

Using Deep Learning to Detect Arabic Fake News on Online Websites

Areej Mosaed Almalki
Master's Student of [Artificial Intelligence]
Midocean University

1. Project Overview

This project focuses on detecting fake news written in Arabic on online news websites using deep learning techniques. Most existing studies have focused on English or Chinese fake news detection, while Arabic remains underrepresented in this area. Misinformation in Arabic digital media can mislead readers, affect public opinion, and undermine the credibility of online journalism. This project aims to build a deep learning model capable of identifying fabricated Arabic news by analyzing linguistic and semantic patterns commonly found in deceptive content [1], [2]. Through this study, an Arabic-specific framework will be developed to enhance accuracy and reliability in fake news detection within the Arabic language domain.

2. Project Goal

The primary goal of this project is to develop a deep learning-based model to accurately detect fake Arabic news articles from online websites. The project also aims to improve the reliability of Arabic fake news detection systems by fine-tuning existing Arabic transformer models, such as MARBERTv2 and AraBERT [3]. Additionally, this project seeks to contribute to the development of automated systems that strengthen the trustworthiness of Arabic digital media.

3. Approach

The proposed approach involves four main components: data collection, data preprocessing, model selection, and evaluation. Arabic news data will be collected from both verified and unreliable online sources. The text will then be cleaned and normalized to handle language variations and remove noise. Transformer-based models, including MARBERTv2 and AraBERT, will be implemented to detect linguistic cues and contextual differences between real and fake news [1]. Model performance will be evaluated using metrics such as accuracy, precision, recall, and F1-score to ensure reliability [2].

4. Project Steps and Timeline

Time	Step	Description
Week 1–2	Collect Arabic news data	From both trusted and unreliable websites
Week 3	Clean and prepare text	Remove noise and process Arabic text
Week 4–5	Build and train model	Fine-tune MARBERTv2 and compare with AraBERT
Week 6	Evaluate results	Test model accuracy and analyze errors
Week 7	Write final report	Summarize and present project

5. Expected Results

It is expected that the developed model will achieve high accuracy in detecting fake Arabic news. The comparison between MARBERTv2 and AraBERT will provide insights into which transformer model performs better in Arabic NLP tasks [3]. Additionally, the project aims to contribute to the creation of a standardized dataset for Arabic fake news detection and offer a foundation for future Arabic-language research in misinformation analysis.

6. Resources Needed

- Python, Jupyter Notebook, and Hugging Face for model training
- Arabic text datasets collected from real websites
- A computer with good processing power (GPU if possible)
- Guidance from my supervisor or AI experts when needed

7. References

- [1] A. Bashaddadh, et al., “Fake News Detection using Deep Learning Models,” Journal of AI Applications, 2025.
- [2] K. Park and J. Chai, “User Behavior Patterns in Fake News Detection,” Computational Intelligence Journal, 2023.
- [3] A. Himdi, et al., “MARBERTv2: Transformer Models for Arabic Language Understanding,” Proceedings of ACL, 2025.