

A Transformer-based Hybrid Model for Implicit Emotion Recognition in Arabic Text

Hanane Boutouta

Department of Computer Science, Faculty of Sciences, Ferhat Abbas University, Setif, Algeria | Artificial Intelligence Laboratory (AI-lab), Faculty of Sciences, Ferhat Abbas University, Setif, Algeria
hanane.boutouta@univ-setif.dz (corresponding author)

Abdelaziz Lakhfif

Department of Computer Science, Faculty of Sciences, Ferhat Abbas University, Setif, Algeria | Artificial Intelligence Laboratory (AI-lab), Faculty of Sciences, Ferhat Abbas University, Setif, Algeria
abdelaziz.lakhfif@univ-setif.dz

Ferial Senator

Department of Computer Science, Faculty of Sciences, Ferhat Abbas University, Setif, Algeria | Artificial Intelligence Laboratory (AI-lab), Faculty of Sciences, Ferhat Abbas University, Setif, Algeria
ferial.senator@univ-setif.dz

Chahrazed Mediani

Department of Computer Science, Faculty of Sciences, Ferhat Abbas University, Setif, Algeria | Artificial Intelligence Laboratory (AI-lab), Faculty of Sciences, Ferhat Abbas University, Setif, Algeria
chahrazed.mediani@univ-setif.dz

Received: 17 January 2025 | Revised: 14 April 2025 and 19 April 2025 | Accepted: 22 April 2025

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: <https://doi.org/10.48084/etasr.10261>

ABSTRACT

Implicit emotion recognition has emerged as an active area of research in modern Natural Language Processing (NLP). Unlike explicit emotions, which are directly expressed through emotional words, implicit emotions are inferred from the surrounding context, making their detection more challenging. While most research in Arabic NLP has focused on recognizing explicit emotions, the study of implicit emotions remains largely unexplored, primarily due to its unique linguistic and morphological characteristics. The current study addresses this gap by compiling an Arabic dataset for the implicit emotion recognition task, named Arabic Implicit Emotion Dataset (AIEmoD), which is curated from existing publicly available explicit emotion datasets. Furthermore, it proposes a novel hybrid deep learning model that integrates the Arabic transformer-based AraBERT model with a Bidirectional Gated Recurrent Units (BiGRU) network to recognize and classify implicit emotions in Arabic text. The proposed AraBERT-BiGRU model was evaluated on two widely used Arabic emotion datasets, AETD and SemEval-2018, in addition to the newly compiled AIEmoD dataset. The results show that the model achieved F1-scores of 79.87% on AETD and 70.67% on AIEmoD, significantly outperforming deep learning baseline methods. Moreover, the proposed model surpassed current state-of-the-art approaches for explicit emotion recognition, even when applied to the more challenging task of implicit emotion detection. These findings highlight the effectiveness and robustness of the proposed AraBERT-BiGRU model in recognizing implicit emotions in Arabic text.

Keywords-natural language processing; implicit emotion recognition; Arabic transformer; AraBERT; BiGRU

I. INTRODUCTION

Emotion recognition in textual data has become a significant research area within NLP, aiming to automatically identify emotions expressed in text. Typically, emotions are conveyed through explicit emotional words, such as happy or sad, but they can also be expressed implicitly, without using

direct emotional terms. Implicit emotion recognition involves inferring emotions indirectly from contextual cues rather than relying on explicit emotional expressions [1].

While most existing work in emotion recognition focuses on detecting explicit emotions, implicit emotion recognition has attracted increasing attention in recent years, especially due

to its real-world relevance in sentiment analysis and affective computing [2]. However, the task remains inherently more challenging and is especially underexplored in Arabic.

Arabic is the fifth most widely spoken language globally, with over 400 million speakers, and is the official language of approximately 22 countries [3]. It is a morphologically rich language characterized by high ambiguity, complex syntactic structures, numerous irregular forms, and a wide variety of dialectal variants. These dialects are often used alongside Modern Standard Arabic (MSA), which lacks standardized writing conventions for colloquial forms. Such linguistic complexity poses significant challenges in accurately identifying emotions, particularly those expressed implicitly [4]. Furthermore, compared to English, Arabic is considered a low-resource language, with limited emotional lexicons and datasets.

Traditional Arabic emotion recognition systems have primarily targeted explicit emotion detection, relying on rule-based methods or machine learning approaches that employ engineered features, such as Part-of-speech (POS) tagging, Term Frequency-inverse Document Frequency (TF-IDF), or Bag-of-words (BoW) representations, combined with classifiers, like Support Vector Machines (SVM) or Naïve Bayes [5–9]. However, these methods are generally insufficient in detecting implicit emotions, struggling to capture subtle and context-dependent cues without explicit emotional keywords. To address these limitations, researchers have turned to deep learning models, particularly Convolutional Neural Networks (CNNs) [10] and Recurrent Neural Networks (RNNs) [11], such as Long Short-Term Memory (LSTM) [12] and GRU [13]. These models, along with their bidirectional variants (BiLSTM, BiGRU), have demonstrated strong capabilities in modeling long-range dependencies and capturing contextual information from both forward and backward directions [14–20].

In parallel, pre-trained contextualized language models, particularly Bidirectional Encoder Representations from Transformers (BERT) [20], showed significant promise in various NLP tasks. Unlike previous approaches, transformer-based models utilize the attention mechanism to assess the importance of each word in a sentence, facilitating a deeper understanding of context and significantly improving performance across many tasks. AraBERT [21] was the first transformer-based pretrained language model in Arabic based on the BERT model. It was pre-trained on more than 70 million Arabic sentences from diverse sources, with preprocessing steps tailored for Arabic morphology and syntax. AraBERT has demonstrated strong performance in a range of NLP tasks, including sentiment analysis and emotion detection.

Nevertheless, implicit emotion recognition in Arabic has not achieved comparable progress as in English [22–27] and in Chinese [28, 29]. To the best of the authors' knowledge, the presented work is among the first attempts to address this gap by leveraging the power of pre-trained language models combined with sequence modeling networks.

A novel hybrid deep learning model is proposed that integrates the contextual word embeddings of AraBERT with

the sequential learning capabilities of a BiGRU network. This combination aims to better capture the nuanced semantics and hidden emotional cues embedded in Arabic text. The model is evaluated across several datasets, including AETD [7], SemEval-2018 [30], and a newly compiled, task-specific dataset, AIEmoD, which focuses specifically on implicit emotion recognition.

II. ARABERT-BIGRU MODEL

The proposed AraBERT-BiGRU uses the AraBERT pre-trained language model to produce the embedding representation for each text input. The word embeddings are then passed to the BiGRU model to capture the long-range dependencies present within them. The outputs of the BiGRU layer are then fed into a dense layer to further learn and refine the relationships between the extracted features and the target emotion classes. Finally, a classification layer with a softmax activation function estimates the probability distribution over the emotion classes. The overall architecture of the hybrid AraBERT-BiGRU model is illustrated in Figure 1.

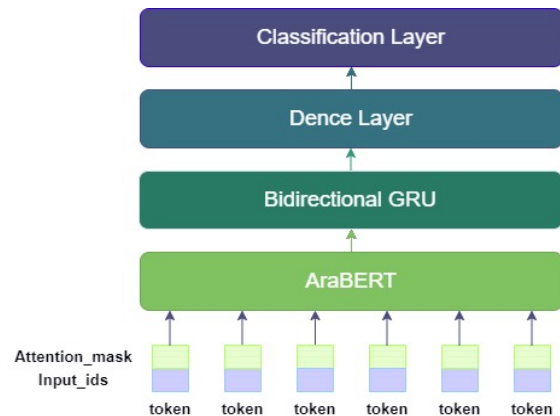


Fig. 1. The architecture of the proposed AraBERT-BiGRU hybrid model.

A. AraBERT Pre-trained Language Model

AraBERT is a BERT-based language model specifically designed for Arabic and has demonstrated high effectiveness in generating both word- and sentence-level embeddings [21]. Unlike traditional word embedding methods, such as Word2Vec and GloVe, which produce static representations regardless of context, AraBERT generates contextualized embeddings that dynamically reflect a word's meaning based on its surrounding context.

In addition to word embeddings, AraBERT is also well-suited for generating sentence embeddings. Each input sentence is first tokenized into N tokens, and a special classification token [CLS] is added at the beginning of the sequence. This [CLS] token serves as a representation of the entire sentence. Specifically, the hidden state corresponding to the [CLS] token in the penultimate layer is typically utilized as the sentence embedding, effectively capturing the overall semantic meaning of the sentence based on its full context.

The key strength of AraBERT lies in its ability to produce dynamic, bidirectional representations that are sensitive to

context, making it particularly effective for complex natural language understanding tasks, such as implicit emotion recognition. Furthermore, its pre-training on large-scale Arabic corpora enables AraBERT to generate high-quality embeddings that are well-suited for capturing the nuanced characteristics of Arabic text required in this task.

B. Bidirectional Gated Recurrent Units

The embedding representation produced by the AraBERT pre-trained language model is used as input for the BiGRU model. A GRU is an advanced variant of the LSTM model, employing gating mechanisms to regulate information flow and mitigate the vanishing gradient problem inherent in traditional RNNs. Unlike LSTMs, which utilize three gates (input, output, and forget gates), a GRU consists of only two gates (a reset gate and an update gate), resulting in fewer trainable parameters. Despite this simplified architecture, GRUs often achieve performance comparable to that of LSTMs while requiring less training time.

The BiGRU architecture effectively captures bidirectional contextual information by employing two independent GRUs: a forward GRU, which processes the sequence in its original order, and a backward GRU, which processes the sequence in reverse. Both GRUs generate forward and backward hidden states from the input word vectors. These states are then combined to form a comprehensive representation of each word within its full context, enabling the model to effectively capture long-range dependencies. Since emotional features are extracted through both forward and backward GRU networks, the BiGRU can better understand the global context of the text, which is essential for recognizing implicit emotions.

C. Classification Layer

In the proposed model, two dense layers are employed. The first layer is integrated to capture the relationships between the BiGRU outputs and the class labels, and then a dropout technique is applied (at a rate of 0.25) to mitigate overfitting. The second dense layer is designed to produce the final classification output. This layer utilizes the softmax activation function to convert the output values into a probability distribution across the classes, with the class having the highest probability being selected as the predicted class for the input sequence.

III. EXPERIMENTS

In this section, the details of the experiments performed in this study and the results obtained are discussed.

A. Dataset

The proposed AraBERT-BiGRU model was evaluated using two well-known publicly available Arabic emotion datasets: the SemEval 2018, Task 1 dataset [30] and the AETD dataset [7]:

- SemEval-2018 Task 1 dataset: The Ar-Ec subtask involves classifying a tweet as either "no emotion" or into one or more of eleven distinct emotions. For this subtask, only tweets annotated with a single emotion were considered. The EI-oc subtask requires classifying a tweet into one of four ordinal intensity levels. The present work focused

solely on the presence of emotion, excluding intensity levels and removing tweets annotated as containing no emotion. The SemEval-2018 Task 1 dataset is publicly available on the official SemEval-2018 Affect in Tweets (AIT) task webpage [31].

- AETD dataset: The dataset comprises 10,000 tweets, which were manually annotated into eight emotional categories: sympathy, joy, love, anger, sadness, fear, surprise, and no emotion (none). This dataset is publicly available on Github [32].

B. Preprocessing and Filtering

Tweets from both datasets were preprocessed using the most common preprocessing techniques, which included removing non-Arabic letters, links, mentions, retweets, numbers, repeated characters, punctuation marks, stop words, and Arabic diacritics. The present work also replaced emojis and emoticons with emotions and applied text normalization as part of the preprocessing steps.

To ensure that the models developed for this study would focus on recognizing emotions from the context rather than relying on emotional words, all tweets that contained explicit emotional words, which were identified from the National Research Council of Canada (NRC) Emotion Lexicon, were filtered out.

This study considered the four emotion categories, anger, fear, joy, and sadness, which are shared between these two datasets. The filtered tweets were divided into three sets: 70% for training the proposed model, 15% for validation, and the remaining 15% for testing the efficiency of the model.

C. Arabic Implicit Emotion Dataset

By combining and carefully filtering the aforementioned datasets, AIEmoD, a curated dataset of 7,684 Arabic tweets was developed. The dataset was specifically curated to mitigate the scarcity of annotated resources in the field of implicit emotion recognition in Arabic text. Further details regarding all the implicit emotion datasets used in this study are presented in Table I. The distribution of emotion categories within the AIEmoD dataset is depicted in Figure 2.

TABLE I. DESCRIPTION OF THE SEMEVAL-2018, AETD, AND AIEMOD IMPLICIT DATASETS

Emotion	SemEval-2018			AETD			AIEmoD		
	Train	Valid	Test	Train	Valid	Test	Train	Valid	Test
Anger	638	153	138	946	216	221	1584	369	359
Fear	407	83	90	641	135	147	1048	218	237
Sadness	726	122	154	778	171	155	1504	293	309
Joy	407	109	85	835	164	163	1242	273	248
Total	2178	467	467	3200	686	686	5378	1153	1153

D. Evaluation Metrics

To evaluate the performance of the proposed model, the weighted average precision (Prec), recall (Rec), and F1-score (F1) were used as the official metrics for implicit emotion recognition in the performed experiments.

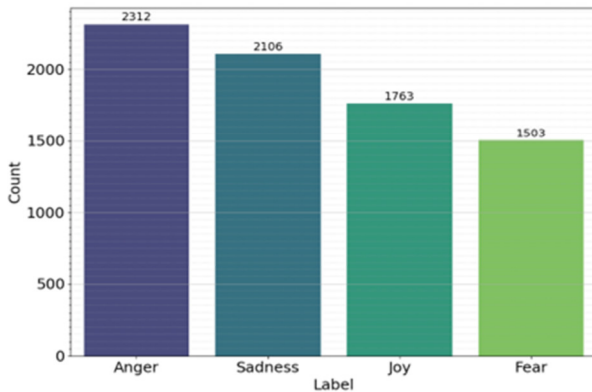


Fig. 2. Emotions distribution in the AIEmoD dataset.

E. Hyperparameters

The hyperparameters of the proposed AraBERT-BiGRU model were carefully selected to optimize performance. Through iterative experimentation with different combinations, the optimal values that yielded the highest accuracy were identified. A summary of the hyperparameters and their corresponding tested and optimal values is provided in Table II.

TABLE II. HYPERPARAMETERS OF THE ARABERT-BIGRU MODEL

Hyper-parameters	Tested values	Optimal values
Learning rate	$2 \cdot 10^{-5}$, $3 \cdot 10^{-5}$, or $5 \cdot 10^{-5}$	$2 \cdot 10^{-5}$
Loss function	Categorical cross-entropy	Categorical cross-entropy
Optimizer	Adam	Adam
Batch size	16, 32, 64	32
Dropout rate	[0.2-0.5]	0.25
GRU	64, 128, 256	256
Epochs	[5-20]	8

For this study's experiment, the AraBERTv0.2-Twitter-based Arabic pre-trained language model was adopted, which includes 12 layers, 768 hidden dimensions, 12 attention heads, and 136M parameters. The model was implemented using the TensorFlow framework. The maximum sequence lengths were set to 48 for the SemEval-2018 and AIEmoD datasets and 35 for the AETD dataset, respectively. To achieve the best results, the training process used a learning rate of $2 \cdot 10^{-5}$, a dropout rate of 0.25, and a maximum of 8 epochs, with an early stopping criterion.

IV. RESULTS AND DISCUSSION

The performance metrics of the proposed AraBERT-BiGRU model using the three datasets SemEval-2018, AETD, and AIEmoD are shown in Table III.

The results demonstrate varying performance across different emotions in the three datasets, with particularly high F1-scores having been achieved for the recognition of joy emotions (88.35%) in the AIEmoD dataset and anger emotions (84.38%) in the AETD datasets. Sadness demonstrates moderate recognition performance, achieving an F1-score of 72.66% on the AETD dataset. Fear emerged as the most challenging emotion to classify, with lower F1-scores across all evaluated datasets. Overall, the model performed best in

recognizing joy emotions, followed by anger, sadness, and finally fear, in descending order of performance across the three datasets.

TABLE III. MODEL PERFORMANCE USING SEMEVAL-2018, AETD, AND AIEMOD IMPLICIT DATASETS

Emotion	SemEval-2018			AETD			AIEmoD		
	Prec (%)	Rec (%)	F1 (%)	Prec (%)	Rec (%)	F1 (%)	Prec (%)	Rec (%)	F1 (%)
Anger	49.39	61.83	54.92	82.89	85.91	84.38	65.84	76.44	70.74
Fear	43.53	46.84	45.12	55.17	53.33	54.24	47.31	47.31	43.35
Sadness	54.10	42.31	47.48	78.95	67.31	72.66	68.60	63.21	65.79
Joy	83.15	78.72	80.87	81.52	89.29	85.23	90.91	85.94	88.35

To compare the efficiency of the proposed hybrid approach against other approaches, a series of experiments was conducted on the three datasets using various deep learning models, including CNN, LSTM, BiLSTM, GRU, BiGRU, and a fine-tuned version of AraBERT. For neural network-based models, this work utilized the pre-trained word embeddings FastText [33] to initialize the networks.

The comparative results are presented in Figure 3 and Table IV. These findings consistently demonstrate that the proposed AraBERT-BiGRU model achieved the best performance metrics across all datasets examined. The superior performance of the AraBERT-BiGRU model over standalone deep learning models is primarily attributed to its ability to generate dynamic, contextualized representations. These allow the model to effectively capture nuanced semantic information by adapting to the linguistic context, thus significantly enhancing word embedding quality. Furthermore, the AraBERT-BiGRU model exhibited approximately a 10% improvement in F1-score compared to neural network models utilizing FastText embeddings. This clearly indicates that AraBERT-BiGRU is more effective in capturing and representing semantic information, leading to superior performance in implicit emotion recognition tasks.

Finally, Table V presents a comparative analysis between the proposed hybrid model and state-of-the-art baseline models for explicit emotion recognition on the AETD dataset. The results clearly show that even when applied to implicit emotions, the proposed model outperforms most baseline models originally designed for explicit emotion recognition.

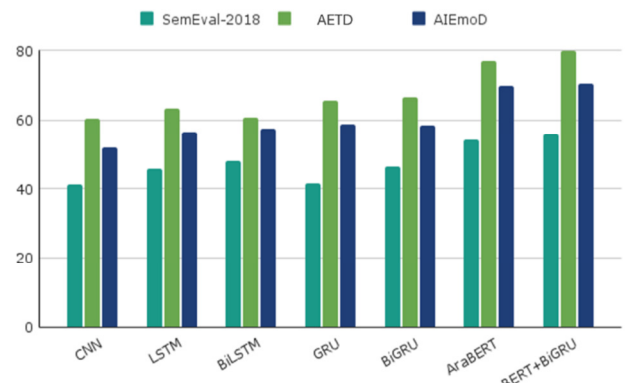


Fig. 3. F1-score comparative results.

TABLE IV. COMPARATIVE RESULTS OF DIFFERENT MODELS ON AETD, SEMEVAL-2018, AND AIEMOD DATASETS

Model	SemEval-2018			AETD			AIEMoD		
	Precision (%)	Recall (%)	F1-score (%)	Precision (%)	Recall (%)	F1-score (%)	Precision (%)	Recall (%)	F1-score (%)
CNN	41.35	42.05	41.18	63.40	61.28	60.31	50.68	55.28	51.99
LSTM	46.40	49.24	45.92	65.96	63.56	63.51	58.07	58.66	56.53
BiLSTM	51.00	51.20	48.14	63.95	61.81	60.83	57.73	59.44	57.64
GRU	43.23	43.36	41.61	67.03	65.16	65.52	59.01	59.83	58.99
BiGRU	46.77	47.28	46.57	67.13	66.47	66.69	60.08	60.12	58.56
AraBERT	55.64	54.35	54.25	76.99	77.18	77.07	69.96	70.25	70.02
AraBERT-BiGRU	56.88	56.09	56.02	79.97	80.14	79.87	70.93	70.83	70.67

TABLE V. COMPARING PERFORMANCES OF THE ARABERT-BIGRU MODEL WITH THE STATE-OF-THE-ART MODELS ON THE AETD DATASET

Study	Method	Explicit/ Implicit	F1-score (%)
[7]	Complement Naïve Bayes	Explicit	65.80
[35]	Human-engineered Feature-based and Deep Feature-based (HEF+DF) model	Explicit	71.80
[34]	AraBERT	Explicit	78.10
[34]	AraBERT Integrated with Named Entity Recognition (NER) and Sentiment Analysis (SA)	Explicit	81.60
Proposed	AraBERT-BiGRU	Implicit	79.87

V. CONCLUSION

The primary contribution of this work lies in the development of an Arabic dataset, named Arabic Implicit Emotion Dataset (AIEMoD), specifically curated for implicit emotion recognition, as well as the integration of Arabic Bidirectional Encoder Representations from Transformers (AraBERT) and Bidirectional Gated Recurrent Units (BiGRU) to address the challenges associated with detecting indirectly expressed emotions. The proposed model was trained and evaluated on three datasets, AIEMoD, AETD, and SemEval-2018. Unlike prior studies that predominantly focused on explicit emotion detection, this work extends the application of modern Natural Language Processing (NLP) models to address implicit emotional cues in Arabic, a morphologically rich language with limited resources in this domain.

The results of the experiments show that the AraBERT-BiGRU hybrid model outperforms all baseline deep learning models, with F1-scores of 79.87 % and 70.67 % on the AETD and AIEMoD datasets, respectively. Compared to existing state-of-the-art models used for explicit emotion tasks, the proposed model achieves superior performance when applied to implicit emotion recognition, underscoring its robustness and adaptability across both explicit and implicit contexts. These findings suggest that contextualized transformer models, when integrated with recurrent architectures, are highly effective at capturing nuanced emotional expressions, particularly in morphologically complex languages, such as Arabic.

For future work, the authors intend to explore the integration of other models to further enhance implicit emotion recognition in Arabic text. For instance, combining a GPT-based model with BERT could potentially improve recognition performance by leveraging GPT's generation capabilities alongside BERT's masked word prediction capabilities.

REFERENCES

- [1] R. Klinger, O. De Clercq, S. Mohammad, and A. Balahur, "IEST: WASSA-2018 Implicit Emotions Shared Task," in *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Brussels, Belgium, 2018, pp. 31–42, <https://doi.org/10.18653/v1/W18-6206>.
- [2] W. Q. A. Saif, M. K. Alshammari, B. A. Mohammed, and A. A. Sallam, "Enhancing Emotion Detection in Textual Data: A Comparative Analysis of Machine Learning Models and Feature Extraction Techniques," *Engineering, Technology & Applied Science Research*, vol. 14, no. 5, pp. 16471–16477, Oct. 2024, <https://doi.org/10.48084/etasr.7806>.
- [3] H. Elfaiik and E. H. Nfaoui, "Deep Bidirectional LSTM Network Learning-Based Sentiment Analysis for Arabic Text," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 395–412, Dec. 2020, <https://doi.org/10.1515/jisys-2020-0021>.
- [4] B. Zemni, M. Zitouni, F. Bouhadiba, and M. Almutairi, "On Ambiguity in the Arabic Language: Scrutinizing Translation Issues through Machine Translation from English and French into Arabic," *Journal of Intercultural Communication*, pp. 203–212, Mar. 2024, <https://doi.org/10.36923/jicc.v24i1.171>.
- [5] O. Rabie and C. Sturm, "Feel the heat: Emotion detection in Arabic social media content," in *The International Conference on Data Mining, Internet Computing, and Big Data (BigData2014)*, Kuala Lumpur, Malaysia, Nov. 2014, pp. 37–49.
- [6] W. A. Hussien, Y. M. Tashtoush, M. Al-Ayyoub, and M. N. Al-Kabi, "Are emoticons good enough to train emotion classifiers of Arabic tweets?," in *2016 7th International Conference on Computer Science and Information Technology (CSIT)*, Amman, Jordan, Jul. 2016, pp. 1–6, <https://doi.org/10.1109/CSIT.2016.7549459>.
- [7] A. Al-Khatib and S. R. El-Beltagy, "Emotional Tone Detection in Arabic Tweets," in *Computational Linguistics and Intelligent Text Processing*, 2018, vol. 10762, pp. 105–114, https://doi.org/10.1007/978-3-319-77116-8_8.
- [8] S. Alzu'bi, O. Badarneh, B. Hawashin, M. Al-Ayyoub, N. Alhindawi, and Y. Jararweh, "Multi-Label Emotion Classification for Arabic Tweets," in *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, Granada, Spain, Oct. 2019, pp. 499–504, <https://doi.org/10.1109/SNAMS.2019.8931715>.
- [9] H. Mulki, C. Bechikh Ali, H. Haddad, and I. Babaoglu, "Tw-StAR at SemEval-2018 Task 1: Preprocessing Impact on Multi-label Emotion Classification," in *Proceedings of The 12th International Workshop on Semantic Evaluation*, New Orleans, Louisiana, 2018, pp. 167–171, <https://doi.org/10.18653/v1/S18-1024>.
- [10] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, "A Convolutional Neural Network for Modelling Sentences." arXiv, Apr. 2014, <https://doi.org/10.48550/arXiv.1404.2188>.
- [11] P. Liu, X. Qiu, and X. Huang, "Recurrent Neural Network for Text Classification with Multi-Task Learning." arXiv, May 2016, <https://doi.org/10.48550/arXiv.1605.05101>.
- [12] A. Graves, "Long Short-Term Memory," in *Supervised Sequence Labelling with Recurrent Neural Networks*, vol. 385, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 37–45.
- [13] T. Mikolov, M. Karafiát, L. Burget, J. Černocký, and S. Khudanpur, "Recurrent neural network based language model," in *Interspeech 2010*,

- Sep. 2010, pp. 1045–1048, <https://doi.org/10.21437/Interspeech.2010-343>.
- [14] M. Abdullah and S. Shaikh, "TeamUNCC at SemEval-2018 Task 1: Emotion Detection in English and Arabic Tweets using Deep Learning," in *Proceedings of The 12th International Workshop on Semantic Evaluation*, New Orleans, Louisiana, 2018, pp. 350–357, <https://doi.org/10.18653/v1/S18-1053>.
- [15] M. Abdullah, M. Hadzikadicy, and S. Shaikhz, "SEDAT: Sentiment and Emotion Detection in Arabic Text Using CNN-LSTM Deep Learning," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Orlando, FL, Dec. 2018, pp. 835–840, <https://doi.org/10.1109/ICMLA.2018.00134>.
- [16] M. Rhanoui, M. Mikram, S. Yousfi, and S. Barzali, "A CNN-BiLSTM Model for Document-Level Sentiment Analysis," *Machine Learning and Knowledge Extraction*, vol. 1, no. 3, pp. 832–847, Jul. 2019, <https://doi.org/10.3390/make1030048>.
- [17] H. Elfaik and E. H. Nfaoui, "Combining Context-Aware Embeddings and an Attentional Deep Learning Model for Arabic Affect Analysis on Twitter," *IEEE Access*, vol. 9, pp. 111214–111230, 2021, <https://doi.org/10.1109/ACCESS.2021.3102087>.
- [18] A. E. Samy, S. R. El-Beltagy, and E. Hassanien, "A Context Integrated Model for Multi-label Emotion Detection," *Procedia Computer Science*, vol. 142, pp. 61–71, 2018, <https://doi.org/10.1016/j.procs.2018.10.461>.
- [19] A. Mansy, S. Rady, and T. Gharib, "An Ensemble Deep Learning Approach for Emotion Detection in Arabic Tweets," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 4, 2022, <https://doi.org/10.14569/IJACSA.2022.01304112>.
- [20] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." arXiv, May 2019, <https://doi.org/10.48550/arXiv.1810.04805>.
- [21] W. Antoun, F. Baly, and H. Hajj, "AraBERT: Transformer-based Model for Arabic Language Understanding." arXiv, Mar. 2021, <https://doi.org/10.48550/arXiv.2003.00104>.
- [22] A. Rozenal, D. Fleischer, and Z. Kelrich, "Amobee at IEST 2018: Transfer Learning from Language Models," in *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Brussels, Belgium, 2018, pp. 43–49, <https://doi.org/10.18653/v1/W18-6207>.
- [23] J. Balazs, E. Marrese-Taylor, and Y. Matsuo, "IIIDYT at IEST 2018: Implicit Emotion Classification With Deep Contextualized Word Representations," in *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Brussels, Belgium, 2018, pp. 50–56, <https://doi.org/10.18653/v1/W18-6208>.
- [24] A. Chronopoulou, A. Margatina, C. Baziotis, and A. Potamianos, "NTUA-SLP at IEST 2018: Ensemble of Neural Transfer Methods for Implicit Emotion Classification," in *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Brussels, Belgium, 2018, pp. 57–64, <https://doi.org/10.18653/v1/W18-6209>.
- [25] Q. Zhou and H. Wu, "NLP at IEST 2018: BiLSTM-Attention and LSTM-Attention via Soft Voting in Emotion Classification," in *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Brussels, Belgium, 2018, pp. 189–194, <https://doi.org/10.18653/v1/W18-6226>.
- [26] H. Fei, Y. Ren, and D. Ji, "Implicit Objective Network for Emotion Detection," in *Natural Language Processing and Chinese Computing*, 2019, vol. 11838, pp. 647–659, https://doi.org/10.1007/978-3-030-32233-5_50.
- [27] M. Noui, A. Lakhfif, and M. A. Laouadi, "Event Detection and Classification in Tweets using Deep Learning," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19977–19982, Feb. 2025, <https://doi.org/10.48084/etasr.9238>.
- [28] D. Li, J. Wang, and X. Zhang, "CIEA: A Corpus for Chinese Implicit Emotion Analysis," in *2019 International Conference on Asian Language Processing (IALP)*, Shanghai, Singapore, Nov. 2019, pp. 90–95, <https://doi.org/10.1109/IALP48816.2019.9037667>.
- [29] Y. Qian, J. Wang, D. Li, and X. Zhang, "Interactive capsule network for implicit sentiment analysis," *Applied Intelligence*, vol. 53, no. 3, pp. 3109–3123, Feb. 2023, <https://doi.org/10.1007/s10489-022-03584-3>.
- [30] S. Mohammad, F. Bravo-Marquez, M. Salameh, and S. Kiritchenko, "SemEval-2018 Task 1: Affect in Tweets," in *Proceedings of The 12th International Workshop on Semantic Evaluation*, New Orleans, Louisiana, 2018, pp. 1–17, <https://doi.org/10.18653/v1/S18-1001>.
- [31] F. Bravo-Marquez, *SemEval-2018 Task 1: Affect in Tweets (AIT-2018) – Datasets*, CodaLab Competitions, 2018. [Online]. Available: https://competitions.codalab.org/competitions/17751#learn_the_details-datasets.
- [32] A. El Khatib, *Emotional-Tone*. GitHub repository. [Online]. Available: <https://github.com/amrmalkhatib/Emotional-Tone>.
- [33] A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov, "Bag of Tricks for Efficient Text Classification." arXiv, Aug. 2016, <https://doi.org/10.48550/arXiv.1607.01759>.
- [34] M. J. Althobaiti, "Emotion Recognition in Arabic: A BERT-Based Transfer Learning Approach Leveraging Semantic Information of Online Comments," *Journal of Theoretical and Applied Information Technology*, vol. 101, no. 9, pp. 3270–3282, May 2023.
- [35] N. Alswaidan and M. E. B. Menai, "Hybrid Feature Model for Emotion Recognition in Arabic Text," *IEEE Access*, vol. 8, pp. 37843–37854, 2020, <https://doi.org/10.1109/ACCESS.2020.2975906>.