset contains 138 claims and more than 9,000 tweets. The researchers tested several machine learning and transformer-based models to evaluate the dataset. Models such as AraBERT and MARBERT were used for classification tasks. The results showed that the MARBERT model performed better than the other models.

One of the strengths of this study is the creation of a useful dataset for Arabic misinformation research. However, the dataset focuses on COVID-19 misinformation. Despite this limitation, the methods used in this study can also be applied to detect misinformation related to political or humanitarian situations such as the Syrian conflict.

Paper 2

Title: Arabic Fake News Detection Using Deep Learning

This study focused on using deep learning models to detect fake news written in Arabic. The researchers collected a dataset from online news websites and social media platforms. The dataset was labeled into two categories: real news and fake news.

Different deep learning models were applied in this study, including Convolutional Neural Networks (CNN) and Long Short-Term Memory networks (LSTM). The results showed that combining CNN and LSTM models improved the classification accuracy compared with traditional machine learning methods.

Although the study achieved good results, the dataset used was relatively small. This may limit the ability of the model to perform well on larger and more diverse data.

Paper 3

Title: Detecting Arabic Fake News Using Machine Learning

This research examined the use of traditional machine learning algorithms for detecting fake news in Arabic text. The dataset used in the study consisted of Arabic news headlines and posts collected from different online sources.

Several machine learning algorithms were tested, including Support Vector Machine (SVM), Naïve Bayes, and Random Forest. The experimental results showed that the SVM model achieved the highest accuracy among the tested models.

The study demonstrated that traditional machine learning methods can still provide useful results. However, these models depend heavily on manually extracted features and may not capture deeper meanings in the text.

Paper 4

Title: AraBERT: Transformer Model for Arabic NLP

This study presented AraBERT, a transformer-based language model designed specifically for Arabic language processing tasks. The model was trained on a large Arabic text corpus and tested on several natural language processing tasks.

The results showed that AraBERT achieved better performance compared with many traditional machine learning models and earlier Arabic language models. Because of this, AraBERT is now widely used in tasks such as sentiment analysis and fake news detection.

Despite its strong performance, AraBERT requires significant computing resources and large datasets for training and fine-tuning.

Paper 5

Title: Detecting Misinformation in Social Media Using Deep Neural Networks

This study explored the use of deep neural networks to detect misinformation on social media platforms. The researchers focused on analyzing textual content and identifying patterns that may indicate misleading information.

The results showed that deep learning models can capture complex relationships in text data and improve the detection accuracy compared with traditional machine learning models.

However, these models require large training datasets and may become less accurate if misinformation patterns change over time.

Paper 6

Title: Fake News Detection Using Natural Language Processing

This study investigated the use of natural language processing techniques for detecting fake news on social media platforms. Several text preprocessing methods were used, including tokenization and TF-IDF feature extraction.

After preprocessing the data, machine learning algorithms such as Logistic Regression and Random Forest were applied for classification. The results showed that combining linguistic features with machine learning algorithms can help detect fake news effectively.

One of the challenges mentioned in the study is the difficulty of dealing with multilingual data and the constantly evolving strategies used to spread misinformation.

3 Comparative Analysis

From reviewing the previous studies, it can be seen that different methods can be used to detect fake news. Some studies rely on traditional machine learning algorithms such as SVM and Naïve Bayes, while other studies use deep learning models. These methods can be applied to different types of misinformation, including political and humanitarian misinformation related to the Syrian situation.

Deep learning models such as CNN and LSTM can learn patterns directly from large datasets and may perform better than traditional models in many text classification tasks. More recent research also uses transformer-based models such as AraBERT, which can better understand the context of Arabic text.

For research related to the Syrian situation, these models could be used to analyze social media posts and detect misleading information about political events, humanitarian issues, or military developments. However, one of the main challenges is the lack of publicly available datasets that focus specifically on misinformation related to the Syrian conflict.

4 Literature Review Summary Table

Study	Authors	Year	Research Design	Methodology	Key Findings	Limitations
ArCOV19-Rumors Dataset	Haouari et al.	2021	Experimental	Dataset creation and transformer models	MARBERT achieved the best performance	Focuses mainly on COVID-19 rumors
Arabic Fake News Detection Using Deep Learning	Al-Yahya et al.	2020	Experimental	CNN and LSTM models	Hybrid models improved accuracy	Small dataset
Detecting Arabic Fake News Using ML	Alkhair et al.	2019	Experimental	SVM, Naïve Bayes, Random Forest	SVM achieved the best accuracy	Limited semantic understanding
AraBERT Model	Antoun et al.	2020	Experimental	Transformer language model	Significant improvement in NLP tasks	High computational requirements
Detecting Misinformation Using Deep Neural Networks	Shu et al.	2017	Experimental	Deep neural networks	Better semantic representation	Requires large datasets
Fake News Detection Using NLP	Ahmed et al.	2018	Experimental	NLP preprocessing and ML classifiers	Linguistic features improve detection	Difficulty handling multilingual data

5 Conclusion

Fake news detection has become an important research area because misinformation spreads quickly on social media platforms. This problem becomes even more serious during political conflicts and humanitarian crises. In the case of the Syrian situation, large amounts of information are shared daily online, and some of this information may be misleading or inaccurate.

Different artificial intelligence techniques such as machine learning, deep learning, and transformer-based models have been proposed to address this problem. These methods can help analyze Arabic social media posts and identify misleading information.

However, detecting misinformation related to the Syrian situation remains challenging because there are limited datasets that focus specifically on this topic. Future research should focus on collecting datasets related to the Syrian conflict and improving models that can better analyze Arabic social media discussions about political and humanitarian events.

References

- Ahmed, H., Traore, I., & Saad, S. (2018). Detecting opinion spams and fake news using text classification. *Security and Privacy Journal*.
- Alkhair, M., Meftouh, K., Smaïli, K., & Othman, N. (2019). An Arabic corpus for fake news detection. *Proceedings of the International Conference on Arabic Language Processing*.
- Al-Yahya, M., Al-Khalifa, H., & Al-Bukhitan, S. (2020). Detecting fake news in Arabic using deep learning models. *International Journal of Advanced Computer Science*.
- Antoun, W., Baly, F., & Hajj, H. (2020). AraBERT: Transformer-based model for Arabic language understanding. *Proceedings of the Workshop on Open-Source Arabic Corpora*.
- Haouari, F., Hasanain, M., Suwaileh, R., & Elsayed, T. (2021). ArCOV19-Rumors: Arabic COVID-19 Twitter dataset for misinformation detection. *ACL Workshop on NLP for COVID-19*.
- 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*.