Arabic Fake News Detection using Machine Learning Approach

1st Tasneem Duridi
Department of Computer Science
Palestine Technical University
Tulkarm, Palestine
tasneem.duridi@ptuk.edu.ps

2nd Derar Eleyan
Faculty of Telecomm. and Information Technology
Nablus University for Technical and Vocational Education
Nablus, Palestine
d.eleyan@nu-vte.edu.ps
Department of Computer Science
Palestine Technical University - Kadoorie
Tulkarm, Palestine
d.eleyan@ptuk.edu.ps

3rd Amna Eleyan
Department of Computing and Mathematics
Manchester Metropolitan University
Manchester M1 5GD, UK
a.eleyan@mmu.ac.uk

4th Tarek Bejaoui
Computer Engineering Department
University of Carthage
Tunisia
tarek.bejaoui@ieee.org

***Abstract*—In the contemporary digital realm, the widespread dissemination of false Arabic news is a significant social concern laden with various risks. Recognizing the seriousness of this issue, our research utilizes cutting-edge technologies, specifically Machine Learning, to distinguish between authentic Arabic news and deceptive counterparts. The consequences of propagating misinformation go beyond compromising social cohesion; they extend to the erosion of digital information's credibility, fostering an atmosphere of mistrust and deception that undermines the very foundations of societal bonds. Through the application of contemporary technologies, this study aims to identify and underscore Arabic news that embodies such risks.**

The intricate characteristics of Arabic language morphology, marked by words carrying multiple meanings based on inflectional forms and the prevalence of numerous diacritical marks, intensify the challenges of text classification. Despite contending with these linguistic intricacies, modern natural language processing approaches offer practical solutions. Notably, our methodology relies on the preprocessing of two available datasets for training and testing, a crucial step for seamlessly integrating a range of Machine Learning techniques, including Random Forest (RF), Logistic Regression (LR), Decision Tree (DT), and Multinomial Naive Bayes (NB). Importantly, the Logistic Regression technique emerged as the most effective, achieving an accuracy of 95.92% in the SANAD dataset for discerning nuanced Arabic news

***Index Terms*—Machine Learning, Arabic, Fake News, NLP, News, Fake News Detection, Text Classification, SANAD, AraNews.**

I. Introduction

The term "fake news," originating in the 1880s, has evolved to encompass any inaccuracies, irrespective of source. In the contemporary digital landscape, characterized by the prevalence of false or misleading information presented as news on the internet, this phenomenon extends beyond political boundaries, posing significant challenges to national security, corporate stability, and economic well-being. The rapid dissemination of deceptive content, facilitated by instant messaging, bots, and mobile networks, exhibits a seven-fold increase in distribution compared to traditional news channels.

Understanding the multifaceted nature of this issue requires a comprehensive grasp of diverse global media content, ranging from streaming videos to niche blogs. Despite this diversity, individuals tend to place trust in content transparency without critically evaluating its source. The online environment, featuring endless scrolling and abundant information choices, intensifies the competition for consumers' attention. Empirical studies reveal the accelerated propagation of fake news on social media platforms [1], surpassing the speed and reach of real news, thereby exacerbating the overarching challenge [2]. This phenomenon extends beyond social media, eroding the foundational pillars of trustworthy knowledge and imperiling the integrity of digital information [3].

Within this dynamic and complex framework, unscrupulous entities exploit vulnerabilities for personal gain. The deliberate dissemination of false and harmful information, defined as disinformation, manifests through sensational headlines or satirical articles that misrepresent reliable information [4]. This has led researchers to propose the term "information disorder" for a more comprehensive description, acknowledging the intricate dynamics of the digital age. Consequently, there arises an imperative need for extensive efforts to combat the broad dissemination of misleading information, safeguard information integrity, promote transparency, and uphold societal well-being.

In response to these challenges, significant efforts in natural language processing have been devoted to developing methods for identifying rumors (fake news) and confirming the accuracy of facts. Conventional manual approaches for detecting fake news are rendered impractical due to time and cost constraints, underscoring the imperative need for automated methods to address the proliferation of unreliable information and to keep the public informed about the potential falsehoods they may encounter. variety of approaches, ranging from simple to complex models, including semantic-based models [5], [6], and artificial intelligence systems [7] [8], have been employed for fake news detection.

Despite the extensive research conducted in the field, the predominant focus of studies on fake news detection has been on datasets in the English language. This is notable, given that Arabic is the second most-spoken language globally, boasting approximately 440 million speakers across 22 countries. Consequently, the detection of fake news in Arabic poses a significant and enduring challenge. This challenge is primarily attributed to two key factors. Firstly, the absence of large, readily available Arabic fake news datasets poses a hindrance. Secondly, the Arabic language introduces a distinct set of challenges, deviating considerably from English. Consequently, there is an evident need for advancements in text representation and embedding techniques to bolster the learning capabilities of fake news detection algorithms tailored for the Arabic language.

This research is committed to the identification of fabricated Arabic news content, filling a notable void in the existing literature that encompasses news articles across a range of subjects, including politics, technology, sports, and more. Deviating from the predominant focus on tweets in current research, this investigation takes a more comprehensive approach by employing various machine-learning techniques across two distinct datasets, Aranews and SANAD. The central objective is to systematically compare multiple performance metrics and pinpoint the algorithm that excels in effectively discerning fake news data. The study meticulously evaluates the performance of diverse machine learning methods to identify the one that attains the highest accuracy in detecting fabricated news.

The publication is structured into several segments, each fulfilling a distinct role. The 'Related Work' section offers an overview of prior research and its outcomes, facilitating comparisons with findings from other studies. The 'Proposed Methodology' section delves into the selected approaches for addressing the identified issue, providing detailed explanations for each choice. The outcomes gathered from each proposed methodology are discussed in the 'Results and Analysis' section, which also comprehensively assesses the model's strengths. The 'Conclusion and Future Work' section summarizes the findings, analyzes the results, and suggests potential avenues for future research.

II. RELATED WORK

The task of identifying false news in Arabic has gained growing attention, with significant advancements outlined in Table I. Early work focused on analyzing content to assess credibility. For example, [9] used Convolutional Neural Networks (CNNs) with word- and character-level embeddings to evaluate weblog posts, achieving an F1-score of 0.63 despite limited training data.

[10] developed the Arabic News Stance (ANS) dataset for claim verification, using LSTM and pre-trained BERT models. LSTM achieved an F1-score of 0.643 for fake claim detection. Similarly, [11] utilized transformers to generate fake news and introduced AraNews, a dataset annotated with parts of speech, achieving a macro F1-score of 70.06 in detecting manipulated Arabic news.

Linguistic features, such as part of speech and polarity, were explored by [7], who used NB, RF, and SVM classifiers for fake news detection, with RF achieving 79% accuracy. Expanding to sentiment analysis, [12] focused on sarcasm detection in Arabic datasets, with BERT models reaching 91% accuracy in binary classification.

Fake news detection on Twitter has also been explored. [13] used RF classifiers to assess tweet authenticity, achieving an F1-score of 0.776. In rumor detection, [14] applied user-based features with the EM algorithm, showing promising results.

For satirical news, [15] trained CNNs using pre-trained embeddings, achieving 98% accuracy. In another recent study, [16] introduced JointBERT for Arabic fake news detection, outperforming models like AraBERT and AraGPT2 on various datasets.

The COVID-19 pandemic intensified the need for fake news detection. Studies such as [17], [18], and [19] investigated misinformation on platforms like Twitter, highlighting the impact of dataset size and model performance. Notably, [17] used an XGB classifier with an F1-score of 0.39, while [18] demonstrated the effectiveness of extra pre-training on transformer models.

Further, [20] and [21] applied neural networks and ensemble methods to detect false COVID-19 news, with QARiB achieving a 95% accuracy on one dataset but lower performance on new data. A two-model approach by [21] achieved an F1-score of 0.935 using ensemble-based classifiers.

The integration of social context features has also gained traction. [22], [23], and [24] incorporated user behavior and profile data to improve classifier performance. In a notable study, [25] achieved an F1-score of 0.956 by combining content and social context for fake news detection on Twitter.

Finally, [26] employed a bidirectional LSTM-RNN model for rumor prediction, achieving an F1-score of 0.87, outperforming traditional LSTM models.

The existing literature on Arabic fake news detection, as shown in Table I, reveals a significant gap, primarily focusing on tweets rather than news articles. In contrast to the extensive attention given to English studies, current research often neglects the unique challenges posed by fake news in news content.

TABLE I
OVERVIEW OF THE CURRENT DATASETS AND METHODS USED FOR ARABIC FAKE NEWS DETECTIONN

Ref	Year	Data Set	Classification Approach	Result	Gaps and Opportunities
[22]	2018	800 Tweets	DT, SVM, NB	Accuracy 0.899	Need for a larger dataset and to assess the influence of adding or removing features
[14]	2019	271,000 Tweets	Gaussian Naïve Bayes	Accuracy 0.78	Single fact-checking source and dataset unavailability
[9]	2019	268 labeled blog posts, 20,392 unlabeled blog posts	CNN	F1-score 0.63	Limited availability of labeled data for training purposes
[13]	2019	9,000 Tweets	DT, RF, SVM, NB	F1-score 0.776	The new feature does not improve classification outcomes
[23]	2019	1,862 Tweets	AdaBoost, LR, RF, DT	Accuracy 0.76	Considering only the emotional aspect of users sharing news tweets and the limited dataset size
[10]	2020	ANS	LSTM, mBERT	F1-score 0.643	Same limitations in word count as tweets
[11]	2020	AraNews, ATB, ANS	AraBERT, mBERT, XLM-RBase, XLM-RLarge	F1-score 0.70	Datasets are machine-generated from real news
[15]	2020	6,895 satirical news articles	XGBoost, NB, CNN	F1-score 0.984	Focused only on political issues related to the Middle East
[24]	2021	1,862 Tweets	LR, LDA, SVM, RF, XGBoost, KNN, DT, NB	Accuracy 0.82	Complexity of Arabic dialects; suggests exploring other meta-heuristic algorithms
[19]	2021	5.5 Million COVID-19 Tweets	XGB, RF, NB, LR, SVM, MLP	F1-score 0.933	Restricted COVID-19 keyword constraints
[18]	2021	10,828 COVID-19 Tweets	DistilBERT-multi, AraBERT COV19, AraBERT, mBERT, mBERT COV19	F1-score 0.9578	Update dataset with the latest COVID-19 events
[20]	2021	ANS, AraNews, COVID-19-Fakes, ArCOV19-Rumors	CNN, RNN, MARBERT, GRU, AraBERT v1, v2, v02, QARiB, Ar-Electra, ArBERT	F1-score 0.95	Small dataset and noise in the data
[17]	2021	8,786 COVID-19 Tweets	XGB, RF, SVM, NB, CNN, RNN, CRNN, SGD	F1-score 0.39	Failed to enhance the performance of all classifiers
[21]	2022	3,157 COVID-19 Tweets	ExtraTree, LR, KNN, CART, AdaBoost, SVM, NB, RF, Bagging	F1-score 0.935	Enhance the dataset by gathering more tweets
[27]	2022	4,299 COVID-19 Tweets	LR, RNN, BiRNN, DT, XGBoost, SVM, KNN, NB, SGD, BiLSTM, GRU, BiGRU, LSTM	Accuracy 0.81	Explore a pre-trained transformer and expand the dataset beyond COVID-19
[7]	2022	1,098 pilgrimage news articles	RF, NB, SVM	F1-score 0.79	Add more textual features
[16]	2022	COVID-19, ANS, AraNews, Satirical News	QARiB, jointBERT, AraGPT2, AraBERT	F1-score 0.85	Enhance for multi-class fake news categorization
[12]	2022	Misogyny, Sarcasm	PAC, LRC, RFC, LSVC, DTC, KNNC, ARABERT	Sarcasm 0.88, Misogyny 0.91	Limited benchmarks, dataset imbalances, and mixed-language challenges
[26]	2022	PHEME	Bi-LSTM, LSTM-RNN	F1-score 0.87	Rigorous tests needed to understand deep learning in rumor detection
[12]	2022	COVID-19	RoBERTa, BERTweet, CT-BERT	F1-score 0.99	Lack of data and unbalanced dataset
[25]	2023	5,000 news Tweets	CNN, MARBERT	F1-score 0.956	Unbalanced dataset; need to identify more user features to distinguish fake news publishers
[28]	2023	644 news	Mini-BERT	F1-score 0.98	Mini-BERT model performance may vary with larger datasets

This trend prioritizes the detection of fake news in tweets, overlooking the need for a comprehensive approach that includes diverse media sources. While both ”fake news” and ”fake tweets” involve false information, their contextual differences warrant targeted research on news articles.

Additionally, the impact of text length poses challenges for Natural Language Processing (NLP) models [29], which may struggle with longer news articles compared to shorter tweets, affecting context comprehension and feature extraction.

Our research aims to implement a machine learning approach to detect fake Arabic news in news datasets. Although some studies, like [20] and [30], have shown the effectiveness of deep learning models, their complexity, extended training times, and high computational requirements must also be considered.

III. METHODOLOGY

In our study, we utilized two openly accessible datasets, AraNewa and SANAD. AraNewa underwent annotation for binary classification tasks (Fake and Real), while SANAD comprises a single classification with two distinct categories: one designated for fake news and the other for real news. Both datasets were purposefully designed to discern between authentic and fabricated news.

Fig. 1. Suggested Structure for Arabic News Recognition

Various machine learning algorithms were employed, and each underwent a thorough evaluation before being applied to train models on the datasets. Following the outlined steps in Figure 1, it becomes possible to preprocess the datasets, train a machine learning model, and assess its performance.

A. Data Set Selection

The initial and pivotal step in our classification method involves sourcing an appropriate dataset. In our investigation, we used the AraNews and datasets

- SANAD: The Single-labeled Arabic News Articles Dataset (SANAD) plays a crucial role in the field of Natural Language Processing (NLP) for Modern Standard Arabic (MSA). Comprising nearly 200,000 articles sourced from reputable news portals like AlKhaleej, AlArabiya, and Akhbarona, SANAD stands as an essential asset for research in Arabic computational linguistics [31].
  Organizing its content into seven distinct categories (culture, finance, politics, religion, sports, and tech), SANAD offers a diverse range of articles for analysis. With meticulous compilation, the dataset is specifically designed to support various NLP tasks, emphasizing single-label text classification. In the research experiment, a subset of available SANAD comprising 10,000 rows is employed in experimental procedures. The research aims to demonstrate the applicability and effectiveness of machine learning models within the scope of 10,000 articles from those seven categories.
- AraNews: a comprehensive dataset, has been developed to explore misinformation in Arabic news. Figure 2 displays a subset of the gathered AraNews News. The dataset, comprising a vast collection of news articles spanning various topics and countries, is meticulously curated. A list of fifty newspapers from fifteen Arab nations, the United States of America (USA), and the United Kingdom (UK) must be manually compiled as part of the creation process [11].

Text	labels
"وطالب مسير المحطة شركة نفطال بمضاعفة التوزيع حتى يتم امتصاص الطلب الكبير الذي ساهمت فيه الاشاعات على حد تعبيره , على اعتبار ان شائعة اضراب	1
"حيث قال السانتور في مداخلة له على هامش مناقشة قانون المالية لسنة 8897 : ان هذه السياسة لا تتماشى مع توصيات وتعليمات الرئيس عبد العزيز بوتفليقة والوز	0
"واطلق عبد نجيب العنان لخياله وافكاره فراح ينتقي البذور وشتلات وشجيرات الورود من كل حدب وصوب من داخل الوطن وخارجه ليحول بذلك قطعة ارض قاحلة	0
" وعن السؤال ذاته يجيب سعيدي بالقول الحركة تعرف اختلافا داخليا بخصوص موقعها من المعارضة , فبينما البعض يراها او يريد لها ان تكون القوة السياسية الاك	0
"الجزائريون كانوا يتوقفون في القدس لم يكن الحج رحلة زيارة لبيت الله الحرام فقط , وشراء للهدايا للاهل والاصحاب , بل كان رحلة علم وسياحة دنيوية وروحية ,	0
"المهم , وصلنا للمرحلة التي نقدر فيها تقييم الامور , ونقول لماذا لم تكتب ديمومة كبرى للدساتير الاربعة في الحياة لسبب واحد انها لم تكن دساتير اعدت باسلوب وا	0
"ويبدو من الرقم الاجمالي للنزاعات العمالية الذي تجاوز نصف مليون نزاع , ان العقد الاقتصادي والاجتماعي , الذي تشكل المركزية النقابية احد اطرافه لم يتكفل بها	0

Fig. 2. sub set of AraNews Dataset

B. Data Pre-processing

- **Normalization of Orthography:**
  Standardize Alef characters and The characters in the input text using the `normalize_hamza`. `normalize_alef`. `normalize_teh`. functions available in the Araby library.
- **Elimination of the English Text:** Make use of the araby library to tokenize every text input. Tokens outside of the Arabic Unicode range should be excluded. To create a coherent Arabic text, combine the remaining Arabic tokens. The processed Arabic texts should be saved in a new list.
- **Extra Preparation**: Use the Araby library's `strip_tashkeel` function to remove tashkeel (diacritical marks) from each Arabic text. To guarantee that the characters in the Arabic text are consistently represented, further implement orthographic normalization. Eliminate characters that are repeated. Make tokens out of the normalized text. Tokens outside of the Arabic Unicode range should be excluded. Arabic tokens can be stemmed using ISRlStemmer. Remove stop words, which are frequently occurring words with little semantic significance, using the nltk library's stop words list as a guide. Keep non-alphabetic tokens out of it. To create a processed text, concatenate the filtered tokens. The edited texts should be saved in a new list.

C. Feature Extraction

In this investigation, the TF-IDF [35] methodology for feature extraction was employed to assess the significance of words within documents, determined by their frequency. TF-IDF is the product of term frequency (TF) and inverse document frequency (IDF). A higher score is assigned to a term when it exhibits frequent occurrence within a document but is infrequent across the entire corpus. $(tf - idf)$ algorithm is calculated using the formula:

$$tf - idf(w) = \text{tf} \times \log\left(\frac{N}{\text{df}(w)}\right)$$

Where:

tf is the term frequency of the word $'w'$ in the document,
N is the total number of documents in the corpus,
$\text{df}(w)$ is the number of documents in the corpus that contain the word '

D. Model Training

- **Data Splitting:** We employ the train˙test˙split function from the scikit-learn library to partition the dataset into distinct training and testing sets. By setting the test˙size parameter to 0.25, a quarter of the dataset is reserved for testing purposes.
- **Text Vectorization with TF-IDF:** In the process of text vectorization with TF-IDF, we utilize the TfidfVectorizer from the scikit-learn feature extraction module. This vectorizer is initialized to convert textual data into TF-IDF features, a widely adopted method in natural language

processing. The fit˙transform method is then employed on the training data (x˙train) to both fit the vectorizer to the data and transform the text into TF-IDF features. The acquired vocabulary is subsequently used to apply the same transformation to the testing data (x˙test) using the transform method. This approach ensures that the text data is effectively represented in a numerical format conducive to machine learning applications.
these steps facilitate the preparation of text data for classification by partitioning it into training and testing sets and employing TF-IDF vectorization to convert text features into numerical representations. The resulting TF-IDF features are consequently primed for training a machine learning model.

E. Performance Assessment

Accuracy alone is insufficient for a complete evaluation of machine learning models. To gain deeper insights, additional metrics like F1 score, precision, and recall are necessary.

Accuracy measures the proportion of correct predictions:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

The F1 score balances precision and recall, making it effective for imbalanced datasets. Precision reflects how many positive predictions were correct:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

Recall shows how well the model identifies true positives:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

The F1 score combines these metrics for a more rounded performance measure:

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

IV. Results and Analysis

Table II presents the findings of the fake news detection experiment, contrasting the efficacy of various machine learning algorithms on the SANAD and AraNews datasets. Metrics including F1-score, Accuracy, Precision, Recall, and Training time are included in the analysis.

For the SANAD dataset, Logistic Regression demonstrates the highest Accuracy at 95.92%, outperforming Decision Tree (85.32%), Random Forest (88.96%), and Multinomial NB (91.68%). Precision and Recall also follow similar trends, with Logistic Regression consistently leading across all metrics. The F1-score further confirms the superior performance of Logistic Regression.

In contrast, the AraNews dataset exhibits lower overall performance across all algorithms. Logistic Regression again achieves the highest Accuracy at 59.13%, followed by Decision Tree (48.6%), Random Forest (47.43%), and Multinomial NB (50.12%). Precision, Recall, and F1-score metrics echo this pattern, with Logistic Regression consistently outperforming other algorithms.

TABLE II
PERFORMANCE COMPARISON OF MACHINE LEARNING MODELS ON ARABIC FAKE NEWS DETECTION DATASETS

Dataset	**Metric**	**Logistic Regression**	**Decision Tree**	**Random Forest**	**Multinomial NB**
SANAD	Accuracy	95.92%	85.32%	88.96%	91.68%
	Precision	0.96	0.85	0.89	0.92
	Recall	0.96	0.85	0.89	0.92
	F1-Score	0.96	0.85	0.89	0.92
	Training Time (s)	1.04	5.15	1.35	0.03
AraNews	Accuracy	59.13%	48.60%	47.43%	50.12%
	Precision	0.59	0.49	0.47	0.50
	Recall	0.59	0.49	0.47	0.50
	F1-Score	0.59	0.49	0.34	0.50
	Training Time (s)	0.24	4.34	4.88	0.01

Training time analysis reveals variations across algorithms. Logistic Regression generally exhibits lower training times compared to Decision trees, Random Forests, for both datasets. Multinomial NB consistently has the lowest training times, indicating its efficiency in model training.

The observed variations in results between the SANAD and AraNews datasets can be attributed to several factors. One notable distinction lies in the word count distribution, where AraNews exhibits a lower word count in its news articles compared to SANAD. This discrepancy echoes the inherent challenges posed by constrained word counts, a similarity shared with the limitations associated with word count constraints in tweets compared to news articles. Furthermore, AraNews represents a publicly available dataset focusing on misinformation in the Arabic language, generated through automated manipulation of authentic stories sourced from 50 newspapers spanning 15 Arab countries, AraNews maintains a balanced composition. However, it is crucial to note that this dataset is synthetically generated, introducing considerations related to its artificial nature. On the other hand, SANAD's data collection process involved the use of Python scripts tailored for scraping three popular news portals. These scripts were designed to extract text and tags from each article's page, providing a more organic and authentic representation of news content compared to the manipulated nature of AraNews.

V. CONCLUSION AND FUTURE WORK

In recent years, detecting fake news has garnered significant attention due to its rapid spread and harmful effects. While the field of Arabic fake news detection is still emerging, it offers promising opportunities for research, particularly given the scarcity of existing studies. The complexities of the Arabic language—such as its diverse dialects, varying orthographic rules, extensive vocabulary, and limited datasets—complicate the identification of fake news.

This study presents a novel approach to determining the authenticity of Arabic news articles. Our exploration of various machine learning algorithms revealed that the Logistic Regression model performed the best in terms of accuracy. However, the limited availability of classified Arabic news datasets presents challenges that may affect the generalizability of our findings.

Looking ahead, we will explore the application of other NLP transformer models, such as BERT, in fake news detection. Our aspirations extend beyond text-based analysis alone. In the future, we aim to enhance our fake news detection capabilities by incorporating image-based features into our models. By combining textual and visual information, we seek to improve the robustness and comprehensiveness of our detection methods. This evolution represents a forward-looking approach, embracing the potential of multi-modal data for more accurate and nuanced fake news identification in the ever-evolving landscape of information dissemination.

REFERENCES

[1] A. A. Abubaker, D. Eleyan, A. Eleyan, T. Bejaoui, N. Katuk, and M. Al-Khalidi, "Social engineering in social network: a systematic literature review," in *2023 International Symposium on Networks, Computers and Communications (ISNCC)*, pp. 1–7, IEEE, 2023.

[2] D. S. Abd Elminaam, A. Abdelaziz, G. Essam, and S. E. Mohamed, "Arafake: A deep learning approach for arabic fake news detection," in *2023 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)*, pp. 1–8, IEEE, 2023.

[3] T. Durid, D. Eleyan, A. Eleyan, T. Bejaoui, and M. Jazzar, "Cybersecurity in fitness devices-a systematic literature review," in *2023 International Conference on Smart Applications, Communications and Networking (SmartNets)*, pp. 1–6, IEEE, 2023.

[4] R. Arul, K. Vishnu, A. Eleyan, and A. K. Bashir, "The authenticity of information on social media," *IEEE Technology Policy and Ethics*, vol. 5, no. 6, pp. 1–6, 2020.

[5] V. Sabeeh, M. Zohdy, and R. Al Bashaireh, "Enhancing the fake news detection by applying effective feature selection based on semantic sources," in *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 1365–1370, IEEE, 2019.

[6] A. B. Hani, O. Adedugbe, F. Al-Obeidat, E. Benkhelifa, and M. Majdalawieh, "Fane-kg: A semantic knowledge graph for context-based fake news detection on social media," in *2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pp. 1–6, IEEE, 2020.

[7] H. Himdi, G. Weir, F. Assiri, and H. Al-Barhamtoshy, "Arabic fake news detection based on textual analysis," *Arabian Journal for Science and Engineering*, vol. 47, no. 8, pp. 10453–10469, 2022.

[8] M. Umer, Z. Imtiaz, S. Ullah, A. Mehmood, G. S. Choi, and B.-W. On, "Fake news stance detection using deep learning architecture (cnn-lstm)," *IEEE Access*, vol. 8, pp. 156695–156706, 2020.

[9] C. Helwe, S. Elbassuoni, A. Al Zaatari, and W. El-Hajj, "Assessing arabic weblog credibility via deep co-learning," in *Proceedings of the Fourth Arabic Natural Language Processing Workshop*, pp. 130–136, 2019.

[10] J. Khouja, "Stance prediction and claim verification: An arabic perspective," *arXiv preprint arXiv:2005.10410*, 2020.

[11] E. M. B. Nagoudi, A. Elmadany, M. Abdul-Mageed, T. Alhindi, and H. Cavusoglu, "Machine generation and detection of arabic manipulated and fake news," *arXiv preprint arXiv:2011.03092*, 2020.

[12] A. Y. Muaad, H. Jayappa Davanagere, J. Benifa, A. Alabrah, M. A. Naji Saif, D. Pushpa, M. A. Al-Antari, T. M. Alfakih, *et al.*, "Artificial intelligence-based approach for misogyny and sarcasm detection from arabic texts," *Computational Intelligence and Neuroscience*, vol. 2022, 2022.

[13] R. Mouty and A. Gazdar, "The effect of the similarity between the two names of twitter users on the credibility of their publications," in *2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*, pp. 196–201, IEEE, 2019.

[14] S. M. Alzanin and A. M. Azmi, "Rumor detection in arabic tweets using semi-supervised and unsupervised expectation–maximization," *Knowledge-Based Systems*, vol. 185, p. 104945, 2019.

[15] H. Saadany, E. Mohamed, and C. Orasan, "Fake or real? a study of arabic satirical fake news," *arXiv preprint arXiv:2011.00452*, 2020.

[16] W. Shishah, "Jointbert for detecting arabic fake news," *IEEE Access*, vol. 10, pp. 71951–71960, 2022.

[17] S. Alqurashi, B. Hamoui, A. Alashaikh, A. Alhindi, and E. Alanazi, "Eating garlic prevents covid-19 infection: Detecting misinformation on the arabic content of twitter," *arXiv preprint arXiv:2101.05626*, 2021.

[18] M. S. H. Ameur and H. Aliane, "Aracovid19-mfh: Arabic covid-19 multi-label fake news & hate speech detection dataset," *Procedia Computer Science*, vol. 189, pp. 232–241, 2021.

[19] A. R. Mahlous and A. Al-Laith, "Fake news detection in arabic tweets during the covid-19 pandemic," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 778–788, 2021.

[20] M. Al-Yahya, H. Al-Khalifa, H. Al-Baity, D. AlSaeed, and A. Essam, "Arabic fake news detection: comparative study of neural networks and transformer-based approaches," *Complexity*, vol. 2021, pp. 1–10, 2021.

[21] S. N. Qasem, M. Al-Sarem, and F. Saeed, "An ensemble learning based approach for detecting and tracking covid19 rumors.," *Computers, Materials & Continua*, vol. 70, no. 1, 2022.

[22] S. F. Sabbeh and S. Y. Baatwah, "Arabic news credibility on twitter: An enhanced model using hybrid features.," *journal of theoretical & applied information technology*, vol. 96, no. 8, 2018.

[23] G. Jardaneh, H. Abdelhaq, M. Buzz, and D. Johnson, "Classifying arabic tweets based on credibility using content and user features," in *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*, pp. 596–601, IEEE, 2019.

[24] T. Thaher, M. Saheb, H. Turabieh, and H. Chantar, "Intelligent detection of false information in arabic tweets utilizing hybrid harris hawks based feature selection and machine learning models," *Symmetry*, vol. 13, no. 4, p. 556, 2021.

[25] S. Alyoubi, M. Kalkatawi, and F. Abukhodair, "The detection of fake news in arabic tweets using deep learning," *Applied Sciences*, vol. 13, no. 14, p. 8209, 2023.

[26] T. Ahmad, M. S. Faisal, A. Rizwan, R. Alkanhel, P. W. Khan, and A. Muthanna, "Efficient fake news detection mechanism using enhanced deep learning model," *Applied Sciences*, vol. 12, no. 3, p. 1743, 2022.

[27] G. Amoudi, R. Albalawi, F. Baothman, A. Jamal, H. Alghamdi, and A. Alhothali, "Arabic rumor detection: A comparative study," *Alexandria Engineering Journal*, vol. 61, no. 12, pp. 12511–12523, 2022.

[28] H. M. Alawadh, A. Alabrah, T. Meraj, and H. T. Rauf, "Attention-enriched mini-bert fake news analyzer using the arabic language," *Future Internet*, vol. 15, no. 2, p. 44, 2023.

[29] M. YAHYA, D. ELEYAN, and A. ELEYAN, "A systematic literature review of automatic keyword extraction algorithms: Textrank and rake," *Journal of Theoretical and Applied Information Technology*, vol. 99, no. 20, 2021.

[30] N. O. Bahurmuz, G. A. Amoudi, F. A. Baothman, A. T. Jamal, H. S. Alghamdi, and A. M. Alhothali, "Arabic rumor detection using contextual deep bidirectional language modeling," *IEEE Access*, vol. 10, pp. 114907–114918, 2022.

[31] O. Einea, A. Elnagar, and R. Al Debsi, "Sanad: Single-label arabic news articles dataset for automatic text categorization," *Data in brief*, vol. 25, p. 104076, 2019.