IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 15 February 2024, accepted 17 April 2024, date of publication 22 April 2024, date of current version 9 May 2024.
Digital Object Identifier 10.1109/ACCESS.2024.3392297

RESEARCH ARTICLE

Improving Prediction of Arabic Fake News Using ELMO's Features-Based Tri-Ensemble Model and LIME XAI

TURKI ALJREES
Department of Computer Science and Engineering, University of Hafr Al-Batin, Hafar Al-Batin 39524, Saudi Arabia
e-mail: tajrees@uhb.edu.sa

ABSTRACT The proliferation of fake news poses a substantial and persistent threat to information integrity, necessitating the development of robust detection mechanisms. In response to this challenge, this research specifically focuses on the detection of Arabic fake news, employing a sophisticated approach that leverages textual features and a powerful stacking classifier. The proposed model ingeniously combines bagging, boosting, and baseline classifiers, strategically harnessing the unique strengths of each to create a resilient ensemble. Through a series of extensive experiments and the integration of Embeddings from Language Models (ELMO) word embedding, the proposed approach achieves remarkable results in the realm of Arabic fake news detection. The model's effectiveness is further heightened by the utilization of advanced stacking techniques, coupled with meticulous textual feature extraction. This capability enables the model to effectively distinguish between real and fake news in Arabic, highlighting its potential to enhance the accuracy of information. The findings of this study hold significant implications for the field of fake news detection, especially within the context of the Arabic language. The proposed model emerges as a valuable tool, contributing to the enhancement of information veracity and fostering a more informed public discourse in the face of misinformation challenges. Furthermore, the excellence of the proposed model is substantiated by its outstanding performance metrics, boasting a 99% accuracy, precision, recall, and F-score. This substantiation is underscored through a comprehensive performance comparison with other state-of-the-art models, affirming the model's reliability in the domain of Arabic fake news detection.

INDEX TERMS Arabic fake news, text mining, ensemble learning, word embedding.

I. INTRODUCTION

The advent of the Internet has revolutionized global connectivity, mainly due to its cost-effectiveness and the absence of regulatory barriers. This expansive medium has ushered in unprecedented avenues for the swift and seamless exchange of information, positioning it as the preferred medium for news consumption, supplanting traditional outlets. However, this rapid expansion has given rise to pressing issues and hurdles, notably the proliferation of fake news [1], [2]. Compounding the issue is the reality that individuals can effortlessly voice their perspectives on any subject and disseminate these views across a multitude of networks and platforms, reaching diverse audiences. This freedom results in a vast influx of unchecked content. To combat the rapid and unchecked dissemination of false narratives, scholarly endeavors have introduced automated or partially automated methods for assessing the truthfulness of such content [3]. Typically, these verification procedures commence with the aggregation of data from various social media platforms including Twitter, Facebook, YouTube, and others [4], [5]. These platforms are often pinpointed as primary conduits for the distribution of fabricated stories across the web. These deceptive pieces can manifest in various guises: from clickbait to fabricated tales and hoaxes, they represent just a few of the numerous incarnations of disinformation found online and in traditional media [6]. It is possible to discern fake news from legitimate reporting. The study [7] suggests

The associate editor coordinating the review of this manuscript and approving it for publication was Yin Zhang.

that the propagation rate of fake news eclipses that of factual reporting. Furthermore, [8] defines fake news as intentionally crafted falsehoods, designed to be perceived as truth for particular motives. A notorious instance of this, as indicated in [9], is the 2016 U.S. presidential election.

The transformation of communication into digital forms has transcended Western boundaries, firmly establishing a foothold in Arab territories. Social media has become a nearly omnipresent force in the Middle East, with 79% of the populace engaging with social media or instant messaging services daily [10]. Furthermore, approximately two-thirds of the Middle Eastern population turn to social media platforms regularly to stay updated with the latest news [11]. Social media platforms have become a critical news and information conduit for young adults. Messaging apps like WhatsApp, Facebook Messenger, Snapchat, and LINE are increasingly the default mediums for personalized news curation. Users are exposed to news stories deemed share-worthy by their chosen circle of friends, creating a filtered news feed tailored to their network's preferences.

Social networks serve as tools for disseminating the latest news to the public, yet they also have the potential to spread fake news. Twitter, in particular, stands out as one of the most extensively used social networking platforms in the Arabian region [12]. Sharing the news on Twitter is economical and swift, with minimal expenses and time required compared to other channels. Its straightforward design and the absence of rigorous content oversight allow for the rapid and widespread dissemination of fake news among users [13]. Fake news is defined as news that is intentionally false, misleading, or fabricated, disseminated with the purpose of deception [14]. The spread of fake news is typically orchestrated to mislead the audience for political, social, or financial benefits. This practice poses substantial threats to individuals, organizations, and governments, as it can influence public opinion, destabilize markets, and disrupt societal harmony [15]. Accordingly, there is a pressing necessity for the development and implementation of effective methods to identify and mitigate fake news on social networks, to avert its detrimental effects.

Numerous fact-checking platforms, like the Anti-Rumors Authority and Misbar, have emerged to assess the truthfulness of news circulating online, representing initial efforts to mitigate the influence of manufactured stories. These sites rely on human experts to manually validate news authenticity, a process that, while thorough, is time-consuming and labor-intensive, making it challenging to scale with the sheer quantity of news that requires verification. To address these limitations, recent trends have seen the adoption of machine learning and deep learning techniques, which have shown promise in addressing complex issues like fake news detection more efficiently [16]. Presently, two distinct learning approaches are utilized to develop automated systems for the detection of false news on social networks: content-based and context-based learning. The content-based approach concentrates on the text's writing style, seeking out syntactic and semantic patterns that can help categorize the news. On the other hand, the context-based approach examines patterns in user behavior and interaction within social media platforms. It gleans insights from user profiles, discourse, and the interconnections within user networks.

The motivation behind this study is rooted in the need to safeguard individuals, organizations, and governments from the detrimental effects of fake news in the Arab context. Fake news has the potential to influence public opinion, destabilize markets, and disrupt societal harmony. It poses not only a threat to the credibility of information but also to the stability and well-being of communities. In the Arab context, where geopolitical and social dynamics are intricate, the impact of misinformation can be even more pronounced. The research seeks to not only understand the linguistic features indicative of deception in Arabic text but also to propose a comprehensive framework that addresses the nuances of the Arab information landscape. To systematically address the multifaceted challenge of fake news in the Arab context, this study is driven by the following research questions, each strategically formulated to guide the investigation and contribute to a comprehensive understanding and mitigation of the issue:

1) What is the prevalence of fake news in Arabic social media platforms, especially focusing on Twitter?
2) How can an automated model effectively classify Arabic fake news through the analysis of linguistic features in Arabic text using natural language processing (NLP) and supervised machine learning techniques?
3) How effective are machine learning and deep learning techniques in comparison to human-driven validation in mitigating the influence of fake news in the Arabic information ecosystem?
4) What are the distinctive linguistic features indicative of deception in Arabic text, and how can they be effectively identified and utilized for fake news detection?

This research aims to construct an automated model capable of classifying Arabic fake news through the analysis of Arabic text. It employs natural language processing (NLP) and supervised machine learning techniques. News articles containing unverified or incorrect information are labeled as fake, while those with fully authenticated information are deemed real. The goal is to identify distinct linguistic features that could indicate deception, using these markers to pinpoint and flag fake news articles. The primary contributions of this research are as follows

- This study introduces a comprehensive framework designed for the identification of Arabic fake news through the utilization of textual data. The prediction of Arabic fake news is carried out through the application of an ensemble approach involving bagging and boosting classifiers.
- For performance comparison, this study employs a diverse set of machine learning models, encompassing

extreme gradient boosting (XGB), random forest (RF), stochastic gradient descent (SGD), K-nearest neighbor (KNN), extra-trees classifier (ETC), and gradient boosting machine (GBM) models. Moreover, to facilitate comparison, the effectiveness of the proposed system is compared with several established state-of-the-art techniques, utilizing commonly recognized evaluation metrics such as accuracy, precision, recall, and F1 score.
- The results are further compared with state-of-the-art approaches and 5-fold cross-validation techniques to show the significance of the proposed model.

The structure of this study is as follows: Section II offers an overview of prior research in this field. Section III explains the overall approach to solving the problem and the steps involved. Section IV presents the assessment of the proposed approach, including experimental results and related discussions. Lastly, Section V serves as the conclusion for this study.

II. RELATED WORK

Arabic language NLP is a fascinating and difficult research area with a variety of topics and tasks. Besides detecting fake news and spam, there are other relevant and significant tasks to start with, such as Arabic sentiment analysis and an Arabic question-answering system. In this regard number of studies have been conducted for the Arabic fake news detection. The most notable and recent studies for Arabic fake news detection are described in this section of the study.

Maha Al-Yahya et al. [17] used the neural network and transformer-based language models for Arabic fake news detection. In this study, the authors also compared the performance of the deep learning models and transformer-based language models. Results of the study show that the transformer-based language models outperformed the deep learning models in terms of accuracy. Transformer-based model QARiB achieved an accuracy of 95% while the deep learning model gated recurrent unit (GRU) achieved the highest accuracy of 83%. Alkair et al. [18] worked on fake news detection in the Arabic language through YouTube comments. The authors used multiple machine learning and deep learning models for classifying the comments into rumour and non-rumour categories. The result shows that the deep learning model convolutional neural network (CNN) achieved an accuracy of 95%.

Mahlous and Laith [19] introduced an automated system for categorizing fake news in Arabic tweets amid the COVID-19 pandemic. To accomplish this task, they gathered a dataset from Twitter spanning from January 2020 to August 2020 and employed several feature engineering techniques for feature extraction. The authors applied various machine learning models, including logistic regression (LR), XGB, support vector machine (SVM), multilayer perceptron (MLP), RF, and Naive Bayes (NB). The findings indicate that the LR model outshone the others, achieving an accuracy of 93.3% when utilizing N-gram-level term frequency-inverse document frequency (TF-IDF) features. Khanam et al. [20] proposed a machine learning-based approach for automatically classifying fake news. For the fake new classification, the authors used the 'politifacto.com' dataset. They used decision tree (DT), RF, NB, KNN, XGB, SVM and linear regression models for experiments. Results indicate that SVM in combination with neural networks achieved the highest accuracy of 99.90%.

Himdi et al. [21] introduced a system for detecting fake news in the Arabic language, employing textual analysis as the foundation. They applied various data preprocessing techniques and feature extraction methods, extracting features like Arabic PoS, emotion, polarity, and linguistic features. To perform the classification task, they utilized machine learning models, including SVM, RF, and NB. The dataset utilized in this study was the Hajj news dataset. The study's findings reveal that the RF attained an accuracy score of 78%. Wotaifi and Dhannonn [22] introduced an effective approach for detecting Arabic fake news, consisting of a deep ensemble model. They utilized deep learning models, including CNN, long short-term memory (LSTM), and an ensemble of CNN-LSTM. The results demonstrated that the CNN-LSTM ensemble model achieved an accuracy of 91.4%.

Najadat et al. [23] proposed a deep learning-based system for fake news detection in Arabic headline-article news data. The study utilized advanced deep learning models, specifically AFND-LSTM and AFND-CNN-LSTM, to automatically identify false news within the Arabic Language Technologies dataset. The findings indicate that the AFND-CNN-LSTM model attained a 70% accuracy in detecting fake news. Nassif et al. [24] introduced a deep contextualized embedding approach tailored for discerning fake news in Arabic. They initially took an English-language dataset of fake news and converted it into Arabic. The models employed in their research included MARBERT, Araelectra, Arabert, GigaBert, Roberta, QaribBert, Arabic-BERT, and ARBERT. Among these, the ARBERT model stood out by reaching an impressive accuracy level of 98.8%, surpassing the performance of the latest advanced models.

In the same vein, Fouad et al. [25] employed deep learning models to automate the detection of fake news in Arabic. They experimented with various neural network architectures, including CNN, LSTM, CNN+LSTM, BiLSTM, and CNN+BiLSTM. Their study was conducted using two distinct datasets, with experiments carried out normalization before classification. ut separately on each dataset. Additionally, a third set of experiments was conducted on a combined dataset, merging the two original sets. The study found that the BiLSTM model was the most accurate compared to other employed models. Alyoubi et al. [26] applied deep learning techniques to identify fake news within Arabic tweets. Their model took into account both the content of the news and the social background of the users spreading the news. They conducted thorough testing with two deep learning methods combined with different word embedding

TABLE 1. Summary of related work.

Ref.	Classifiers	Dataset	Achieved Accuracy
[17]	CNN, RNN, GRU, and transformers	ANS, ArCOV19-Rumors, and AraNews datasets.	95% use QARiB transformers
[18]	SVM,CNN, DT, BiLSTM MNB, LSTM	You tube self-collected	CNN 95%
[19]	NB, LR, XGB, RF, MLP, SVM	Twitter data Self-collected	93.3 % LR with TF-IDF at the n-gram level
[20]	DT, RF, NB, k-NN, XGB, SVM and Linear regression	Politifact dataset	99.90% SVM
[21]	NB, RF, SVM	Hajj related news	78% RF
[22]	CNN, LSTM, CNN-LSTM	AraNews	91.4% CNN-LSTM
[23]	AFND-LSTM, AFND-CNN-LSTM	Arabic language Technologies,	70% AFND-CNN-LSTM
[24]	Roberta, Araelectra, ARBERT, GigaBert, Arabert, Arabic-BERT, MARBERT, and QaribBert	kaggle	98.8% ARBERT
[25]	CNN, LSTM, CNN+LSTM, BiLSTM, CNN+BiLSTM	Arabic news portals dataset and Francisco study dataset	74.27% BiLSTM on the Francsco dataset, 77.32% on the merged dataset using BiLSTM, and 84.82% BiLSTM on the Arabic news portals dataset
[26]	CNN, BiLSTM	Self-created Twitter dataset	CNN with a MARBERT embedding model of 95.64%.

models to find the most effective one for fake news detection. The results showed that combining MARBERT with a CNN led to the highest accuracy and achieved an F1-score of 0.9564. Dahou et al. [27] proposed a novel framework for the detection of Arabic fake news. They applied a modified feature selection algorithm to Twitter data. Another research work [28] focused on the aspect level of Arabic text using Seq2Seq normalization before classification.

Almandouh et al. [29] developed a news corpus containing 3185 fake news and 1453 real news and applied an ensemble model to detect fake news in Arab and Egyptian countries. Wotaifi and Dhannoon [22] proposed a deep learning-based hybrid model to improve the detection of Arabic fake news with 91.4% accuracy. Hawashin et al. [30] applied feature selection for Arabic fake news detection. These findings collectively emphasize the necessity for culturally informed strategies to effectively address the intricacies of misinformation in the Arab world. it is crucial to acknowledge the potential limitations inherent in current studies. The reliance on machine learning algorithms, as observed in many research works, introduces challenges related to bias, especially considering the diverse linguistic styles and expressions within the Arabic language. Additionally, the scarcity of labeled datasets for training models poses a constraint on the scalability and generalizability of detection systems. Recognizing these limitations is imperative for refining methodologies and fostering a more nuanced understanding of the complexities associated with fake news detection in the Arabic language. For a comprehensive overview, a summary of prior studies is presented in Table 1.

III. MATERIALS AND METHODOLOGY

This section describes the proposed methodology for fake news detection, and the dataset utilized for the experiments. The architecture of the proposed Arabic Fake News framework is shown in Figure 1.

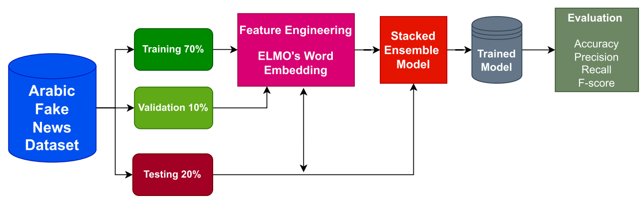

FIGURE 1. Architecture of the proposed Arabic fake news detection framework.

A. DATASET

Several low-resource datasets are available for the identification of false news. The Arabic Fake News Dataset (AFND)" on Kaggle is publicly accessible and is used in this study for performance evaluation of the proposed approach [31]. It is a benchmark dataset containing public Arabic news articles. It has 606912 news items gathered from 134 distinct Arabic news websites that are open to the public. The articles are divided into three categories by 'Misbar', a public Arabic news fact-checking platform: credible, not credible, and undecided. Out of 606912 news items, 207310 are credible, 167233 are not credible, and 232369 are undecided. The average number of words in each news ranges from 217-254.

While the 134 news websites are freely accessible to the public, their Uniform Resource Locator (URL) identities have been substituted with generic labels such as "source_1," "source_2," and so forth within JSON objects. These objects are compiled in a file named "sources.json" located in the main directory, ensuring the anonymization of the news sources. Each JSON object contains an anonymous identifier for the news source along with a corresponding label denoting its credibility status (credible, not credible, or undecided). The scraped Arabic articles from these public sources are organized into 134 individual sub-directories within the "Dataset" directory. Each sub-directory is named after the anonymous identifier (e.g., "source_1") assigned to the respective news source. Within each sub-directory, there exists a file named "scraped_articles.json," which contains an array of JSON objects storing information about the scraped articles. Each object includes details such as the article's title, text, and publication date. An example of such a JSON object, representing a scraped article from an anonymous news source.

B. WORD EMBEDDING

Word embedding is the process of encoding words as numerical vectors in a continuous vector space. Because textual input is intrinsically discrete, this enables machine learning models to comprehend and interpret it in a way that captures the semantic links between words. Conventional approaches such as bag-of-words (BoW) or term frequency-inverse document frequency (TF-IDF) handle individual words independently, disregarding their contextual connections. Conversely, word embeddings use a high-dimensional vector space, where words with similar meanings are close together to capture the context and meaning of individual words [32], [33]. Word embeddings help models grasp the subtleties of language, which is especially useful for sentiment analysis. This method of encoding words allows the model to learn not just the sentiment of individual words but also the feeling that is conveyed by the way those words interact in sentences. Word2Vec, global vectors for word representation (GloVe), and, more recently, contextualized embeddings like embedding from language models (ELMo) and bidirectional encoder representation from transformer (BERT) are popular word embedding models. We adopted the contextualized embedding because of the importance of the ELMo. The accuracy of sentiment analysis models has increased dramatically thanks to ELMo embeddings, which has made it possible for these algorithms to comprehend and analyze the complex nature of human sentiment as it is conveyed in text.

C. EMBEDDING FROM LANGUAGE MODELS

A ground-breaking method for word representation in natural language processing is called ELMo. ELMo takes into account the context in which a word appears, in contrast to conventional approaches which assign fixed vectors to words independent of context [34]. It does this by processing complete phrases using a bidirectional LSTM neural network, which considers both words that come before and after. ELMo is unique in that it uses numerous biLSTM layers. A multi-layered representation of each word is produced by combining the outputs of each layer, which each captures a distinct part of the context. Because of this, ELMo can capture complex contextual data, which makes it extremely useful for a variety of tasks like question answering, sentiment analysis, and machine translation. A word can have many representations based on its context according to ELMo's dynamic contextual embeddings, which provide a more sophisticated comprehension of language. ELMo is still an essential tool for natural language comprehension even though newer models like BERT have since gained popularity. It illustrates how context is crucial for processing and understanding human language.

TABLE 2. Details of machine learning models' hyperparameters.

Classifier	Hyperparameter
RF	n_estimators = 120, max_depth = 45, criterion='entropy',
SGD	epsilon = 0.2, Larning_rate='optimal',
ETC	n_estimators = 120, max_depth = 45, criterion='entropy',
KNN	leaf_size = 45, n_neighbors = 5,
GBM	n_estimators = 120, max_depth = 45, learning_rate = 0.2 ,
XGB	n_estimators = 120, max_depth = 45, learning_rate = 0.2 ,

D. MACHINE LEARNING MODELS FOR FAKE NEWS DETECTION

Results from text analysis with machine learning classifiers have been encouraging. Consequently, there are a great deal of algorithms and their variants in the literature. In the current study, bogus Arabic news is classified using ETC, KNN, RF, GBM SGD, and XGB. These algorithms are implemented using the scikit-learn library. A number of hyperparameters have been adjusted to improve the performance of these algorithms. Table 2 provides a list of the parameters and their respective values that yield the maximum accuracy. This subsection offers a succinct explanation of various algorithms.

1) RANDOM FOREST

RF is a supervised machine learning model and is based on DT. The ultimate forecast of a class is made by RF using multiple decision trees, each of which operates independently. The prediction is based on the class that gained the majority of votes. Because there is little association between the trees in Random Forest, the error rate is minimal [35]. Several strategies are available in random forests to determine the split in decision trees according to the classification or regression challenges. The Gini Index is used as the cost function for the classification problem. This determines the dataset's division. To calculate the Gini Index, deduct the total square probabilities for every class from one. The Gini Index can be achieved by:

$$G_{ind} = 1 - \sum_{i=1}^{c} (P_i)^2 \tag{1}$$

2) STOCHASTIC GRADIENT DECENT

A well-known optimization approach called gradient descent (GD) discovers the optimal values of model parameters at each iteration to reduce the cost function (c_f). SGD is a variation of GD that focuses on random probability, or stochastic so that a single sample is chosen for the model's training at each iteration. To obtain local minima, it takes a lot shorter training time to discover the critical function (c_f) of a single training sample (x_i) at each iteration [36]. For every x_i

and matching target class y_i, it updates the model's parameters to achieve this.

$$\theta_j = \theta_j - \alpha(Y^{i'} - Y^i)x_j^i \tag{2}$$

where α denotes the model's learning rate and θ_j represents its parameter. SGD uses several hyper-parameters to enable its operation on the data that is being analyzed.

3) EXTRA TREE CLASSIFIER

An arbitrary subset of characteristics is used in the ensemble of unpruned decision trees known as ETC to divide nodes. In contrast to RF, it does not bootstrap data; rather, it integrates all available data to form a decision tree [37]. The number of randomized input features chosen at each node (K), the minimum sample size needed to split a node (n_{min}), and the total number of decision trees in the ensemble are the two primary parameters that are involved (M). Because the split points are chosen at random, decision trees in ETC are less likely to be associated. In the event of regression, ETC generates a final forecast by averaging the predictions made by the decision trees in the ensemble. The majority voting of decision trees' predictions to produce a final prediction is relevant to this study since it deals with binary categorization. The adjustment of the ETC involves multiple hyper-parameters.

4) GRADIENT BOOSTING MACHINE

The classifier called the GBM builds a group of ineffective learners in an additive manner to raise the learning model's precision and effectiveness. Every weak learner in GBM aims to lower the error rate of the weak learner before them. By integrating the gradients and the loss function, it achieves this [38]. It effectively manages missing values in the data. GBM, which links to the statistical framework, is a formulation of boosting methods based on gradient descent. GBMs' learning mechanism fits new models one after the other to produce an estimate of the sample data that is more accurate. This method's fundamental idea is to create new base learners that have the highest possible correlation with the loss function's negative gradient, which is coupled to the entire ensemble. Because boosting techniques are so easy to use, you can experiment with various model designs. The updated prediction is multiplied by the learning rate to get the starting forecast.

5) K-NEAREST NEIGHBOR

A dependent variable is not required for the prediction of a result on particular data when using the KNN unsupervised learning algorithm. Enough data is supplied to this system to train it and allow it to determine which specific data points are relevant. It calculates the distance between the new data points and their closest neighbors. It also calculates the votes of its neighbors based on the value of K. If K is equal to 1, new data points are assigned to the class with the closest distance. Using the following formula, one may determine the separation between the two spots [39].

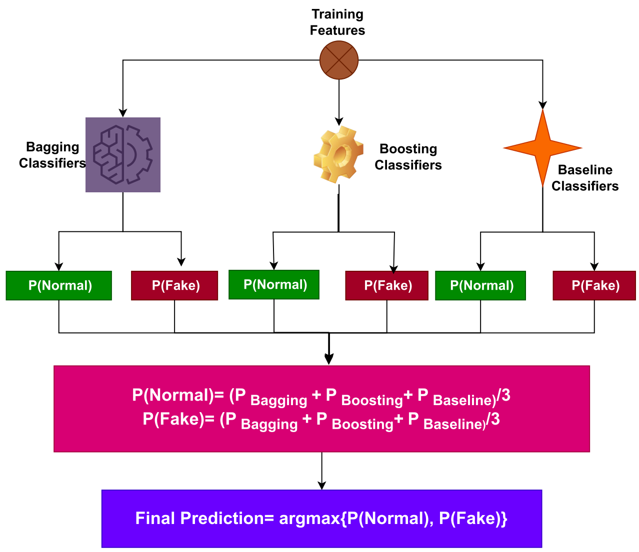

FIGURE 2. **Architecture of the proposed Arabic fake news ensemble learning.**

6) EXTREME GRADIENT BOOSTING

This work uses a high-speed supervised learning algorithm called XGBoost to classify Arabic fake news with high accuracy and precision. Regularized learning features, which help to adjust final weights and reduce overfitting risk, are one of XGBoost's main advantages [40]. In this case, the particular algorithm in use is as follows:

$$\Omega(\theta) = \sum_{t=1}^{n} d(y_t, \hat{y}_t) + \sum_{k=1}^{k} \beta(f_k) \tag{3}$$

The variables in the current context have the following definitions: For the loss function, use "d". The regularisation term is indicated by "b". The expected value is represented by "y_i". The number of instances in the training set is denoted by "n". The number of trees is "k".

E. PROPOSED FRAMEWORK

This section outlines the arrangement of the proposed methodology and its components as implemented within the experiment. Figure 2 presents a detailed depiction of the structure of the proposed framework. The proposed approach comprises two phases. In the first phase, all learning algorithms are implemented, and the proposed stacked ensemble model, a fusion of bagging, boosting, and baseline classifiers, undergoes training on the dataset. Transitioning to phase 2, the proposed model results are compared with state-of-the-art models.

The rationale behind adopting this stacking approach lies in harnessing the advantages of bagging, boosting, and baseline classifiers. The objective is to craft a more resilient and precise framework that facilitates the early detection of Arab fake news. In the text mining domain, prediction accuracy holds immense significance. The complete algorithm of the voting classifier is shown in Algorithm 1.

Algorithm 1 Ensemble Learning Classifier With Bagging, Boosting, and Baseline Models

```
Data: Training dataset D, base model M, number of
      base models N
Result: Final prediction for a given instance
Bagging Phase:;
for i = 1 to N do
    Sample Subset: Sample a random subset D_i from
    D with replacement;
    Train Base Model: Train a base model M_i on D_i;

Boosting Phase:;
Initialize weights for training instances: w_i = 1/N,
  where 1 ≤ i ≤ N;
for i = 1 to N do
    Train Base Model with Weighting: Train a base
    model M_i on D with instance weights w_i;
    Compute Error: Compute the classification error
    ϵ_i of M_i on D;
    Update Weights: Update instance weights based
    on ϵ_i: w_i = w_i · exp(α_i · ϵ_i);
    Normalize Weights: Normalize the weights so
    that they sum to 1;

Baseline Phase:;
Initialize weights for training instances: w_i = 1/N,
  where 1 ≤ i ≤ N;
for i = 1 to N do
    Train Base Model with Weighting: Train a base
    model M_i on D with instance weights w_i;
    Compute Error: Compute the classification error
    ϵ_i of M_i on D;
    Update Weights: Update instance weights based
    on ϵ_i: w_i = w_i · exp(α_i · ϵ_i);
    Normalize Weights: Normalize the weights so
    that they sum to 1;

Voting Phase:;
for i = 1 to N do
    Make Predictions: Use each base model M_i to
    predict the class of the instance;

Majority Voting: Combine predictions using
  majority voting (or weighted voting);
return Final prediction
```

F. EVALUATION PARAMETERS

This study utilizes multiple evaluation criteria, such as accuracy, F1 score, recall, and precision, to analyze the effectiveness of transfer learning models. Furthermore, the research makes use of confusion matrices to assess the performance of these algorithms. A confusion matrix, also known as an error matrix, is a tabular representation commonly used to illustrate the classifier's performance on test data, offering a visual representation of algorithm performance.

A "true positive (TP)" refers to instances in which the model made an accurate prediction for the positive class. In contrast "true negative (TN)" signifies cases where the model correctly predicted the negative class. Conversely, "false positive (FP)" corresponds to situations where the model made an incorrect prediction for the positive class when the actual class was negative. Likewise, "false negative (FN)" denotes instances where the model inaccurately predicted the negative class when the true class was positive.

The model's overall prediction accuracy is determined by evaluating the ratio of correct predictions to the entire dataset's total instances. This accuracy metric can be computed through the following

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

Precision serves as a metric that gauges the proportion of positive instances that were accurately predicted out of all the instances that the model identified as positive. Its central goal is to reduce false positives, providing insight into the model's capacity for correctly identifying positive cases. Precision is determined through the following

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

Recall, which is also referred to as the true positive rate or sensitivity, evaluates the proportion of positive instances that were correctly predicted about the total number of actual positive instances within the dataset. It quantifies the model's effectiveness in accurately capturing positive cases. Recall is computed using the following

$$Recall/Sensitivity/TPR = \frac{TP}{TP + FN} \tag{6}$$

The F1 score represents the harmonic average of precision and recall, offering a well-balanced assessment of the model's comprehensive performance by simultaneously accounting for both precision and recall. Its computation involves the following

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{7}$$

IV. RESULTS AND DISCUSSION

The results of various classifiers for Arabic fake news identification are presented in this section. The tests were conducted on an Intel 7th-generation Core i7 CPU running on Windows 10 on a Jupyter Notebook, with the machine learning models constructed using Python 3.8. The models' performance was evaluated through the use of the F1 score, accuracy, precision, and recall.

A. RESULTS OF MACHINE LEARNING MODELS FOR FAKE NEWS DETECTION

A thorough assessment of supervised machine learning classifiers is carried out using all of the features in the Arabic fake news dataset. Some classifiers performed below expectations, while others performed better than expected.

TABLE 3. Full feature set findings from machine learning models.

Model	Accuracy	Precision	Recall	F1 score
RF	87.86	88.55	89.76	88.42
SGD	87.14	87.92	88.07	87.85
ETC	89.09	88.72	88.72	88.72
GBM	89.34	89.78	89.73	89.39
KNN	80.14	79.63	80.56	80.18
XGB	89.28	89.67	90.20	89.93

This study covered a variety of statistical, tree-based, and regression-based false news detection methods. Table 3 shows the performance evaluation of each machine learning model utilizing all of the available features.

The accuracy numbers in the table below range from 80.14% to 89.34%. Though accuracy by itself is insufficient to emphasize a model's performance, more accuracy does indicate better overall performance of the model. To demonstrate the model's capacity for genuine positives and true negatives, precision and recall are crucial. Precision evaluates how well the models predict favorable outcomes. Between 79.63% and 79.78% are the precise numbers. Fewer false positives are indicative of higher precision. Recall, sometimes referred to as sensitivity or true positive rate, assesses how well the models can recognize all pertinent positive examples. The recall %ages are between 80.56% and 90.20%. A higher recall indicates fewer false negatives. A balanced indicator of a model's performance that takes into account both false positives and false negatives is the F1 score. The values of the F1 score vary from 80.18% to 89.93%. Recall and precision are better balanced when the F1 score is higher.

According to the data, the GBM classifier scored 86.89% accuracy, 86.52% recall, precision, and F1 score. The RF ensemble model achieved an F1 score of 86.22% and an accuracy of 85.66%. Except for the recall score, the GBM classifier surpassed the competition and achieved 87.14% accuracy, 87.53% recall, 87.58% for precision, and 87.19% F1 Score. With 78.36% recall, 77.94% accuracy, 77.43% precision, and 77.98% F1 score, the KNN is the lowest classifier for Arabic fake news detection.

B. RESULTS OF ENSEMBLE MACHINE LEARNING MODELS

Two baseline models SGD and KNN, two bagging classifiers, ETC and RF, and two boosting classifiers, GBM+XGB are employed in this work to detect fake news. Additionally, the group of these models for the categorization of fake news is studied. Table 4 displays the ensemble models' outcomes.

According to the ensemble model results, the boosting models' ensemble (GBM+XGB) fared better than other ensemble models, achieving 90.44% accuracy, 90.88% precision, 90.69% recall, and a 90.89% F1 score. With an accuracy of 88.14%, the baseline classifiers (SGD+KNN) come in second. Out of all the ensemble models, the bagging classifiers ensemble had the lowest accuracy, at 87.13%.

This work presented a stacking classifier for the Arabic false news categorization. A combination of bagging and boosting approaches make up the suggested system. The results obtained using the suggested method are displayed in Table 4.

TABLE 4. Results of the ensemble learning models.

Model	Accuracy	Precision	Recall	F1 score
Baseline Classifier(SGD+KNN)	88.14	87.87	87.76	87.81
Bagging Classifier (ETC+RF)	87.13	86.83	86.65	86.96
Boosting Classifier(GBM+XGB)	90.44	90.88	90.69	90.89
Proposed Stacking Classifier	**99.07**	**99.04**	**99.12**	**99.08**

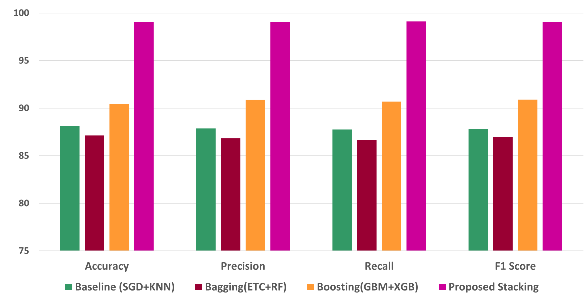

FIGURE 3. Result comparison of ensemble learning models.

Figure 3 illustrates the result comparison of the ensemble learning models. According to experimental results, the suggested stacking ensemble model surpassed previous learning models in terms of accuracy and other metrics, achieving an accuracy of 99.07%. Better results were also obtained by the suggested stacking classifier for the remaining assessment criteria, including 99.04% precision, 99.12% recall, and 99.08% F1 score. Figure 4 presents the ROC-AUC Curve of the proposed Arabic fake news ensemble learning model which exhibits exceptional performance, boasting an impressive AUC of 0.99. This curve vividly illustrates the model's ability to discriminate between true and false news with high accuracy across different threshold settings. The steep ascent and a near-perfect AUC signify the robustness and effectiveness of the ensemble learning approach in distinguishing between genuine and fake news in the Arabic language. The model's outstanding performance, as depicted by the ROC-AUC curve, underscores its reliability and efficacy in addressing the complex task of fake news detection.

C. RESULTS OF K-FOLD CROSS-VALIDATION

Using the k-fold cross-validation, the models are validated for their precision and robustness. Five-fold cross-validation is applied and results are shown in Table 5. Results suggest that the proposed model performs better than alternative models in terms of accuracy, precision, recall, and F1 score.

D. THEORETICAL FOUNDATIONS OF THE PROPOSED MODEL

Theoretical foundations of the proposed tri-ensemble learning models of bagging, boosting, and baseline classifiers can be traced back to several fundamental concepts in machine learning and statistical theory. Here are some key theoretical foundations for each of these ensemble learning techniques:

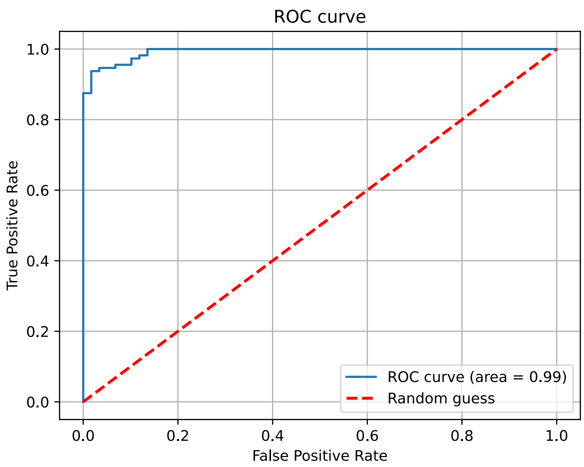

FIGURE 4. ROC-AUC curve of the proposed arabic fake news ensemble learning model.

TABLE 5. Five-fold-cross-validation results for the proposed approach.

Fold	Accuracy	Precision	Recall	F1 score
1st fold	98.95	99.38	96.27	98.36
2nd fold	98.95	99.50	98.58	98.47
3rd fold	98.16	99.83	99.50	98.33
4th fold	99.50	99.94	98.87	99.050
5th fold	99.59	99.58	97.79	98.64
Average	**99.31**	**99.79**	**97.65**	**98.56**

1) BAGGING (BOOTSTRAP AGGREGATING)

- Bootstrap Sampling: Bagging relies on the principle of bootstrap sampling, where multiple datasets are created by randomly sampling with replacement from the original dataset. This technique allows for the generation of diverse training sets.
- Variance Reduction: Bagging aims to reduce the variance of a model by averaging the predictions of multiple models trained on different bootstrap samples. According to the law of large numbers, averaging the predictions of multiple models tends to reduce the overall variance.
- Independence of Base Learners: Bagging works best when the base learners (individual models) are diverse and make independent errors. This diversity ensures that errors made by one model are compensated for by others, leading to improved generalization performance.

2) BOOSTING

- Sequential Learning: Unlike bagging, boosting operates by sequentially training a series of weak learners, where each subsequent learner focuses on correcting the errors made by the previous ones.
- Weighted Training Instances: Boosting assigns higher weights to misclassified instances, thereby directing subsequent learners to focus more on correcting these mistakes. This adaptive reweighting scheme ensures that the ensemble progressively learns to better handle difficult instances.

TABLE 6. Accuracy comparison with other feature engineering techniques.

Model	ELMo	GLOVE	BERT	FastText
Baseline Classifier(SGD+KNN)	88.14	80.32	86.57	84.49
Bagging Classifier (ETC+RF)	87.13	82.67	87.07	85.48
Boosting Classifier(GBM+XGB)	90.44	86.38	88.99	87.75
Proposed Stacking Classifier	**99.07**	**97.34**	**98.89**	**98.54**

- Exponential Loss Minimization: Boosting algorithms such as AdaBoost aim to minimize an exponential loss function, which emphasizes the importance of correcting the mistakes made by earlier models.

3) BASELINE CLASSIFIERS

- Simplicity and Interpretability: Baseline classifiers provide simple and interpretable models that serve as benchmarks for comparison with more complex ensemble methods like bagging and boosting.
- Naive Approaches: Baseline classifiers often include simple algorithms such as majority voting, random guessing, or assigning class labels based on the class distribution in the training data.
- Understanding Model Performance: Baseline classifiers help in understanding the baseline performance that can be achieved without employing sophisticated ensemble techniques. They provide insights into the minimum level of classification accuracy or error rate that can be expected.

Overall, these theoretical foundations underpin the rationale behind the design and operation of tri-ensemble learning models, which combine the strengths of bagging, boosting, and baseline classifiers to achieve improved predictive performance and robustness in various machine learning tasks.

E. PERFORMANCE COMPARISON WITH OTHER FEATURE ENGINEERING TECHNIQUES

To further investigate the significance of the proposed model and to show variety in experimentation, this research makes use of 3 other feature engineering techniques (GLOVE, BERT, and FastText) to check the accuracy of the proposed model. The results of the all ensemble learning model including the proposed model are shown in Table 6. The results from Table 6 show that the proposed model always gives the best accuracies no matter which features engineering technique is used with them. The best results are still obtained when ELMO word embedding is used with the proposed model. The second best-performing word embedding technique is BERT with an accuracy of 98.89%.

F. PERFORMANCE COMPARISON WITH EXISTING MODELS

Results are compared with existing models to demonstrate how well the proposed model performs in comparison to earlier state-of-the-art models. In order to achieve this goal, the three most relevant research studies using cutting-edge models intended to increase accuracy were chosen. For

TABLE 7. Performance comparison with other studies.

Ref.	Technique	Accuracy
[17]	QARiB transformer	95%
[18]	CNN	95%
[19]	LR using n-gram-level TF-IDF	93%
[21]	RF	78%
[22]	CNN-LSTM	91%
[23]	AFND-CNN-LSTM	70%
[24]	ARBERT	98%
[25]	BiLSTM	84%
[26]	CNN using MARBERT embedding model.	95%
Proposed	**Stacking Classifier (bagging, boosting, baseline)**	**99%**

example, the study [17] obtained a 95% accuracy for Arabic fake news identification using the QARiB transformer. The study [19] obtained the maximum accuracy of 93.3% by employing the LR using N-gram-level TF-IDF. In a similar vein, the CNN-LSTM ensemble deep learning model employed by [22] reported an accuracy score of 91.4%. The study [24] employed the word embedding technique ARBERT to classify fake news, achieving a 98.8% accuracy. The performance comparison of the suggested and current studies is displayed in Table 7. Comparison results demonstrate that the suggested model performs better than existing models.

G. LIME

Local interpretable model-agnostic explanations (LIME) have emerged as a powerful tool in the realm of Explainable AI, particularly for text classification tasks. LIME operates by generating perturbed instances of input data and observing the changes in the model's predictions. In the context of text classification, LIME provides local, human-interpretable explanations for individual predictions, allowing users to understand how a specific decision was reached. For instance, in natural language processing applications, LIME can highlight the key terms or phrases within a document that heavily influenced the classification outcome. This interpretability is crucial for building trust in AI models, especially in fields where transparency and accountability are paramount, such as healthcare or finance [41].

LIME's ability to shed light on the decision-making process of complex models enhances its utility in diverse applications and promotes the responsible deployment of AI systems. Its contributions to transparency and interpretability make it a valuable tool for researchers, practitioners, and stakeholders seeking to comprehend and validate the outcomes of text classification models [42].

Tweet translation in English

Hakimi is close to the Italian League title. Find out about the professional matches today, Saturday

from lime.lime_text import LimeTextExplainer
importing LIME library
explainer = LimeTextExplainer(class_names=
Target_Class)
getting explanations of target class prediction
idx = 0
index getting
exp = explainer.explain_instance(newsgroups_test.data
[idx], c.predict_proba, num_features=6)
print('Document id: print('Probability(med) =', c.
predict_proba([newsgroups_test.data[idx]])[0,1])
print('True class:
Output Document id: 0
Probability(Original) = 0.815
True class: Original

The classifier got this example right (it predicted original news). The explanation is presented below as a list of weighted features.

exp.as_list()
Output
[('Hakimi', 0.007521),
('is', 0.083925),
('close', 0.045346),
('to', 0.034576),
('the', 0.068439),
('Italia', 0.035984),
('League', −0.024727)
('title', 0.025674),
('Find', −0.015352),
('out', −0.002563),
('about', 0.059689),
('professional', −0.031854),
('matches', 0.036450),
('today', −0.034259),
('Saturday', −0.082589)]

V. CONCLUSION

This research presents a robust and effective solution for Arabic fake news detection through the integration of textual features and a sophisticated stacking classifier. By leveraging the collective strengths of bagging, boosting, and baseline classifiers, the proposed model demonstrated exceptional performance, achieving an impressive accuracy, precision, recall, and F1 score of 99%. The meticulous extraction of textual features, combined with the power of ensemble learning, enables the model to discern between authentic and fake news in the Arabic language with unparalleled accuracy. The significance of these findings is paramount in the current information landscape, where the proliferation of fake news threatens the integrity of shared knowledge. This study not only provides a highly accurate detection model but also contributes to advancing the field of fake news detection in the Arabic language. The achieved results underscore the efficacy of the proposed approach, offering a promising tool for maintaining information veracity and fostering a more trustworthy information ecosystem. The future work direction is to develop a large language dictionary by making a feature fusion of large word embedding dictionaries.

REFERENCES

[1] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," *J. Econ. Perspect.*, vol. 31, no. 2, pp. 211–236, May 2017.

[2] T. Rasool, W. H. Butt, A. Shaukat, and M. U. Akram, "Multi-label fake news detection using multi-layered supervised learning," in *Proc. 11th Int. Conf. Comput. Autom. Eng.*, Feb. 2019, pp. 73–77.

[3] K. Comeforo, "Review essay: Manufacturing consent: The political economy of the mass media," *Global Media Commun.*, vol. 6, no. 2, pp. 218–230, Aug. 2010.

[4] A. Zubiaga, M. Liakata, R. Procter, G. W. S. Hoi, and P. Tolmie, "Analysing how people orient to and spread rumours in social media by looking at conversational threads," *PLoS One*, vol. 11, no. 3, Mar. 2016, Art. no. e0150989.

[5] K. P. K. Kumar and G. Geethakumari, "Detecting misinformation in online social networks using cognitive psychology," *Hum.-Centric Comput. Inf. Sci.*, vol. 4, no. 1, p. 14, Dec. 2014.

[6] S. Helmstetter and H. Paulheim, "Weakly supervised learning for fake news detection on Twitter," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Aug. 2018, pp. 274–277.

[7] X. Zhang and A. A. Ghorbani, "An overview of online fake news: Characterization, detection, and discussion," *Inf. Process. Manage.*, vol. 57, no. 2, Mar. 2020, Art. no. 102025.

[8] K. Park and H. Rim, "Social media hoaxes, political ideology, and the role of issue confidence," *Telematics Informat.*, vol. 36, pp. 1–11, Mar. 2019.

[9] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD Explor. Newslett.*, vol. 19, no. 1, pp. 22–36, 2017.

[10] D. Radcliffe and H. Abuhmaid, "Social media in the middle east: 2019 in review," Jan. 2020. [Online]. Available: https://ssrn.com/abstract=3517916, doi: 10.2139/ssrn.3517916.

[11] Mideastmedia. (2017). *Social Media · Media Use in the Middle East, 2017*. [Online]. Available: https://www.mideastmedia.org/survey/2017/chapter/social-media/

[12] 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, Mar. 2021.

[13] Y. Liu and Y.-F.-B. Wu, "FNED: A deep network for fake news early detection on social media," *ACM Trans. Inf. Syst.*, vol. 38, no. 3, pp. 1–33, Jul. 2020.

[14] 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," 2021, *arXiv:2101.05626*.

[15] R. K. Kaliyar, A. Goswami, and P. Narang, "DeepFakE: Improving fake news detection using tensor decomposition-based deep neural network," *J. Supercomput.*, vol. 77, no. 2, pp. 1015–1037, Feb. 2021.

[16] M. Al-Sarem, A. Alsaeedi, F. Saeed, W. Boulila, and O. AmeerBakhsh, "A novel hybrid deep learning model for detecting COVID-19-related rumors on social media based on LSTM and concatenated parallel CNNs," *Appl. Sci.*, vol. 11, no. 17, p. 7940, Aug. 2021.

[17] 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, Apr. 2021.

[18] M. Alkhair, A. Hocini, and K. Smaïli, "Spotting fake news in Arabic with machine and deep learning techniques," *Int. J. Sci. Eng. Develop. Res.*, vol. 8, no. 2, pp. 605–611, 2023.

[19] A. R. Mahlous and A. Al-Laith, "Fake news detection in Arabic tweets during the COVID-19 pandemic," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 6, pp. 778–788, 2021.

[20] Z. Khanam, B. Alwasel, H. Sirafi, and M. Rashid, "Fake news detection using machine learning approaches," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 1099, no. 1, 2021, Art. no. 012040.

[21] H. Himdi, G. Weir, F. Assiri, and H. Al-Barhamtoshy, "Arabic fake news detection based on textual analysis," *Arabian J. Sci. Eng.*, vol. 47, no. 8, pp. 10453–10469, Aug. 2022.

[22] T. A. Wotaifi and B. N. Dhannoon, "An effective hybrid deep neural network for Arabic fake news detection," *Baghdad Sci. J.*, Jan. 2023.

[23] H. Najadat, M. Tawalbeh, and R. Awawdeh, "Fake news detection for Arabic headlines-articles news data using deep learning," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 4, p. 3951, Aug. 2022.

[24] A. B. Nassif, A. Elnagar, O. Elgendy, and Y. Afadar, "Arabic fake news detection based on deep contextualized embedding models," *Neural Comput. Appl.*, vol. 34, no. 18, pp. 16019–16032, Sep. 2022.

[25] K. M. Fouad, S. F. Sabbeh, and W. Medhat, "Arabic fake news detection using deep learning," *Comput., Mater. Continua*, vol. 71, no. 2, pp. 3647–3665, 2022.

[26] S. Alyoubi, M. Kalkatawi, and F. Abukhodair, "The detection of fake news in Arabic tweets using deep learning," *Appl. Sci.*, vol. 13, no. 14, p. 8209, Jul. 2023.

[27] A. Dahou, A. A. Ewees, F. A. Hashim, M. A. A. Al-qaness, D. A. Orabi, E. M. Soliman, E. M. Tag-eldin, A. O. Aseeri, and M. A. Elaziz, "Optimizing fake news detection for Arabic context: A multitask learning approach with transformers and an enhanced nutcracker optimization algorithm," *Knowl.-Based Syst.*, vol. 280, Nov. 2023, Art. no. 111023.

[28] M. E. Chennafi, H. Bedlaoui, A. Dahou, and M. A. A. Al-qaness, "Arabic aspect-based sentiment classification using seq2seq dialect normalization and transformers," *Knowledge*, vol. 2, no. 3, pp. 388–401, Aug. 2022.

[29] M. E. Almandouh, M. F. Alrahmawy, M. Eisa, and A. Tolba, "Ensemble based low complexity Arabic fake news detection," *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, no. 2, pp. 1022–1031, 2023.

[30] B. Hawashin, A. Althunibat, T. Kanan, S. AlZu'bi, and Y. Sharrab, "Improving Arabic fake news detection using optimized feature selection," in *Proc. Int. Conf. Inf. Technol. (ICIT)*, Aug. 2023, pp. 690–694.

[31] M. Yaseen. *Arabic Fake News Dataset (AFND)*. Accessed: Nov. 2, 2023. [Online]. Available: https://www.kaggle.com/datasets/murtadhayaseen/arabic-fake-news-dataset-afnd

[32] M. Umer, S. Sadiq, H. Karamti, A. A. Eshmawi, M. Nappi, M. U. Sana, and I. Ashraf, "ETCNN: Extra tree and convolutional neural network-based ensemble model for COVID-19 tweets sentiment classification," *Pattern Recognit. Lett.*, vol. 164, pp. 224–231, Dec. 2022.

[33] M. Umer, S. Sadiq, M. M. S. Missen, Z. Hameed, Z. Aslam, M. A. Siddique, and M. Nappi, "Scientific papers citation analysis using textual features and SMOTE resampling techniques," *Pattern Recognit. Lett.*, vol. 150, pp. 250–257, Oct. 2021.

[34] M. S. I. Malik, T. Imran, and J. Mona Mamdouh, "How to detect propaganda from social media? Exploitation of semantic and fine-tuned language models," *PeerJ Comput. Sci.*, vol. 9, p. e1248, Feb. 2023.

[35] M. Manzoor, M. Umer, S. Sadiq, A. Ishaq, S. Ullah, H. A. Madni, and C. Bisogni, "RFCNN: Traffic accident severity prediction based on decision level fusion of machine and deep learning model," *IEEE Access*, vol. 9, pp. 128359–128371, 2021.

[36] L. Bottou, "Stochastic gradient descent tricks," in *Neural Networks: Tricks of the Trade*, 2nd ed. Berlin, Germany: Springer, 2012, pp. 421–436, doi: 10.1007/978-3-642-35289-8_25.

[37] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely randomized trees," *Mach. Learn.*, vol. 63, no. 1, pp. 3–42, Apr. 2006.

[38] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, Oct. 2001.

[39] J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy K-nearest neighbor algorithm," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, no. 4, pp. 580–585, Jul. 1985.

[40] A. Juna, M. Umer, S. Sadiq, H. Karamti, A. A. Eshmawi, A. Mohamed, and I. Ashraf, "Water quality prediction using KNN imputer and multilayer perceptron," *Water*, vol. 14, no. 17, p. 2592, Aug. 2022.

[41] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?' Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 1135–1144.

[42] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, and F. Herrera, "Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Inf. Fusion*, vol. 58, pp. 82–115, Jun. 2020.

TURKI ALJREES is currently with the Department of Computer Science and Engineering, University of Hafr Al-Batin, Saudi Arabia. His research interests include machine learning, artificial intelligence, pattern recognition, image processing, web search, and web mining.