# Arabic Sentiment Analysis Using AraBERT Transformer-Based Model

Ammar Yasser Mohamed, Ali Mohamed Ali Elrawy, Hassan Mohamed Hassan, Ahmed Mohamed Abdelmaniem, Ahmed Mostafa Hassan, Omar Mohamed Abdelhakam, Abdelrahman Ibrahim Abdallah, Ziad Ahmed Ezzat, Mostafa Mohamed Abdellatef
Faculty of Engineering – New Ismailia National University, Ismailia, Egypt

*Abstract—*

*Arabic sentiment analysis is a challenging Natural Language Processing (NLP) task due to the linguistic richness, dialectal diversity, and morphological complexity of the Arabic language. This paper presents a Transformer-based sentiment analysis system using AraBERT for binary classification of Arabic text into positive and negative sentiments. A labeled Arabic dataset was collected and preprocessed through extensive normalization, noise removal, and text cleaning procedures. The AraBERT v2 model was fine-tuned using a supervised learning approach. Experimental results achieved an accuracy of 93.19%, precision of 93.64%, recall of 93.00%, and F1-score of 93.23%, demonstrating the effectiveness of transformer-based models for Arabic sentiment analysis.*

*Keywords—Arabic Sentiment Analysis, AraBERT, Transformers, NLP*

## I. Introduction

Sentiment analysis aims to identify and classify opinions expressed in textual data. Arabic sentiment analysis is challenging due to morphology and dialectal variation. Transformer-based models such as AraBERT address these challenges effectively.

## II. Related Work

Previous studies relied on classical machine learning and deep learning models. Recent transformer-based approaches such as AraBERT achieved state-of-the-art performance.

## III. Dataset Description

The dataset consists of Arabic text labeled as positive or negative. Data preprocessing included normalization, noise removal, and tokenization. An 80/20 train-test split was used.

## IV. Methodology

AraBERT v2 was fine-tuned for binary sentiment classification. Tokenization was performed using AraBERT tokenizer with a maximum length of 128 tokens.

## V. Experiments and Evaluation

The model was evaluated using accuracy, precision, recall, and F1-score metrics.

## VI. Results and Discussion

The model achieved strong performance with an F1-score exceeding 93%, outperforming traditional approaches.

## VII. Conclusion and Future Work

The proposed approach demonstrates the effectiveness of AraBERT for Arabic sentiment analysis. Future work includes dataset expansion and multi-dialect support.

## References

[1] W. Antoun et al., "AraBERT: Transformer-based Model for Arabic Language Understanding," 2020.

[2] J. Devlin et al., "BERT: Pre-training of Deep Bidirectional Transformers," NAACL, 2019.