# Arabic Poetry Authorship Attribution and Verification Using Transfer Learning

**Alaa M. El-Halees**

The Islamic University of Gaza, Gaza, Palestine
alhalees@iugaza.edu.ps

## Abstract

This paper employs a transfer learning approach to attribute and validate the authorship of Arabic poetry. Poetry authorship attribution is a technique of identifying a poet from an anonymous poem related to certain traits. Whereas, poetry verification entails determining whether a poem was written by a specific poet. There are a few works in Arabic poetry recognition, but their fundamental flaw is that they rely primarily on manual feature extraction. This is due to the fact that authors used traditional machine learning approaches. Transfer learning, as part of deep learning, has the ability to extract features from text automatically. Furthermore, using transfer learning to improve text classification accuracy in English has lately shown encouraging results. We employed Arabic-BERT as a transfer learning method for the Arabic language in our research. Arabic-BERT's four models were employed for authorship attribution and authorship verification. In terms of authorship attribution, we found that the best Arabic-BERT model is the large model, which has an f-score of 85%. It outperforms k-nearest neighbours, Naive Bayes, and Support Vector Machines, which are traditional machine learning approaches. In addition, we achieved a score of 96% in authorship verification, outperforming traditional machine learning techniques.

**Keywords:** *Arabic Poetry, Authorship Attribution, Authorship Verification, Transfer Learning, Machine Learning.*

## 1. Introduction

Authorship attribution or authorship identification is the process of recognizing the writer of a particular text from a collection of writers [8]. It is the study of determining an author's traits based on the features of documents authored by that author [14]. It is particularly interested in determining who the true author of a contested anonymous document is. Authorship attribution is regarded as a text categorization or text classification problem in the literature.

Poetry is the most popular form of Arabic literature. It is the main source for presenting Arab social, political, and intellectual life [16]. Although significant effort has gone into categorizing Arabic literature, categorizing Arabic poetry is not the same as categorizing other texts. That because Arabic poetry has structural features such as shape, meter, rhyme, and weight [4].

Two identification issues for Arabic poetry were investigated in this study. The first is authorship detection, which involves identifying the author of a poem from a list of poets. The second issue is the authorship verification problem, which entails providing some Arabic poetry and determining if it belongs to a single poet or not [15].

There are a few efforts in Arabic poetry recognition, where their primary flaw relies mostly on manual feature extraction. These days, manual feature extraction is obsolete. It is time consuming and the results have low performance. Deep learning models are now capable of extracting features automatically. Furthermore, because poems are short texts that rely heavily on implicit linguistic aspects, utilizing automated feature extraction for poetry is more successful. Poems are defined by their ability to make words signify meanings more than they do [18]. Furthermore, because the amount of text from a single author is usually limited, authorship concerns are frequently data constrained [25].

In our research, we used the BERT (Bidirectional Encoder Representations from Transformers) language model for transfer learning. BERT-based models have been shown to be highly effective at comprehending language when pre-trained on a large corpus [9]. For most NLP tasks, such models were able to set new benchmarks and reach state-of-the-art outcomes. In this research, Arabic-BERT, a pre-trained model for the Arabic language, has been utilized [21].

Finally, traditional machine learning methods such as k-Nearest Neighbours, Naive Bayes, and Support Vector Machine were used to compare our findings.

The rest of the paper is structured as follows: section two about the related works, section three addresses the research background, section four gives our research methodology, section five about experiments and results, while section six indicates the conclusion and future works.

## 2. Related Works

There is some research on detecting and verifying the authorship of Arabic poetry; however, it mostly relies on traditional machine learning methods. For instance:

Al-Falahi et al. in [1] posed the task of attribution of authorship to Arabic poetry. They employed Characters, Sentence length, Word length, Rhyme, and the first word in the sentence as features. For implementation, they employed a Markov Chain classifier with a large number of texts for validation. The results of their experiments indicate a precision of 96.96.%. The same authors in [2] reported an authorship attribution in Arabic poetry using Machine Learning. They used Naive Bayes and Support Vector Machine. The same features as in the [1] had been used. The findings of the experiment reveal a precision of 98.63%. In addition, from the same authors, in [3] they explored an authorship attribution in Arabic poetry using text mining categorization. The text mining categorization methods they used were Naive Bayes, Support Vector Machine, and Sequential Minimal Optimization. The greatest findings came from the experiment that has a classification precision of 98.96 %.

Also, for Arabic poetry authorship authentication, Omer and Oakes in [19] employed Arud, the metrical method used in traditional Arabic poetry. The study presented a method for automatically discriminating authors based on an Arud encoding. The Arud-based characteristics exceeded the baseline, which was the frequency of the most common terms, which are frequently used linguistic features.

In addition, some studies have looked into the authorship attribution of Arabic texts in general. For example, Altheneyan and El-Menai in [6] investigated the feasibility of Naïve Bayesian classifiers and their effect on event models for authorship attribution of Arabic literature. They investigated their application to this problem using several event models, including simple Naive Bayes (NB), Multi-Variant Poisson Naïve Bayes (MVPB),

Multivariant Bernoulli Naïve Bayes (MBNB), and Multinomial Naïve Bayes (MNB). They compared the performance of these models to other current techniques using an Arabic dataset collected from novels by ten different writers. The findings revealed that MBNB produces the top results, correctly identifying the author of a text with an accuracy of 97.43%. Furthermore, Hajja and Yahi in [13] investigated the topic of Arabic author attribution, which is concerned with identifying a specific author of an Arabic publication from a set of probable writers. For the training and testing of the models, many factors were considered. They examined how factors such as part of speech tags, stylistic concerns such as punctuation and sentence characteristics, word kinds, and word variety affected the results. In general, they found that the task's most informative characteristics were part of speech, n-grams, and stop words.

## 3. Background

In this study, transfer learning for author detection and verification of Arabic poetry was suggested. Transfer learning is a machine learning subfield that has been studied for over three decades [23]. It deals with the capability to use pre-existing models to learn new data. It is one method for obtaining knowledge that has proven to be effective. In transfer learning, a neural network is initially trained on a specific data set and task, then the features learned by the network are reused and transferred to another network to be trained on a different task. Recently, transfer learning has been employed in deep learning approaches to fulfil various tasks after being trained on large dataset [23].

One popular technique for transfer learning is BERT. BERT developed by Google, as a transformer-based machine learning algorithm for Natural Language Processing (NLP) pre-training stage [9]. It uses the Transformer and Attention mechanisms to learn contextual relationships between words (or word pieces) in a document. The transformer consists of two processes: an encoder that accepts text as input and a decoder that produces task predictions. Because BERT's goal is to develop a language model, just the encoder approach is necessary [9].

The first stage of BERT is pretraining. It is performed in an unsupervised manner and comprises of two primary tasks: Masked Language Modelling (MLM) and Next Sentence Prediction (NSP). In MLM task, a certain number of tokens in a sequence are replaced by [MASK] token. Then, the system attempts to detect the masked tokens. In the NSP task, two sentences A and B are simultaneously entered into BERT to predict whether sentence B occurs before or after sentence A in the same text. B is the genuine next phrase that follows 50% of the time during training, and it uses a random sentence from the text the other 50% of the time [9].

The second stage is fine tuning. Using BERT for a certain task in the fine tuning is relatively simple. BERT adds a simple layer to the main model and may be utilized for a large range of linguistic activities. In text categorization like authorship attribution, for example, a classification layer is applied on the Transformer output for the [CLS] token.

## 4. Methodology

In our work we used Experimental methodology. The following steps were engaged in developing a model for authorship attribution and verification:

### 4.1 Collecting Data

The dataset for our training and testing were taken from [5]. The information was scraped from adab.com. The dataset contains about 58K poetry (verses) for 652 writers from the 6th century to the current days. The majority of the poets have only a few lines of poetry. As a result, only poems for writers with the largest 10 texts were chosen. The names of the poets, the number of poems utilized, and the total number of words are listed in table 1.

**Table 1: The names of the poets, the number of poems, and the number of poems' words**

| No. | Poet Name | No. of poem text | No. of poem words |
|---|---|---|---|
| 1 | Ibn Al-Rumi | 2142 | 346,799 |
| 2 | Abu-al-'Ala' al-Mu'ri | 1593 | 103,067 |
| 3 | Ibn-Nabatah al-Masri | 1534 | 136,419 |
| 4 | Jubran Khalil Jubran | 1148 | 192,594 |
| 5 | 'Abd-al-Ghani al-Nabulsi | 1115 | 115,908 |
| 6 | al-Buhturi | 959 | 152,258 |
| 7 | Muhyi-al-Din Bin-'Arabi | 860 | 89,051 |
| 8 | Khalil Mutran | 857 | 152,741 |
| 9 | Abu-Nawwas | 821 | 54,352 |
| 10 | Safi-al-Din al-Huli | 811 | 80,960 |
| | **Total** | **11,840** | **1,231,555** |

### 4.2 Data Pre-processing

The collected dataset underwent various pre-processing stages before being fed as input to the model, including: Removing Punctuations, extra whitespaces, diacritics, and non-Arabic letters. Also, the data with orthographical variances have been standardized. Only for machine learning methods, Stop words, which are unimportant words that appear too frequently in the data collection, have been removed. Also, an algorithm for light stemming is used. It eliminates the most frequent suffixes and prefixes while maintaining the word's form.

For traditional machine learning methods, Term Frequency-Inverse Document Frequency (TF-IDF) is used to represent text in a numerical form. The TF-IDF calculates the importance of a word by considering its frequency of occurrence in the text and calculating how frequently the same word appears in other documents. If a term appears frequently in one document but not in others, it is likely to be extremely important to that text and is thus given greater weight [12].

### 4.3 Authorship Attribution Using Arabic BERT

Figure 1 depicts the high-level architecture for applying Arabic-BERT for the Authorship Attribution model. In the figure *[CLS]* donated the classification token $TOK_1 \ldots TOK_k$ are the Wordpieces tokens, $C$ is the class which is the class label. $A_1..A_n$ are the probabilities each Author.

The system receives an unknown Arabic poem as input. The text is tokenized by the system. Before feeding the model, the Arabic-BERT model, like the original BERT, requires a particular format for the input. At the start of each sentence, a special token called *[CLS]* is added, and at the end of each sentence, a special token called *[SEP]* is added. Wordpieces [24] was chosen as the tokenizer for Arabic tokenization since it was also utilized during BERT's pretraining.

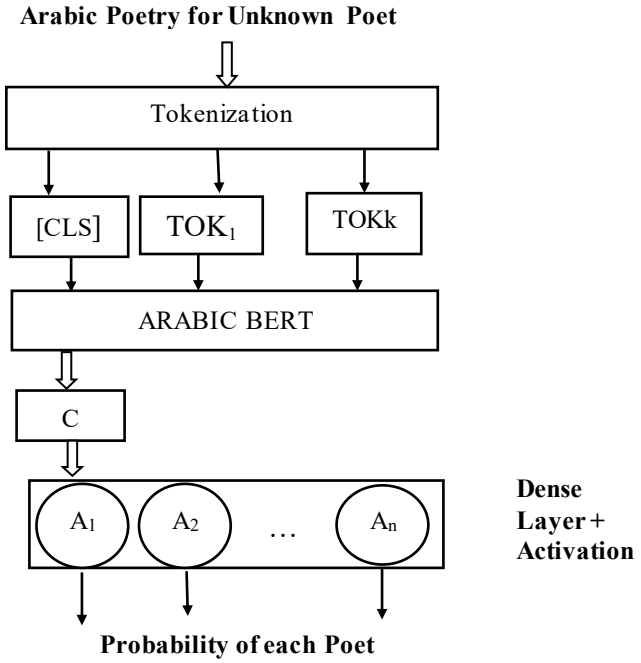


**Figure 1: The high-level architecture for the Authorship Attribution model using Arabic-BERT.**

The four Arabic-BERT language models that built from scratch for pretraining, were employed. Arabic-BERT is made up of models of various sizes that have been trained using masked language modelling and entire word masking [9]. The trained models of sizes Large, Base, Medium, and Mini, were used for experimentations [21]. The models' descriptions are shown in table 2. A vocabulary set of 32,000 Wordpieces was created using a corpus that included the unshuffled version of OSCAR data [22] and a recent data dump from Wikipedia, which totalled 8.2 billion words.

**Table 2: Descriptions of Arabic-BERT Models [21]**

| | Mini | Medium | Base | Large |
|---|---|---|---|---|
| Hidden Layers | 4 | 8 | 12 | 24 |
| Attention Heads | 4 | 8 | 12 | 16 |
| Hidden Size | 256 | 512 | 768 | 1024 |
| Parameters | 11M | 42M | 110M | 340M |

In the fine-tuning step, a dense layer is formed over the pre-trained Arabic-BERT model, as shown in Figure 1, and this dense layer is fed the final hidden vector of the classification token [CLS]. To accomplish multi-label classification of each poem, the sigmoid function is used as an activation function and binary cross-entropy is used as a loss function. The probabilities of all classes are produced by the model, with the greatest probability being the poet's name for the supplied Arabic Poetry.

### 4.4 Authorship Verification Using Arabic- BERT

Authorship verification is the process of finding if a poem was created by a certain poet. In this case, the system's input is an Arabic poem and the poet's name; the output will be true if the poem belongs to that author, or false if it doesn't. We applied the same model as in figure 1 for this sort of problem, with the exception that the output is binary.

In binary text classification, we aim to anticipate whether a piece of text or a phrase belongs in one of two categories, true or false in our model. For implantation, we utilized the soft max activation function with a sparse categorical cross entropy loss function.

### 4.5 Traditional Machine Learning

To compare the transfer learning technique to standard machine learning, we used the same dataset with three well-known machine learning algorithms for both authorship attribution and authorship identification. k-Nearest Neighbours, Naïve Bayes, and Support Vector Machine have been used for comparison.

k-Nearest Neighbours is a classification algorithm. Documents must be indexed and transformed to vector representations during the training phase. To categorize a new document *d*, its document vector must be compared to each document vector in the training set. The k-Nearest Neighbours are found by calculating similarity, which may be done using Euclidean distance [11].

Naive Bayes classifiers are commonly used in text classification because it is simple and computational efficient. It employs training technique that include approximating the relative occurrence of terms in a text as probabilities and using these probabilities to categorize the content. Naive Bayes decomposes the term $P(d \mid c)$, where *d* is the document and *c* is the class, by assuming conditionally independent features [11].

Support Vector Machine is a machine learning method presented by [7]. It is a hyperplane that divides a collection of positive instances from a collection of negative instances with greatest margin in its simplest linear form. The locations of test documents with relation to the hyperplanes are used to classify them [11].

### 4.6 Evaluation

There are many techniques for determining the effectiveness of our work; nevertheless, in our area, accuracy and recall are the most popular. Precision metric is the ratio of categorized poems that is successfully classified. Recall metric is the proportion of all poems properly categorized for a specific class. F-score is a combining metric that takes both precision and recall into account [20].

## 5. Experiments and Results

Using Arabic-BERT and machine learning techniques, we conducted many experiments for Arabic authorship attribution and Arabic authorship verification.

### 5.1 Arabic-BERT Models

Our dataset for Arabic poem attribution was subjected to the four models of Arabic-BERT: Mini, Medium, Base, and Large. We used the following parameters for Arabic-BERT fine tuning: Batch size 32, epoch 5, and 0.00001 Learning Rate.

Figure 2 summarizes the findings. It is clear that the larger the model size for transfer learning, the more accurate the findings. As a consequence, the f-score for mini is the lowest (73%) while the result for large is the greatest (85%).

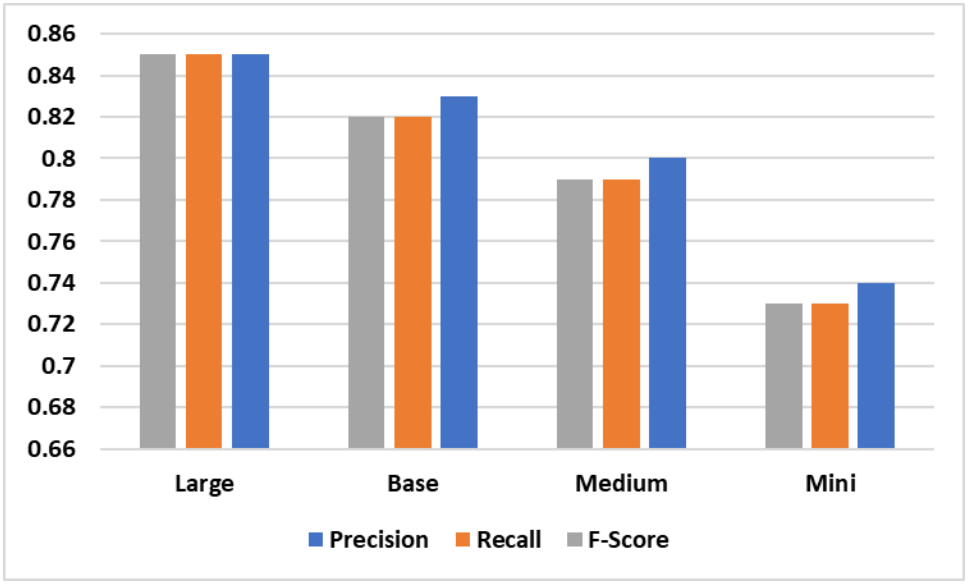


**Figure 2: Arabic Poem Attribution Dataset to the four Arabic-BERT models**

### 5.2 Poet Name Attribution

The method assigns the poet's name to an unknown poem. Table 3 presents the findings for all ten poet names using the Arabic-BERT large model, which produces the best results as demonstrated in figure 2. Except for al-Buhturi, the poem's accuracy is near the norm, which is about 85%. We found that poetry for al-Buhturi's included the following characteristics: "The different facets in which he presents his poems are religious, literary, and political" [17]. That suggests, he wrote in a variety of fields, which means we may require more training data in the future to detect his poems more accurately.

**Table 3: Performance of Poet Attribution**

| Poet Name | Precision | Recall | f-score |
|---|---|---|---|
| Ibn Al-Rumi | 0.8 | 0.82 | 0.81 |
| Abu-al-'Ala' al-Mu'ri | 0.91 | 0.89 | 0.90 |
| Ibn-Nabatah al-Masri | 0.88 | 0.89 | 0.88 |
| Jubran Khalil Jubran | 0.89 | 0.86 | 0.88 |
| 'Abd-al-Ghani al-Nabulsi | 0.91 | 0.86 | 0.88 |
| al-Buhturi | 0.69 | 0.73 | 0.71 |
| Muhyi-al-Din Bin-'Arabi | 0.83 | 0.86 | 0.85 |
| Khalil Mutran | 0.87 | 0.83 | 0.84 |
| Abu-Nawwas | 0.85 | 0.85 | 0.85 |
| Safi-al-Din al-Huli | 0.87 | 0.88 | 0.87 |
| **Average** | **0.85** | **0.85** | **0.85** |

### 5.3 Machine Learning Methods

We compared our transfer learning approach to traditional machine learning using the same dataset. We picked k-nearest Neighbours, Naive Bayes, and Support Vector Machine, all of which are well-known text mining algorithms. As illustrated in table 3, Arabic-BERT with the large model has a f-score of 85% that outperforms other techniques.

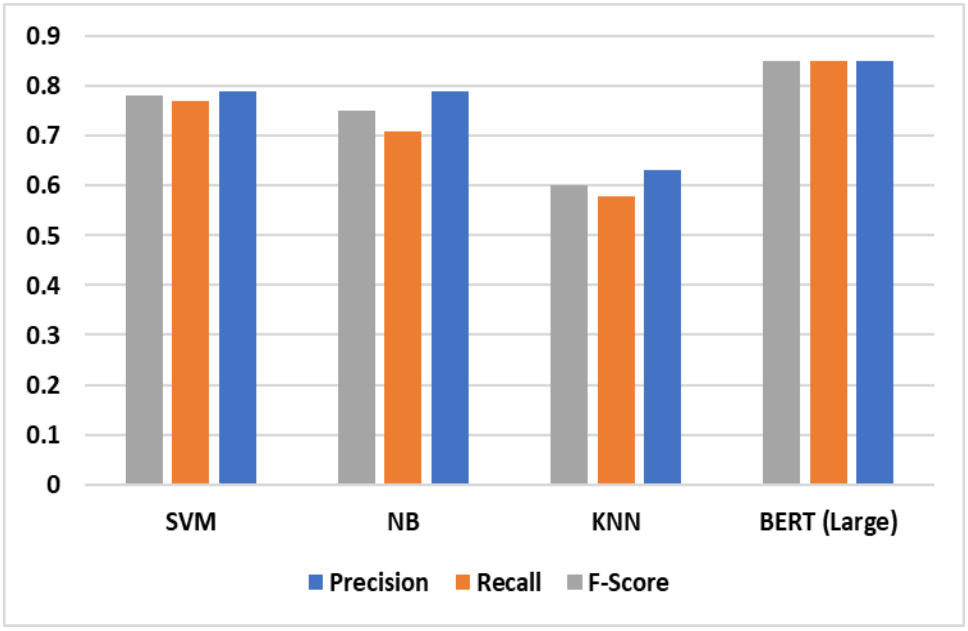


**Figure 3: Compare Arabic-BERT with machine learning methods**

Where KNN is K-Nearest neighbours , NB is Naïve Bays and SVM is Support Vector Machine.

### 5.4 Poet Name Verification

We utilized Arabic-BERT with a large model to verify if an unknown Arabic poem belonged to a specific author, and we achieved a very good f-score of 96%. As demonstrated in figure 4, Arabic-BERT outperforms KNN, NB, and SVM when compared to classical approaches.

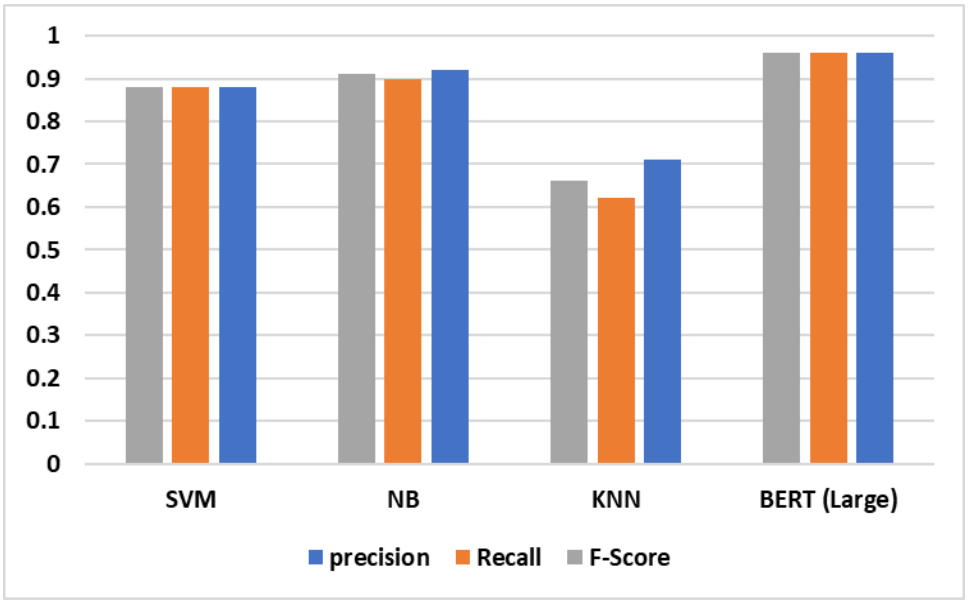


**Figure 4: Compare Poet Name Verification using Arabic-BERT with traditional machine learning**

## 6. Conclusion

Arabic poetry contains structural traits that distinguish it from other texts, making it a challenging problem to solve. Authors in this field usually extract poems features manually. However, nowadays, most work in text mining is based on transfer learning and automated feature extraction. We employed Arabic-BERT as a transfer learning approach to recruit and evaluate Arabic poets. We analysed four models for authorship attribution and found that the largest model, with an f-score of 85 %, is the best. Also, when we compared our findings to those of traditional machine learning with manual feature extraction, we discovered that the transfer learning technique outperformed the others. The other goal of this study is to determine whether an Arabic poem belongs to a specific poet. We acquired an f-score of 96% using transfer learning, which is significantly higher than other traditional machine learning approaches.

For greater performance in the future, we will need to identify or construct a larger language model than the one now in use. We also require more training data, particularly for the poetry, which has little data.

## References

[1]. A. Al-falahi, M. Ramdanim, M. Bellafkih and M. Al-Sarem, "Authorship attribution in Arabic poetry," In the10th International Conference on Intelligent Systems: Theories and Applications (SITA), 2015, pp. 1-6.

[2]. A. Al-falahi, M. Ramdanim and M. Bellafkih, "Machine Learning for Authorship Attribution in Arabic Poetry", International Journal of Future Computer and Communication 6(2):42-46. January 2017.

[3]. A. Al-falahi, M. Ramdanim and M. Bellafkih, "Authorship Attribution in Arabic Poetry Using NB,SVM,SMO". In 2016 11th International Conference on Intelligent Systems: Theories and Applications (SITA). Mohammedia, Morocco.

[4]. A. Ahmed, R. Mohamed and Mostafa, B. "Arabic Poetry Authorship Attribution using Machine Learning Techniques". Journal of Computer Science, 15(7), 1012-1021. 2019.

[5]. F. Alhazmi "Arabic Poetry Dataset (6th - 21st century)", https://www.kaggle.com/fahd09/arabic-poetry-dataset-478-2017. [accessed 10/8/2021]

[6]. A. Altheneyana, and M. Menaib "Naive Bayes classifiers for authorship attribution of Arabic texts", Journal of King Saud University - Computer and Information Sciences Volume 26, Issue 4, December 2014, Pages 473-484.

[7]. C. Cortes, and V. Vapnik " Support-Vector Networks", Machine learning, 20, 3, pp. 273-297, 1995. Kluwer Academic Publishers.

[8]. R.M. Coyotl-Morales, L. Villaseñor-Pineda, M. Montes-y-Gómez and P. Rosso, "Authorship Attribution Using Word Sequences." In: Martínez-Trinidad J.F., Carrasco Ochoa J.A., Kittler J. (eds) Progress in Pattern Recognition, Image Analysis and Applications. CIARP 2006. Lecture Notes in Computer Science, vol 4225. Springer, Berlin, Heidelberg.

[9]. J. Devlin, M. Chang, K. Lee and K. Toutanova " BERT: Pre-training of deep bidirectional transformers for language understanding". In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational

Linguistics: Human Language Technologies, Volume 1 pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

[10]. [11] A. El-Halees, A. "A Comparative Study on Arabic Text Classification". Egyptian Computer Science Journal. Vol. 30, no. 2 (2008).

[11]. V. Gudivada and C. R. Rao " Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications", , in Handbook of Statistics, Volume 38, Pages 2-515 (2018).

[12]. M. Hajja and A. Yahya" Authorship Attribution of Arabic Articles" 7th International Conference on Arabic Language Processing ICALP2019. Nancy, France, 16-17 October 2019.

[13]. G. Juola " Authorship attribution", Foundations and Trends in Information Retrieval 1(3):233-33. March 2008.

[14]. Koppel, M., Schler, J. "Authorship verification as a one-class classification problem", Machine Learning, Proceedings of the Twenty-first International Conference (ICML 2004), Banff, Alberta, Canada, July 4-8, 2004

[15]. Loya, A. (1974). "The Detribalization of Arabic Poetry". International Journal of Middle East Studies, 5(2), 202–215. http://www.jstor.org/stable/162590

[16]. Manna, H. "Al-Buhtari's life and poetry", Dar Al-Fikr Al-Arabi for Printing and Publishing, 2001.( In Arabic)

[17]. Mignot S. "Predicting Gender Poets with Deep Learning Methods. https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1184/reports/6846198.pdf. [accessed 10/8/2021].

[18]. Omer, A. and Oakes ,P. "Arud, the Metrical System of Arabic Poetry, as a Feature Set for Authorship Attribution," 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), 2017, pp. 431-436.

[19]. Powers, D. "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation" (PDF). Journal of Machine Learning Technologies. 2 (1): 37–63. 2011.

[20]. Safaya, A., Abdullatif, M., Yuret, D. " BERT- CNN for Offensive Speech Identification in Social Media", Proceedings of the Fourteenth Workshop on Semantic Evaluation", Barcelona (online)", International Committee for Computational Linguistics",2020. pages = "2054--2059",

[21]. Suárez, P. , Romary L., Sagot, B. "A Monolingual Approach to Contextualized Word Embeddings for Mid-Resource Languages", Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics",2020, Association for Computational Linguistics", pages = "1703--1714".

[22]. Weiss, K. , Khoshgoftaar, T. , Wang, D. " A survey of transfer learning", Journal of Big Data volume 3, Article number: 9, 2016.

[23]. Wu, Y. Schuster, M. , Chen, Z., Le,Q., Norouzi, M., Macherey, W. , Krikun, M. , Cao, Y., Gao,Q., Macherey K. "Google's neural machine translation system: Bridging the gap between human and machine translation" . CoRR abs/1609.08144 2016 .

[24]. Zhang Y., Boumber D., Hosseinia M., Yang F., Mukherjee A." Improving Authorship Verification using Linguistic Divergence", Published in ROMCIR 2021. Workshop held as part of ECIR 2021. March 28 - April 1, 2021.