

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

**Student:** Eman Alhabashi

## Table of Contents

Detecting Fake News Related to the Syrian Conflict on Social Media Using Machine

|                                              |           |
|----------------------------------------------|-----------|
| Learning and Deep Learning Techniques .....  | 3         |
| <b>Chapter 1: Introduction .....</b>         | <b>4</b>  |
| <b>1. Background .....</b>                   | <b>4</b>  |
| <b>2. Problem Statement.....</b>             | <b>5</b>  |
| <b>3. Research Objectives .....</b>          | <b>6</b>  |
| <b>4. Research Scope.....</b>                | <b>8</b>  |
| <b>Chapter 2: Literature Review .....</b>    | <b>9</b>  |
| <b>New Advancements (2022-2025): .....</b>   | <b>9</b>  |
| <b>Chapter 3: Research Methodology .....</b> | <b>11</b> |
| <b>1. Research Phases .....</b>              | <b>11</b> |
| <b>2. Expected Outcomes .....</b>            | <b>11</b> |
| <b>3. Gantt Chart.....</b>                   | <b>13</b> |
| <b>References.....</b>                       | <b>14</b> |

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

## **Abstract**

The rapid spread of misinformation regarding the Syrian conflict on social media platforms has significant societal and humanitarian consequences. While general Arabic fake news detection has achieved high accuracy (up to 99%), models specifically tailored to the Syrian context often struggle with a performance gap due to dialectal nuances and limited datasets. This research proposes a hybrid approach combining Transformer-based models (AraBERT) with Deep Learning architectures (Bi-LSTM/Bi-GRU) to enhance detection accuracy. The study aims to collect a specialized dataset and develop a robust classification framework. The expected results will provide a more reliable tool for identifying misinformation in conflict zones, thereby mitigating the negative impact of digital deception.

## Chapter 1: Introduction

### 1. Background

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.

## **2. Problem Statement**

With viral fake news on social media, it is framing the Syrian conflict. False information lies and manufactured plots do not only affect what people believe, they can also disrupt societies, shift how countries respond and intensify existing tensions.

Old-school machine learning methods such as SVM, RF, and NB notice some of these, but miss the nuances of Arabic dialects and

the semantic context of posts on the conflict. More recently, deep learning models such as LSTM, BiLSTM and transformer-based models such as AraBERT have outperformed. But they are limited by small datasets and the ever-changing social media environment (Hussein et al., 2021).

Disinformation influences public opinion, humanitarian response, and policy.

Its objectives are to survey existing methods, assess their strengths and weaknesses, and design an effective fake news detection system for the Syrian conflict on social media using state-of-the-art machine learning and deep learning models.

### **3. Research Objectives**

- **To comprehensively analyze existing methods for Arabic fake news detection**
  - To identify datasets, preprocessing techniques, and linguistic challenges specific to Arabic content.
- **To investigate the failure of current ML and deep learning models to detect conflict misinformation.**
  - To evaluate machine learning and deep learning models on Syrian conflict data

- To assess traditional ML methods (SVM, Random Forest, Naïve Bayes) for accuracy, precision, recall, and F1-score.
- To capture Arabic dialects, semantics, and time in Deep learning models (LSTM, BiLSTM, CNN-LSTM, transformers)
- **To identify out the critical hurdles and blind spots in Syrian conflict-specific fake news detection**
  - To highlight issues such as data scarcity, imbalance, dialectal diversity, temporal dynamics, and ethical considerations. These problems affect both schema model robustness and real-world utility.
- **To provide a more holistic model for fake news detection regarding the Syrian war**
  - To integrate insights from previous studies and best-performing models to develop a hybrid detection approach. including preprocessing, feature extraction, and model selection for the language and context of Syrian social media.
  - To guide tips for data set generation, model evaluation, and ethical deployment in political contexts.

#### **4. Research Scope**

The study focuses on Arabic-language content (MSA and Syrian dialect) on X (Twitter). The temporal scope covers posts related to the Syrian conflict, focusing on political events and humanitarian reports.

## Chapter 2: Literature Review

Recent studies have explored various techniques for Arabic fake news detection. Early works by **Ahmed et al. (2018)** and **Alkhair et al. (2019)** focused on text classification and the creation of initial Arabic corpora. However, these models relied on hand-crafted features which often missed semantic depth.

### New Advancements (2022-2025):

1. **Ensemble Deep Learning: Alrahmawy et al. (2024)** proposed ensemble-based deep learning models that achieved high performance, suggesting that combining multiple architectures can stabilize detection across different topics.
2. **Hybrid Contextual Embeddings: Asiri and Shoaib (2025)** demonstrated that combining Bi-LSTM with BERT embeddings significantly improves accuracy on Twitter data by capturing both context and sequence information.
3. **Multitask Frameworks: Shehata et al. (2024)** introduced "ArabFake," a framework using MARBERTv2 for multi-dialectal news. Their work highlights the importance of using models pre-trained on social media dialects rather than just standard Arabic.
4. **Contextual Deep Learning: Nassif et al. (2022)** analyzed various versions of BERT for Arabic and concluded that deep

contextualized embeddings are superior for identifying politically motivated misinformation in the Middle East.

**5. Traditional vs. Deep Learning: Alazab et al. (2022)** compared machine learning algorithms, finding that while traditional methods are faster, they lack the robustness required for the evolving nature of social media fake news.

These studies indicate a clear trend toward hybridizing Transformers with Deep Learning to handle the nuances of dialectal Arabic, which this research aims to apply to the Syrian context.

## Chapter 3: Research Methodology

### 1. Research Phases

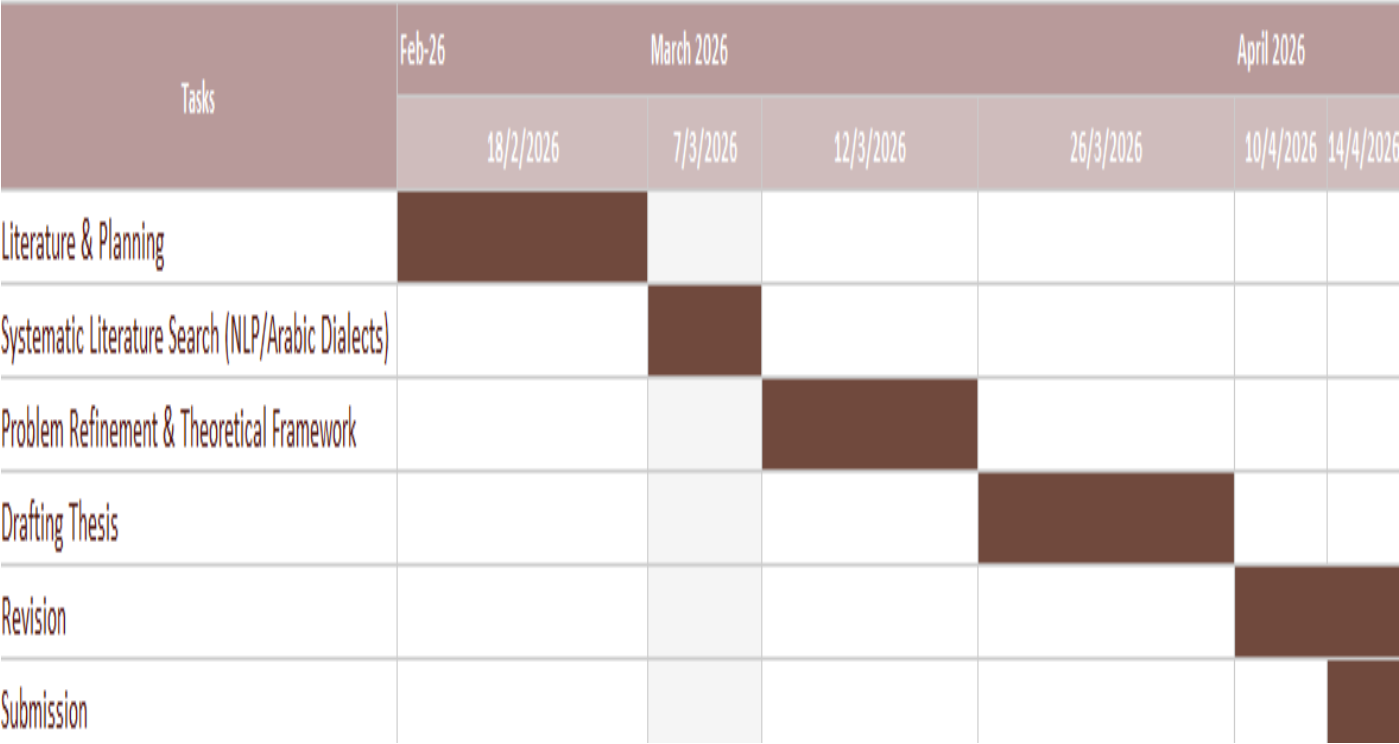
- **Phase 1: Data Acquisition:** Utilizing Twitter API to scrape posts using specific hashtags and keywords related to Syrian conflict (e.g., Arabic tokens for Syria, cities, groups, major events).
- **Phase 2: Preprocessing:** Applying Arabic-specific normalization, removing noise (remove URLs, handle emojis and hashtags, and apply light stemming/lemmatization only when appropriate), and performing dialectal morphological analysis.
- **Phase 3: Model Architecture:** Implementing a Hybrid Model where **AraBERT** extracts contextual features, which are then fed into a **Bi-LSTM** layer to learn sequential patterns.
- **Phase 4: Training and Testing:** Using an 80/20 split for training and validation, employing F1-score and Accuracy as primary metrics.

### 2. Expected Outcomes

- A specialized, high-quality dataset of Syrian-related fake news.

- A hybrid model that outperforms current Syrian-specific benchmarks.
- Insights into the common linguistic patterns used in Syrian conflict misinformation.

3. Gantt Chart



|                 | Submission | Revision  | Drafting Thesis | Problem Refinement & Theoretical Framework | Systematic Literature Search (NLP/Arabic Dialects) | Literature & Planning |
|-----------------|------------|-----------|-----------------|--------------------------------------------|----------------------------------------------------|-----------------------|
| Start Date      | 14/4/2026  | 10/4/2026 | 26/3/2026       | 12/3/2026                                  | 7/3/2026                                           | 18/2/2026             |
| Duration (Days) | 1          | 4         | 15              | 14                                         | 5                                                  | 17                    |

## References

1. Ahmed, H., Traore, I., & Saad, S. (2018). Detecting opinion spams and fake news using text classification. *Security and Privacy Journal*.
2. 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*.
3. 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*.
4. Antoun, W., Baly, F., & Hajj, H. (2020). AraBERT: Transformer-based model for Arabic language understanding. *Proceedings of the Workshop on Open-Source Arabic Corpora*.
5. 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*.
6. 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*.