Al-Nahrain Journal of Science
ANJS, Vol.26(3), September, 2023, pp. 60-66

AutoKeras for Fake News Identification in Arabic: Leveraging Deep Learning with an Extensive Dataset

Raed S. Matti*, Suhad A. Yousif

Department of Computer Science, College of Science, Al-Nahrain University, Baghdad, Iraq

Article's Information

Received: 25.06.2023
Accepted: 21.08.2023
Published: 15.09.2023

Keywords:
AutoKeras
AutoML
Deep learning
Fake news

Abstract

Social media and the World Wide Web have led to a worrying rise in spreading false information, which presents a significant worldwide issue. Identifying and preventing false information is crucial in promoting an informed and knowledgeable society. The identification of false information, specifically in the Arabic dialect, presents inherent difficulties due to its diverse characteristics and linguistic intricacies. This study implements AutoKeras, a deep learning-based machine learning framework. Using advanced optimization techniques, the neural network architecture search, hyperparameter adjustments, and model selection can all be automated in AutoKeras. Therefore, it is suitable for our fake news detection task. The methodology employs proficient deep learning algorithms and natural language processing methods to acquire distinct characteristics that enable accurate differentiation between genuine and fake news. The present study uses various sources, including news websites, social media platforms, and blogs, to construct the dataset. The AutoKeras-based approach is superior to multiple state-of-the-art approaches to detecting fabricated news in Arabic, as evidenced by the experimental results. The suggested method outperforms 93.2% accuracy in identifying fake news, demonstrating its superior efficacy. This demonstrates the great promise of the deep learning-based Auto model for detecting false information.

DOI: 10.22401/ANJS.26.3.09

*Corresponding author: raedmatti.wk@gmail.com

1. Introduction

The spread of false information, commonly called fake news, can significantly affect society, particularly within politics [1]. The identification of fabricated news has attracted considerable interest owing to its extensive proliferation on digital platforms and social media, which presents severe problems [2]. The identification of false information in English has been tackled through diverse methodologies, including natural language processing (NLP) [3], machine learning (ML) [2], [4], and deep learning (DL) [5–7]. However, investigating this problem in other languages, particularly Arabic, has been comparatively limited. Prior research on detecting fake news in Arabic has employed machine learning algorithms. However, it has faced certain obstacles, such as the requirement for annotated datasets and the necessity to customize models for different areas and languages [8–12]. Additional investigation is necessary to tackle these constraints and examine diverse platforms and domains [9]. However, even for humans, more is needed to depend exclusively on NLP for misleading information identification; further fact-checking is required to assure accuracy. Limitations in use may also be seen in models based on deep learning that depend only on natural language processing. AutoML and Deep Learning are two robust technologies that provide considerable benefits in the realm of artificial intelligence. AutoML offers an innovative method to improve the procedure of developing and instructing deep learning models. The model selection and parameter tuning process is automated, enhancing the user experience. The Auto Model feature of Deep Learning enables the creation of self-learning models. The utilization of this methodology results in the automated design of complex neural network architectures, thereby enhancing the efficacy and effectiveness of deep learning tasks. Several studies have utilized AutoML techniques such as Auto-Weka [13], Auto-Arima [14], Auto-RapidMiner [15], and many others.

This study provides a dual contribution. Initially, AutoKeras, a framework for automated machine learning based on deep learning, is implemented to augment the precision of Arabic fake news classification. Furthermore, the proposed approach is comprehensively evaluated using a significant Arabic fake news dataset. This study aims to enhance the effectiveness and dependability of detecting fake news in Arabic by utilizing AutoKeras and an extensive dataset. Most previous research focused on utilizing algorithms for deep learning and machine learning to differentiate between fake and authentic news. From a review of previous research, it has been determined that currently, there is a limited of studies on identifying misleading information in the Arabic language. Wotaifi and Dhannoon (2022) aimed to improve the detection of false news in Arabic by creating a more efficient learning model. To accomplish this, the authors increased the amount of data utilized in [8] from 1862 to 3000 tweets while retaining the same features.

The Term Frequency-Inverse Document Frequency (TF-IDF) method was used to convert phrases into features and identify the most important ones. Compared to other machine learning algorithms, the upgraded random forest approach was shown to have a higher accuracy overall. The results showed that the improved method detected false news with an accuracy of 89%. The limitation of this study involves the fairly small amount of the dataset as well as the restricted level of predictive accuracy [9]. Khouja (2020) generated a new publicly accessible corpus for investigating misinformation in Arabic using textual participation approaches. The corpus consisted of two distinct versions, one containing 4,547 claims that were either true or false and the other containing 3,786 groups of claims and corresponding evidence. The authors developed two machine learning models for claim verification and stance prediction tasks. The model that utilized pre-trained BERT representations achieved F1 scores of about 76.7% and 64.3% in stance prediction and claim verification, respectively. The model based on LSTM architecture attained a maximum accuracy of 70.6% [10]. Furthermore, Sabbeh and Baatwah (2018) proposed a machine learning-oriented methodology to evaluate the credibility of Arabic news disseminated on Twitter. The methodology integrated both subject-specific and focused on user variables to assess news articles' truthfulness. A corpus comprising 800 Arabic language news articles procured from Twitter was subjected to manual annotation to evaluate their credibility. The study evaluated the effectiveness of three distinct machine learning algorithms, namely Decision Tree (DT), Support Vector Machine (SVM), and Naive Bayes (NB). The results indicated that DT performed better than the other two algorithms, achieving an accuracy score of 90.1% [11]. Fouad et al. (2022) introduced a model framework that employs deep learning and machine learning algorithms to identify fabricated news in the Arabic language. Various deep-learning models were assessed, including CNN, LSTM, and BiLSTM. However, no single machine learning model exhibited superior performance across all datasets. The BiLSTM model attained the maximum accuracy of 77% among all the models and datasets. The study employed three distinct datasets, namely Dataset 1 (consisting of 1980 tweets) and Dataset 2 (comprising 2578 tweets), along with Dataset 3 (a combined dataset of the earlier two, containing 4561 tweets) [12].

2. Methodology

This section provides a comprehensive introduction to automated machine learning and introduces the proposed model of the present research, which is referred to as AutoKeras. This study employed the Arabic fake news dataset (AFND) to identify misleading articles written in Arabic. The dataset went through preprocessing steps, which included the elimination of emojis and non-alphanumeric characters, normalizing text, removing stop words, and using stemming. After data cleaning, the processed data was used to train the AutoKeras model. The performance of the model was assessed by evaluating its accuracy, precision, recall, and F1 score metrics. The methodology employed in this research is shown in Figure 1.

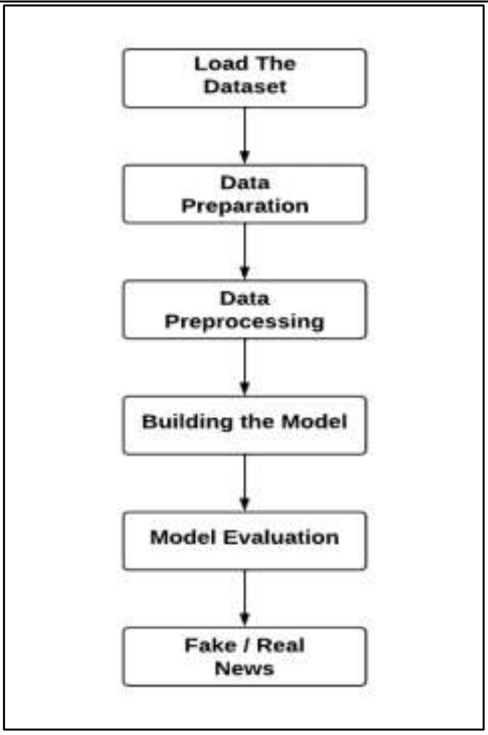

Figure 1. Workflow of Fake News Detection.

2.1. Automated Machine Learning (AutoML)

AutoML automates the whole procedure of applying ML to practical situations. Data preparation, feature extraction, and feature selection are just a few of the many ML application activities that must be performed appropriately to ensure the data is suitable for the intended ML activity. To get desirable results, practitioners should also try different learning models, fine-tune existing hyperparameters, and optimize the standard parameters. Due to the difficulty of these tasks for non-experts, AutoML is proposed as an AI-based solution to the growing issue of ML applications. When the whole ML application process is automated, the results are models that routinely beat those built by hand, more straightforward solutions, and quicker production. If successful, AutoML might make it much easier for non-experts to benefit from ML approaches [16]. These methods are a relatively new way of tackling this problem; they aim to find the optimal collaboration of the algorithms for classification and hyperparameter settings to maximize prediction effectiveness for an input dataset. Various AutoML frameworks are available today; researchers can choose from them [17].

2.2. AutoKeras

AutoKeras leverages the Keras and TensorFlow frameworks to construct models for machine learning. The KerasTuner framework is designed for hyperparameter tuning in Keras and offers the necessary infrastructure for implementing both the search space and the search algorithm. AutoKeras is constructed using KerasTuner as its foundation. It incorporates a set of meticulously crafted exploration areas, search algorithms tailored to specific tasks, and user-friendly application programming interfaces [18]. The fundamental process of AutoKeras currently encompasses the subsequent stages. Initially, AutoKeras analyzes the training data to ascertain various attributes, such as whether a structured data feature is categorical or numerical, whether the image data comprises a channel dimension, or whether or not the classification labels require encoding. Afterward, the previous information establishes an appropriate exploration area encompassing neural architecture designs and typical hyperparameters. Ultimately, the search algorithm successfully identifies hyperparameter values that exhibit superior performance [18]. AutoKeras employs a unique search algorithm that utilizes a previous understanding of the search space rather than approaching hyperparameter tuning as an instance of a black-box optimization issue. The primary concept involves initializing the search process with proficient configurations, where a configuration refers to a comprehensive collection of hyperparameter values utilized for constructing and training a model and persistently exploring surrounding areas of such proficient configurations [18]. Figure 2 shows the process of Autokeras task.

2.3. Dataset

AFND dataset [20] was utilized as the main data source in the present study and was provided in JSON (JavaScript Object Notation) format. The dataset was transformed into a tabular format to facilitate the research process. The AFND is a comprehensive compilation of Arabic language Fake News Datasets that is both extensive and annotated, sourced from publicly accessible Arabic websites. The dataset comprises 606,912 publicly available articles gathered from 134 accessible news websites across 19 Arab nations during a six-month period. To facilitate the experiment, a subset of the dataset was chosen and partitioned into two distinct categories: genuine news and fake news. The study's sample size comprised 83,298

Al-Nahrain Journal of Science
ANJS, Vol.26(3), September, 2023, pp. 60-66

records, of which 42,381 records were classified to be real news, along with 40,917 records that were classified as false news. Subsequently, the dataset was partitioned into two subsets, where 80% of the data was allocated for training purposes, and the remaining 20% was reserved for testing. The distribution of real and fake news in the dataset is shown in Figure 3.

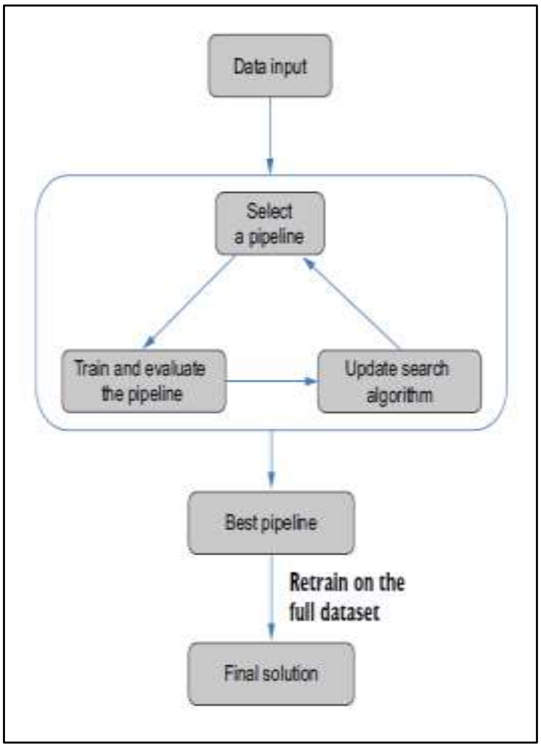

Figure 2. AutoML Process of the AutoKeras [19].

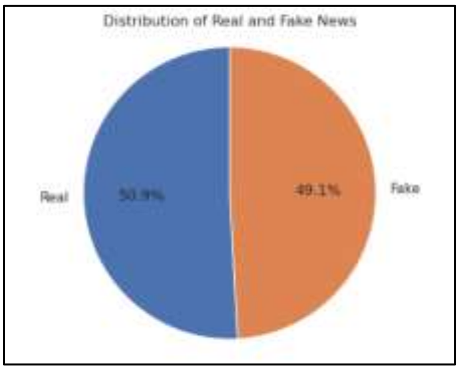

Figure 1. The AFND Dataset.

Figure 4 illustrates a word cloud that displays the most frequently used words in the dataset, divided into two categories: true news and fake news.

Figure 2. Word Cloud.

2.4. Preprocessing Phase

The present study involved executing multiple preprocessing procedures on the AFND Dataset to facilitate its analysis. The initial stage included mixing the headline and body of the news articles and converting numerical figures expressed in Arabic text to their English language equivalents. Subsequently, all emojis and special characters were eliminated from the text. Afterward, text normalization techniques were employed on the Arabic text, removing punctuation marks and standardizing letter form. Furthermore, Arabic stop words were eliminated from the text. These words are frequently utilized but do not substantially enhance the text's significance. In the end, the Tashaphyne: Arabic Light Stemmer [21] performed stemmization on the complete text. Stemming is a widely employed approach in natural language processing that reduces words to their basic form, resulting in a streamlined analysis process and enhanced accuracy. Table 1 summarizes the preprocessing steps.

Table 1. Text Preprocessing.

Preprocessing	Text
Raw text	"كَيفَ ضَيعتُكِ في زَحمةِ أيَّامي الطَّويلة - ١"
Removing Unwanted Characters	كَيفَ ضَيعتُكِ في زَحمةِ أيَّامي الطَّويلة
Text Normalization	كيف ضيعتك في زحمه ايامي الطويله
Removing Stop Words	ضيعتك زحمه ايامي الطويله
Stemming	ضيع زحم يوم طله

2.5. Proposed Model

Following the completion of data preparation and preprocessing, the subsequent stage in the suggested approach involves the development of the AutoKeras model. Initially, the parameters are established, encompassing the AutoModel using the input node and output node as the outputs. The 'overwrite' parameter has been configured with a Boolean value of True to enable the model to be retrained during each trial. The maximum number of trials has also been set to 5 through the 'max-trials' parameter. Furthermore, the quantity of epochs is established as 5, along with a validation partition of 0.2%, and the training split is 0.8%. Upon completion of the process, the graphic representation of the optimal model attained appears in Figure 5. The model includes multiple layers and operations. A variable-length text sequence is input to the layer. The data is reshaped by an expand-last-dim layer. Text data is vectorized to produce a fixed-length representation by the text-vectorization layer. Data is transformed through dense vectors of 128 dimensions by the embedding layer. Overfitting is prevented by dropout regularization. To extract text data features, the model uses separable convolutional layers and max-pooling layers. A dense layer with a sigmoid activation function is added as a classification head to make a binary prediction. AutoKeras can export and store the finished model. After training and improving the model, it can be converted and saved for accessible archive and retrieval.

2.6. Evaluation Metrics

The present study employed a confusion matrix [22] to assess the model's performance. Various evaluation metrics can be derived from the confusion matrix, including but not limited to F-measure, accuracy, precision, and recall. The equations (1-4) specify the formula for them. This depends upon the underlying principles:

- The model accurately predicts the negative or positive class, resulting in True Negative (TN) and True Positive (TP).
- The model's incorrect classification of negative and positive classes is called False Negative (FN) and False Positive (FP).

On the other hand, a Receiver Operating Characteristics (ROC) was utilized. The representation highlights the difference between the True Positive Rate (TPR), also known as recall, and the False Positive Rate (FPR) shown by the model [22].

$$\text{Accuracy} = \frac{TP + TN}{All\ Samples} \quad \dots (1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad \dots (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad \dots (3)$$

$$\text{F1} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad \dots (4)$$

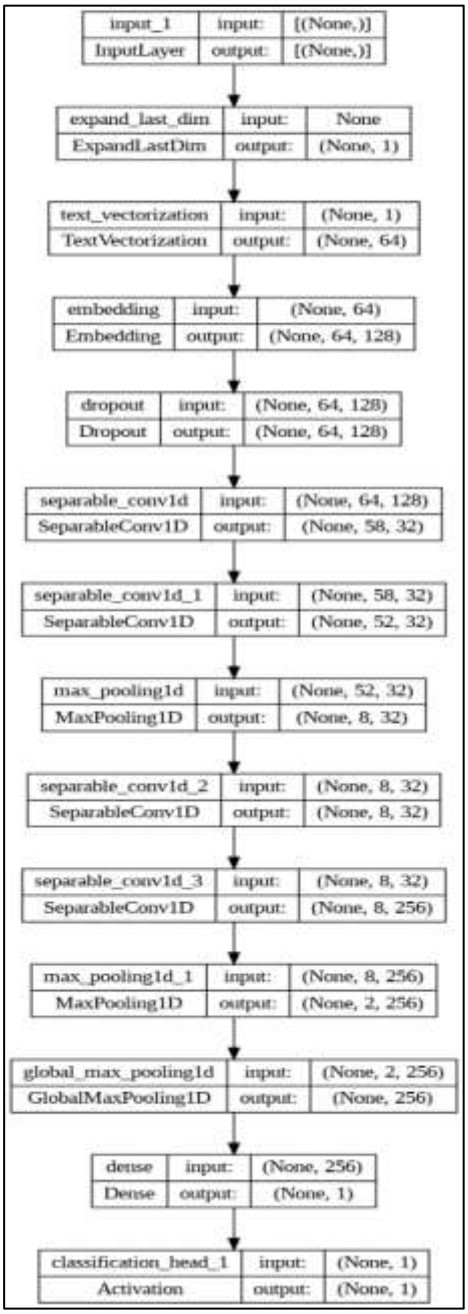

Figure 3. The Proposed Model.

3. Results and Discussion

This study aims to enhance the accuracy of identifying fabricated news. Data preparation and preprocessing techniques are employed to eliminate irrelevant data from the text before conducting news analysis. Subsequently, the Stemming methodology transforms the words into their respective base form. The subsequent phase involves utilizing the AutoKeras model and selecting the optimal number of trials and epochs. Ultimately, the construction of models begins, followed by their training utilizing the dataset. The research outcomes represent classifying news items into two distinct groups: true and fake. The confusion matrix (CM) demonstrates that the model accurately classified many articles into true or fabricated news categories. According to the true positives and true negatives, the model has accurately identified 7606 articles of fake news and 7916 articles of real news. The model exhibited false positives and false negatives, as evidenced by its misclassification of 485 genuine news articles as fake news and 653 fabricated news articles as authentic. The CM of the model is presented in Figure 6.

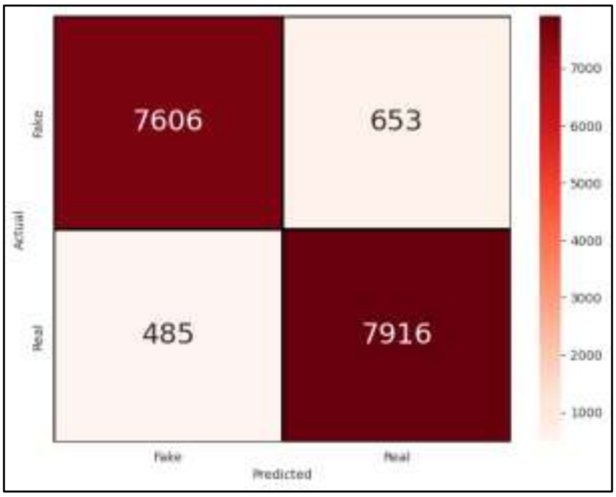

Figure 4. CM of the Proposed Model.

The outcomes of the model's classification on the dataset, after preprocessing and the implementation of the stemming technique, are presented in table 2. The ROC curve for the proposed model, which exhibits an accuracy of 0.932%, demonstrates its proficiency in detecting fabricated news. The findings indicate that the algorithms indicate considerable proficiency in discriminating between authentic and misleading news. The ROC curve of the model appears in Figure 7.

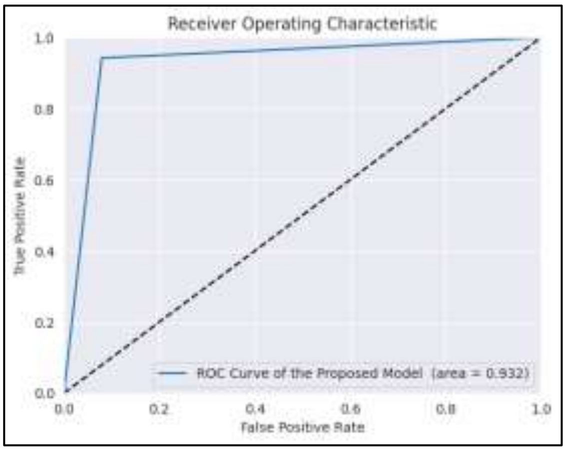

Figure 5. ROC Curve of the Proposed Model.

Table 2. Classification Results of the Model.

Metric	Proposed Model
Accuracy	93.2 %
Precision	93 %
Recall	93 %
F1-score	93 %

The effectiveness of the proposed model in predicting fake news articles is demonstrated by the classification outcomes presented in Table 2. The obtained metrics of accuracy, precision, recall, and F1-score were consistently high, with an accuracy rate of 93.2%, a precision rate of 93%, a recall rate of 93%, and an F1-score rate of 93%. The table mentioned shows that this study had surpassed prior research attempts to identify fabricated news in the Arabic language with high accuracy of 93.2%. When comparing current work with previous research, in [12], the most remarkable accuracy achieved using the BiLSTM model was 77%, but the model obtained from the current study produced a higher accuracy of 93.2%.

4. Conclusions

This study aimed to predict Arabic fake news articles using AutoKeras, an automatic machine learning framework and investigated the effectiveness of different deep learning architectures. The results demonstrated the efficacy of AutoKeras in the task of fake Arabic news prediction, showcasing its potential as a powerful tool in the field. In addition, the research found that AutoKeras models outperformed traditional approaches. The results obtained from this research study make a valuable addition to the existing research on the automated detection of fake news, underscoring the effectiveness of AutoKeras as a dependable framework for accomplishing this objective. The promising results encourage the adoption of AutoKeras and its deep learning architectures in combating the spread of fake news and promoting trust and accuracy in news dissemination. While this research primarily focused on fake news detection in Arabic language datasets, further investigation is necessary to fully explore the capabilities of AutoKeras models. Future studies should delve into the generalizability and adaptability of AutoKeras models across diverse languages and domains, expanding the scope of automatic credibility analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] Fatema, A.; Shaikhah, A.; Aisha, A.; Amal, E.; Monther, A.; "FALSE: Fake News Automatic and Lightweight Solution"; https://arxiv.org/abs/2208.07686v1, 2022.
[2] Reham, J.; Suhad, A.Y.; "Fake News Classification Using Random Forest and Decision Tree (J48)"; ANJS 23(4), 49-55, 2020.
[3] Zhixuan, Z.; Huankang, G.; Meghana, M.B.; Justin, H.; "Fake news detection via NLP is vulnerable to adversarial attacks". arXiv:1901.09657, 2019.
[4] Reham, J.; Suhad, A.Y.; "Classification of fake news using multi-layer perceptron". AIP Conf. Proc. 2334(1): 2021.
[5] Nishant, R.; Deepika, K.; Naman, K.; Chandan, R.; Ahad, A.; "Fake News Classification using transformer based enhanced LSTM and BERT". Int. J. Cogn. 3: 98-105, 2022.
[6] Pritika, B.; Preeti, S.; Raj, K.; "Fake news detection using bi-directional LSTM-recurrent neural network". Procedia Comput. Sci. 165: 74-82, 2019.
[7] Youssef, T.; Adelmoutalib, M.; Fouad, M.; "Automatic fake news detection based on deep learning, FasTtext and news title". IJACSA 13(1): 2022.
[8] Ghaith, J.; Hamed, A.; Momen, B.; Douglas, J.; "Classifying Arabic Tweets Based on Credibility Using Content and User Features". JEEIT: 596–601, 2019.
[9] Tahseen, A. W.; Ban, N. D.; "Improving Prediction of Arabic Fake News Using Fuzzy Logic and Modified Random Forest Model". Karbala Int. J. Mod. Sci. 8(3): 477-485, 2022.
[10] Jude, K.; "Stance prediction and claim verification: An Arabic perspective"; arXiv:2005.10410, 2020.
[11] Sahar, F. S. and Sumaia, Y. B.; "Arabic news credibility on Twitter: An enhanced model using hybrid features". J. Theor. Appl. Inf. Technol. 96(8): 2018.
[12] Khaled, M.F.; Sahar, F.S.; Walaa, M.; "Arabic Fake News Detection Using Deep Learning". Comput. Mater. Contin. 71(2): 2022.
[13] Venus, W.S.; Suhad, A. Y.; Nadia, M. G.; "Intrusion Detection System: An Automatic Machine Learning Algorithms Using Auto-WEKA". 2022 IEEE 13th Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, Malaysia, 42-46, 2022.
[14] Redha, A.A.; Suhad, A.Y.; "High performance time series models using auto autoregressive integrated moving average". Indones. J. Electr. Eng. Comput. 27: 422-430, 2022.
[15] Suhad, A.Y.; Venus, W.S.; Nadia, M.G.; "Automatic Machine Learning Classification Algorithms for Stability Detection of Smart Grid". BDAI: 34-39, 2022.
[16] Lin, Z.C.; "How Can Machine Learning and Optimization Help Each Other Better?". J. Oper. Res. Soc. 8(2): 341-351, 2020.
[17] Mustafa, A.; Rahimi, A.M.; "Automated machine learning for healthcare and clinical notes analysis"; Comp. 10(2): 24, 2021.
[18] Haifeng, J.; Francois, C.; Qingquan, S.; Xia, H.; "AutoKeras: An AutoML Library for Deep Learning"; J. Mach. Learn. Res. 24(6): 1-6, 2023.
[19] Qingquan, S.; Haifeng, J.; Xia, H.; "Automated Machine Learning in Action". MEAP, 2022.
[20] Ashwaq, K.; Moath, J.; Monther, A.; Manar, J.; "AFND: Arabic fake news dataset for the detection and classification of articles credibility". Data Br. 42, 2022.
[21] Taha, Z.; "Tashaphyne, Arabic light stemmer"; pypi.python.org/pypi/Tashaphyne/0.2, 2012.
[22] Muhammad, F.M.; Ashfia, J.K.; Abdul, M.H.; Muhammad, M.M.; Saifur, M.R; "A comprehensive review on fake news detection with deep learning". IEEE Access 9: 156151-156170, 2021.