

Article

A Transformer-Based Approach to Authorship Attribution in Classical Arabic Texts

Fetoun Mansour AlZahrani * and Maha Al-Yahya 

Information Technology Department, College of Computer and Information Sciences, King Saud University,
P.O. Box 145111, Riyadh 4545, Saudi Arabia; malyahya@ksu.edu.sa

* Correspondence: 442202914@student.ksu.edu.sa

Abstract: Authorship attribution (AA) is a field of natural language processing that aims to attribute text to its author. Although the literature includes several studies on Arabic AA in general, applying AA to classical Arabic texts has not gained similar attention. This study focuses on investigating recent Arabic pretrained transformer-based models in a rarely studied domain with limited research contributions: the domain of Islamic law. We adopt an experimental approach to investigate AA. Because no dataset has been designed specifically for this task, we design and build our own dataset using Islamic law digital resources. We conduct several experiments on fine-tuning four Arabic pretrained transformer-based models: AraBERT, AraELECTRA, ARBERT, and MARBERT. Results of the experiments indicate that for the task of attributing a given text to its author, ARBERT and AraELECTRA outperform the other models with an accuracy of 96%. We conclude that pretrained transformer models, specifically ARBERT and AraELECTRA, fine-tuned using the Islamic legal dataset, show significant results in applying AA to Islamic legal texts.

Keywords: authorship attribution; Islamic law; Arabic language; deep learning; Arabic transformer; AraBERT; AraELECTRA; ARBERT; MARBERT



Citation: AlZahrani, F.M.; Al-Yahya, M. A Transformer-Based Approach to Authorship Attribution in Classical Arabic Texts. *Appl. Sci.* **2023**, *13*, 7255. <https://doi.org/10.3390/app13127255>

Academic Editor: Douglas O'Shaughnessy

Received: 15 May 2023

Revised: 8 June 2023

Accepted: 15 June 2023

Published: 18 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the widespread availability of textual content online, such as research studies, literary writings, and user-generated text, and in light of the ease of reproducing and sharing these texts, plagiarism and literary theft have become much easier. To help in identifying authors of anonymous or unattributed text and in verifying authorship, literary scholars have been using an approach called “Authorship Attribution” (AA). This approach has roots from the 19th century [1], when Mendenhall [2] made the first recorded attempt to identify an author based on their writing style in 1887. Zipf [3] and Yule [4] further explored this in later years using basic statistical methods. However, with technological advancements, new computational methods have been investigated and used to address AA, starting from machine learning approaches and culminating in the use of deep learning and transformer-based approaches.

In the early 1960s, the approach of Mosteller and Wallace [5] was considered the foundation of computer-assisted stylometry. Stylometry includes many related fields, such as AA, authorship verification, authorship discrimination, plagiarism detection, and authorship profiling, described as follows [5,6]:

- AA is defined as a method of identifying the author of an unknown text. It can be used to detect an author, discriminating between writers’ styles, and resolving disputed work that might be attributed to different authors [7].
- Authorship verification involves verifying whether a given text is written by a specific author, and is considered to be a binary classification problem [8].
- Authorship discrimination is used to check whether different texts have the same author [9].

- Plagiarism detection is concerned with content similarity detection; through identifying cases of plagiarism in a given text, often via searching for a sentence or paragraph that has one or more than one writer [9].
- Authorship profiling involves extracting the features of the author's style to determine certain demographics, such as author gender, without identifying the author [5].

AA is well-known as a task in the field of natural language processing (NLP) [10]. According to Juola [11], AA is considered both one of the oldest and most persistent problems in the field of information retrieval; questions have been raised about authenticating documents as long as the documents themselves have existed [11]. AA has gained wide popularity in the literature, because it is key to solving many problems in the areas of authorship forensics, plagiarism detection, and the identification of anonymous authors. The main concept involves recognizing that every author has their own distinct writing style. This is referred to as author's stylometry or stylo-features [12]. Accordingly, AA approaches are concerned with selecting and extracting the best possible features that will provide the best performance of authorship identification using various classification methods [13].

According to [12], AA can be classified into three different categories: one-class attribution, binary-class attribution, and multi-class attribution. In one-class attribution, the purpose is to determine whether a specific document is written by the target author. In binary-class attribution, the purpose is to determine which of two authors wrote the given unattributed documents, based on evidence that the documents in question were written by one of two particular authors. In multi-class attribution, the purpose is to identify whether numerous documents have been written by more than two authors.

AA is a research area that aims to attribute a given text to its correct author. Not only literary scholars, but also politicians, historians, forensic experts, and religious scholars may have interests in the issue of authorship [14]. Many studies and research papers have examined the use of AA for the Arabic language [14–22]. However, most of these studies have focused on various Arabic collections such as articles, poems, Tweets, and other text genres from Arabic books. Although Islamic law text, which includes Islamic legal rulings [23,24] and Islamic doctrine affiliation [25], is considered to be among the Arabic literary genres, studies of AA related to this genre are limited.

Numerous approaches and algorithms have been utilized for the task of AA. An early method was the rule-based approach, in which a set of rules were identified to help the system learn how to categorize and classify given data. Later, machine learning (ML) models were introduced [26] to support the development of solutions for various NLP tasks; when used in classification problems, ML needs to train on the data using known features [27]. However, this technique requires experts to extract those features, which is considered to be costly and time-consuming. Surveying and reviewing the literature [15,16,28] has revealed that most studies have used the traditional machine learning classifiers, such as naïve Bayes, support vector machine (SVM), SMO-SVM, linear discriminant analysis (LDA), K-nearest-neighbors (KNN), logistic regression, and the Gaussian Bayes and decision trees classifiers. More recent ML approaches have shifted to using deep learning methods. Deep learning is a subfield of machine learning that works to stimulate the human brain for analytics learning [27]. There are many deep learning models used with NLP systems such as convolutional neural network (CNN) [29], recurrent neural network (RNN) [30], and sequence-to-sequence (Seq2Seq) models [31]. Some AA studies, such as the work by Apoorva and Sangeetha [32], have considered using DNN models in their experiments. Another study, conducted by Modupe et al. [33], proposed a regularized deep neural network (RDNN) method that enhances the accuracy of the AA for online text snippets, such as posts, called the PAA system.

Although AA tasks have utilized deep learning models, the main challenge of using these models is that their training requires a large amount of labeled data, which are expensive to generate [26]. This problem has therefore led to a shift toward using transformer pretrained models to overcome the drawbacks of deep learning models [34].

Several pretrained transformers are available. The most popular is the bidirectional encoder representations from transformers (BERT) model, the first to be based on transformer decoder and encoder layers [26]. Another model, ELECTRA, is a pretrained model that trains two transformer models: a generator and a discriminator. Recently, the difference between the pretrained language representation models, such as BERT or ELECTRA, has either been based on replaced token detection (RTD) or masked language modeling (MLM) [35]. The BERT-based models are pretrained via MLM, in which the model is asked to find original tokens that have been hidden randomly in the input sequence [35]. ELECTRA methods are pretrained based on replaced token detection (RTD), which is considered to be more efficient than MLM [35]. Because BERT has achieved significant results in language understanding for English, the Arabic computing community has worked on Arabic variants of BERT (AraBERT [36], ARBERT [37], and MARBERT [37]) for the Arabic language. The Arabic computing community has also worked on Arabic variants of ELECTRA, i.e., AraELECTRA [35], which is an Arabic model that has been used for many NLP tasks.

Recent works on NLP tasks in general, and AA in specific, have shown that pretrained transformer methods can achieve satisfactory results. Although several studies in the literature have focused on Arabic AA in general, the utilization of AA for Islamic text has not gained similar attention. This study investigates the recent state-of-the-art (SOTA) Arabic transformers in a rarely studied domain that has seen limited research contributions: the domain of Islamic legal texts. As mentioned above, given the importance of AA for Islamic law texts, the available research is limited, compared to that on general Arabic language AA. This work aims to address this research gap by investigating the effectiveness of leveraging different Arabic pretrained transformer models for addressing multi-class AA in Islamic law texts. We aim to answer the following research question: to what extent can Arabic pretrained transformer models provide satisfactory performance for the task of Arabic authorship attribution in Islamic legal texts?

We adopt an experimental approach to our investigation. Because no dataset has been designed specifically for this task, we design and build our own dataset using Islamic legal text resources from Al-Maktaba Al-Shamela. We conduct several experiments using four Arabic transformer-based models: AraBERT, AraELECTRA, ARBERT, and MARBERT. We conclude that pretrained transformer models, specifically ARBERT and AraELECTRA models, have shown significant results for the addressed task in this work.

The contributions of this work are summarized as follows:

1. Survey the literature on approaches used for AA in Arabic language works, including Islamic studies.
2. Develop a new, specialized Islamic legal texts dataset.
3. Implement four different Arabic pretrained transformer models (SOTA approaches) for classical Arabic AA. We fine-tune the models and compare their performance.
4. Recommend the best values for the hyperparameters for classical Arabic AA by analyzing the obtained results.

The remainder of the paper is structured as follows. In Section 2, we review related work in the area of Arabic AA. Section 3 presents our methodology, provides details of the dataset collection process, and discusses our experimental setup and evaluation metrics. Section 4 discusses the results. Finally, Section 5 presents the conclusions of this study and suggestions for future work.

2. Related Work

Studies on Arabic AA have been conducted on various text genres and forms, such as news articles [15,38], poetry [28], Tweets [16–19], and classical Arabic texts [14,20–22].

For Arabic text articles, Omar and Ibrahim [38] evaluated the effectiveness of stemming using the stylometric approach for Arabic AA. Their approach uses hierarchical cluster analysis with Ward linkage using Euclidean distance measure. The authors used three different Arabic stemmers: Light 10, Khoga, and GOLD. They collected 2400 text articles

written by 97 authors from various newspapers. To assess the efficiency of stemming, the authors used cluster analysis methods on both stemmed and non-stemmed articles. Their results indicate that stemming was ineffective in Arabic stylometric authorship applications. Their study shows that Light 10 achieved 67% accuracy; however, Khoga achieved 64%, while the accuracy rates of the GOLD and non-stemmed datasets were 61% and 78%, respectively. Hajja et al. [15] conducted an additional study in which they analyzed modern standard Arabic articles (70 articles by seven authors). They investigated the impact of certain defined text features on the authors' styles. These features include part of speech (PoS) tags, sentence characteristics, punctuation marks usage, word diversity, and word types. In addition, this study investigated other factors, such as the number of authors, the number of articles written by an author, and the size of text chunks used. The experiment was conducted using traditional machine learning methods: SMO-SVM, naïve Bayes, and decision trees. SVM yielded the best results: 98.24% in macro precision, 98.10% in macro recall, and 98.17% in macro F-score.

For Arabic poetry texts, Ahmed et al. [28] built an Arabic poetry authorship attribution model (APAAM) for poet identification using traditional machine learning techniques such as naïve Bayes, SVM, and linear discriminant analysis (LDA). The authors have tested the impact of several features on the task of AA, including lexical, character, structural, poetry, syntactic, semantic, and specific word features. They tested the models on 21,929 Arabic poems by 114 poets. Their study concludes that the LDA technique has a significantly high accuracy rate of 98% compared to other machine learning approaches. However, the researchers recommended further investigation with additional focus on different algorithms.

For short texts such as Tweets, Rabab'ah et al. [16] applied two approaches in attributing Arabic Tweets to their true author: bag-of-words (BOWs) and stylometric features (SFs). The researchers used these to create sets of features of a given text via different traditional machine learning classifiers: SVM, naïve Bayes, and decision trees. They concluded that combining all feature sets provided the best results. The SVM outperformed other classifiers, achieving the highest accuracy, 68.67%, on the combined feature set. In another study, Altakrori et al. [17] adapted an event visualization tool with Arabic tweets. The researchers compared the profile-based approach based on n-grams as features with instance-based classification techniques using the random forests, SVM, naïve Bayes, and decision trees classification methods with stylometric features. This study investigated the effectiveness of the n-gram approach based on different syntactic levels: word, characters, and PoS. It also examined the impacts on the attribution process before and after removing diacritics from a given text. Their experiment was conducted based on varying numbers of authors, starting with two authors, then five, followed by ten, and finally twenty authors. Their results show that the accuracy decreased gradually as the number of authors increased. The results of the study also showed that the diacritics have an insignificant effect on the attribution method; moreover, character-level and word-level n-grams are more effective than part-of-speech tags for the task of AA.

Abuhammad et al. [18] conducted another study, which involved training a machine learning model on Arabic Tweets. The dataset used in this study was collected from the Tweets of 45 authors. The aim of this study was to investigate AA for generating authors' writing style profiles. The study showed an accuracy of 99.24% by utilizing TF-IDF vectorizer and the SVM model. Jambi et al. [19] conducted a further study, which aimed to predict the authorship of different Arabic Tweets. The researchers collected 150 Tweets for every 1000 users for a total of 1.5 million Tweets. The authors used three classifiers: SVM, KNN, and random forests. Their results showed that the SVM and random forests classifiers performed better with regard to the accuracy measure compared to the KNN classifier.

Regarding studies of classical Arabic texts using the Universal Library (Elwaraq), three studies [14,20,21] were conducted to test AA within this library. The first of these [20] is a comparative survey conducted by Ouamour and Sayoud [20] on old Arabic texts. The dataset was collected from the Elwaraq library, extracted from 10 different books by

10 authors. The researchers created a dataset, which they named the AA of Ancient Arabic Texts (AAAT) set. The study used three different classifiers: multilayer perceptron (MLP), SVM, linear regression, and a proposed new approach called vote-based fusion (VBF). They found that the VBF approach produced a high AA accuracy of 90%, higher than the original classifier score utilizing only one feature.

The second study in [21] aimed to limit the noise words from a document to ensure a fair authorship identification process. The researchers constructed a dataset called Authorship Attribution for Ancient Arabic Philosophers (A4P). They investigated character 3 g and words as a feature extraction from the text. Their experiment employed five different classifiers: MLP, SVM, linear regression, Stamatatos distance, and Manhattan distance. This study found that classification performances is mainly based on the utilized features, the classification method, and the level of noise. Thus, the noise limit of 450 words per document appears to be the maximum for obtaining an accurate authorship identification. The third study, by Althenayan and Menai [14], conducted an experiment with several models of naïve Bayes classifiers, simple naïve Bayes, multinomial naïve Bayes, multivariate Bernoulli naïve Bayes, and multivariate Poisson naïve Bayes. The dataset was constructed from 30 books by 10 authors (Alfarabi, Alghaxali, Aljahedh, Almasaody, Almeqrezi, Altabary, Altowhedy, Ibnaljiawzy, Ibnrshd, and Ibnseña) from the Elwaraq website. The authors concluded that the multivariate Bernoulli naïve Bayes approach provided the best accuracy at 97%.

A distinctive study on classical Arabic texts, which uses another dataset, was conducted by Boukhaled [22]. Boukhaled investigated AA using machine learning methods for classical Arabic texts. The dataset used was derived from OpenITI, which contains 700 books by 20 authors. The experiments utilized three different machine learning classifiers, KNN, logistic regression, and Gaussian Bayes, in a comparative study using different types of style markers based on the different classifiers. This study relied on syntactical information for the experiment, such as function word features, PoS-based features, and character-based features. The findings indicate that these style markers can effectively impact the results of the AA task.

In reviewing the published works, we find numerous and considerable studies which address the task of AA for Arabic. However, for classical Arabic, specifically Islamic legal texts, studies are limited. In the following paragraphs, we highlight the studies on the specific Islamic texts subdomain of AA.

In 2018, Sayoud and Hadjadj [39] investigated AA for seven Arabic Islamic books, the Holy Quran, Hadith, Alghazali, Alquaradawi, Abdelkafy, Al-Qarni, and the Amr Khaled text collection. To improve the classification performance, they used fusion methods. The approach was applied in two different forms: the fusion of classifiers and the fusion of features. The fusion of classifiers used four different classifiers: Manhattan distance, MLP, SVM, and linear regression. The results showed that AA was satisfactory, with an accuracy of 96% to 99% based on classifiers without using the fusion approach. Nevertheless, the fusion approaches increased the accuracy to nearly 100%, especially for AA of the Quran and Hadith. Hence, this finding suggests the fusion strategy is strongly recommended for AA approaches requiring a high level of precision. In 2021, Hadjadj and Sayoud [13] proposed a new hybrid approach aiming to enhance the performance of AA for unbalanced data. The new approach was based on combining two algorithms: the principal components analysis and the synthetic minority oversampling technique. The researchers proposed three features for their experiments: function words, starting bigrams, and starting trigrams. The principal components analysis reduced the set of features' dimensionality, and the results were then used via the synthetic minority oversampling technique to construct the balanced dataset. In addition, Seven Arabic Books–dataset two (SAB-2) was created for Arabic AA; this dataset includes seven Arabic books written by seven Islamic scholars. They assessed the experiments using two different classifiers, SVM and naïve Bayes. The findings indicated that combining the principal components analysis and synthetic minority oversampling technique algorithms enhances the classification accuracy results by using

SVM classifiers. Their findings indicate that the proposed hybrid approach can effectively resolve the problem of unbalanced datasets.

Regarding a different type of dataset, Islamic legal rulings, a few studies in the literature address AA for such texts. A 2019 study by Al-Sarem et al. [23] described experiments on AA classifiers for short Arabic textual documents extracted from Dar Al-ifta Al Misriyyah. The study aimed to assess the performance of AA classifiers. The authors collected a dataset from Dar Al-ifta Al Misriyyah, consisting of 4631 legal rulings. They divided the dataset into four categories based on the number of words per text. To this end, the authors experimented with six classifiers: decision tree C4.5, naïve Bayes, KNN, the hidden Markov model, SMO, and the Burrows delta method. With the same objective, the researchers experimented via a combination of numerous features. Their findings indicate that combining the word-based lexical features with structural features achieved a high accuracy percentage of 86.39%. Another finding from this study indicates the superiority of the NB classifier over the other classifiers.

In another study, Al-Sarem et al. [24] showed how ensemble methods can support the AA task. The authors argued that this method combines multiple classifiers, and that this may lead to improved results. The research used the technique for order preferences by similarity to ideal solution (TOPSIS) to select the base classifier of the ensemble method. The TOPSIS model was used to choose the AA base classifiers by measuring five attributes, including average classifier accuracies, the ability to handle high-dimensional data, and sensitivity-to-noise data. The study indicated that the base classifier for ensemble methods is the SMO-SVM classifier. The authors suggested that new attributes be added using the TOPSIS model and the different ensemble methods utilized for Arabic language AA. In 2020, Al-Sarem et al. [40] applied another approach to assess the performance of a deep-learning-based artificial neural network for AA of Arabic text. While most of the previous works have used the traditional machine learning approach, this study by Al-Sarem et al. made a significant contribution to the field by using deep learning for AA in the Arabic language. The collected dataset was similar to the previous works, including 4686 rulings of 15 authors from Dar Al-ifta AL Misriyyah. Their experiment used an artificial neural network model with five-fold cross-validation to compare machine learning models, such as Bernoulli naïve Bayes, SVM, DT, and random forests. Their results indicate that using the artificial neural network outperformed other classifiers, using different metrics such as F-score, accuracy, precision, and recall. Nevertheless, these studies by Al-Sarem et al. are restricted to Islamic rulings only.

Focusing on ontologies, a study by El Bakly et al. [41] approached the task of AA with a model that uses Arabic ontology with semantic features to attribute a ruling to its jurisprudence doctrine (legal school). The authors introduced a new dataset, “ElWafaa LiFocahaa”, which is considered a valuable contribution. The results of evaluating the proposed model produced an accuracy of 90%. Another investigation was conducted in [25] for Islamic school AA. The purpose of the study was to examine how stylometric features can be used to predict the attribution of a given text to a legal Islamic school. The study used two approaches: unsupervised cluster analysis and supervised machine learning techniques. The dataset consisted of 135 books from the four Islamic schools, Hanbali, Hanafi, Shafi’I, and Maliki. The results indicate SVM achieved the best performance, with an accuracy of 97%. Nevertheless, both studies are constrained to Islamic school attribution and do not address the task of attributing an Arabic Islamic legal text to a specific author.

Juola et al. [42] demonstrated how research into AA was mainly concerned with authorship and time attribution of Arabic and Islamic texts. Their project examined some factors that may affect writing style, such as genre and time of composition. The data were collected from a new dataset called CLAUDia for works written between the 9th and 11th centuries. The texts they investigated included multiple genres, such as Arabic literature, Islamic text, and linguistics. The experiment focused on testing the effect of character n-grams, word n-grams, the most common words, rare words, and the length of the word as well. The results indicate that the established attribution graphs seem to contain one

central vertex that demonstrates the prototypical author. Regarding Islamic text in this study, it is limited to eight authors, and uses a standard authorship analysis tool (JGAAP).

Upon reviewing the literature, we find that a limited number of studies have incorporated AA for Islamic texts. Although some of these have achieved significant results, they have not utilized pretrained transformer models. In addition, the Islamic legal text dataset is novel and has not been used in any previous studies, particularly with transformers.

To summarize the review of the relevant work, AA is a well-established task in the natural language processing literature. However, compared to other languages, relatively few studies examine AA for Arabic, and fewer still investigate it for the genre of Islamic texts. Surveying the literature shows that 85% of Arabic AA research papers have used either statistical methods or traditional machine learning-based approaches [15,16,19,28]. The traditional machine learning classifiers such as SMO-SVM, linear regression, and MLP achieved a high level of accuracy, as in [6], which reached 100%. In contrast, the KNN machine learning classifier achieved a low level of accuracy of 35% in [19]. One study, [40], used a deep learning approach by using an artificial neural network for the task of Arabic AA. However, no existing studies have used transformer- or BERT-based models for Islamic legal text AA.

3. Methodology

We have adopted an empirical approach to design and conduct several experiments, analyze findings, and reach conclusions. Initially, we collected and prepared the dataset, which contains Islamic legal texts and their corresponding authors. We next developed a baseline transformer model with the default settings, enabling us to compare other models. Then, we conducted several experiments to determine whether we can improve upon the baseline. Using different optimizations, we developed four transformer-based models by fine-tuning various pretrained transformers. We then compared the performance of the models to our baseline using AA evaluation metrics: macro-F1 and accuracy.

3.1. Dataset

Since no standard or publicly available dataset existed for this task, we developed our own dataset, which contains the Islamic legal text with corresponding authors. The dataset is extracted from the books collection at Al-Maktaba Al-Shamela (<https://shamela.ws/>, accessed on 1 January 2022), an online library for various Arabic and Islamic collections. According to [43], the number of authors is a significant factor in AA studies; therefore, we prepared four datasets, with varying numbers of authors.

Regarding the size of the texts, the books are of different lengths; to prepare the dataset for fine-tuning the transformer models, the data needed to be segmented into smaller sizes [6,13]. Since the performance of the transformer-based models has been proven to be more efficient when they are pretrained on a large number of data items [44], we segmented the books further. This process has been applied in prior AA literature when large texts are involved, as in [17,19]. After splitting the text, we performed simple preprocessing tasks (discussed below). The experiments were conducted on four datasets of different sizes, with 200 texts per author. Dataset A is an eight-author set, dataset B is a 16-author set, dataset C is a 32-author set, and finally, dataset D is a 40-author set.

3.2. Preprocessing

Although preprocessing is an important step that can impact model performance, according to [45,46], there is no need for deep preprocessing, such as removing the frequent usage of stop words and performing the stemming of words, particularly for the task of AA, as these aspects are considered integral to the authors' writing styles. In this context, we have used a few basic regular expressions to remove all unused characters from the Arabic texts. We removed all numbers, English letters, newline delimiters, and special (non-alphanumeric) characters, and removed additional spaces between the words. Figure 1 provides some examples of the dataset.

C	B	A	
Author Name	Full name	Text	
الكوسج	المعروف بالكوسج	زوج المرأة على أن يطعمها سورة من القرآن وفي رواية أخرى أنه لا يجوز جعل تعليم القرآن صدقا ويبدأ القول قال إسحاق بن منصور بن يهرام، أبو يعقوب المروزي، المعروف بالكوسج	1
الكوسج	المعروف بالكوسج	رحمه الله وجزاه خيرا كتب الترجمة ذكرت أن إسحاق بن منصور المروزي قد سأل الإمام أحمد وإسحاق بن راهو إسحاق بن منصور بن يهرام، أبو يعقوب المروزي، المعروف بالكوسج	2
الكوسج	المعروف بالكوسج	ن الثوق يجري والفرعية تلغى للإسهام بيمين العقل في إظهار جزء صغير من آثار أولئك الأئمة من نفس المخط إسحاق بن منصور بن يهرام، أبو يعقوب المروزي، المعروف بالكوسج	3
الكوسج	المعروف بالكوسج	ورق على القول مكتوبة على القصب وله نوار إلى العمرة تتخلقه من أود فيها حب منور إلى البيضاء وليس صحيح إسحاق بن منصور بن يهرام، أبو يعقوب المروزي، المعروف بالكوسج	4
الكوسج	المعروف بالكوسج	ل السارق بلى يا رسول الله قال رسول الله صلى الله عليه وسلم لا خير فيكم من أودعكم الله هذا الحديث صحيح إسحاق بن منصور بن يهرام، أبو يعقوب المروزي، المعروف بالكوسج	5
الكوسج	المعروف بالكوسج	فأما وجهه ذلك أن الوارث إذا أقر بين علي مورثه قبل إقراره بغير خلاف دخله ويعلق ذلك بركة الميت كما لو إسحاق بن منصور بن يهرام، أبو يعقوب المروزي، المعروف بالكوسج	6
الكوسج	المعروف بالكوسج	لغة الماء وراجع الإخراج لأن رجب من هذا الأثر عن عمر رضي الله عنه أنه وضع على أرض السواد الغرا إسحاق بن منصور بن يهرام، أبو يعقوب المروزي، المعروف بالكوسج	7
الكوسج	المعروف بالكوسج	عليما خيرا آية من سورة النساء راجع المقنع بحاشيته كشف القناع المعنى الميزج الإصناف فإن رأى الفرقه فرق إسحاق بن منصور بن يهرام، أبو يعقوب المروزي، المعروف بالكوسج	8
الكوسج	المعروف بالكوسج	ومولى قوم من أنفسهم لهم عليهم الصلوة أيضا وقد قيل الولاء لعمدة كلمة النسب فكان هذا بيان لما أزيد من إسحاق بن منصور بن يهرام، أبو يعقوب المروزي، المعروف بالكوسج	9
الكوسج	المعروف بالكوسج	قال إسحاق لا يعجبني في الفدية والباقي كما قال قلت إذا قتل المحرم الصيد لم أكله قال كفارة واحدة قال إسحاق كما إسحاق بن منصور بن يهرام، أبو يعقوب المروزي، المعروف بالكوسج	10
الكوسج	المعروف بالكوسج	الجاري في باب صيد المراض وقال الحرفي في المختصر ولا يؤكل ما قتل بالندق أو الحجر لأنه موقود إسحاق بن منصور بن يهرام، أبو يعقوب المروزي، المعروف بالكوسج	11
الكوسج	المعروف بالكوسج		12

Figure 1. Examples of the dataset created to explore authorship attribution in Islamic texts.

3.3. Transformer Models

Pretraining transformer-based models such as BERT-based and ELECTRA-based models showed a significant result on many downstream tasks [47]. Therefore, we experimented with several base pretrained transformers: AraELECTRA, AraBERT, ARBERT, and MARBERT. AraELECTRA [35] uses a replaced token detection objective on a large Arabic text dataset. The performance results of AraELECTRA on several downstream NLP tasks show that it outperforms current SOTA Arabic language representation models.

AraBERT is an Arabic pretrained language model that achieved SOTA performance on Arabic NLP tasks when compared with multilingual BERT [36]. It was tested on different tasks, such as named entity recognition, sentiment analysis, and question-answering tasks. AraBERT has six model versions [48], which are AraBERTv0.1-base, AraBERTv0.2base, AraBERTv0.2-large, AraBERTv1-base, AraBERTv2-base, and AraBERTv2-large. AraBERT is publicly available for various Arabic NLP tasks.

BERT-based models ARBERT [37] and MARBERT [37] were introduced by Abdul-Mageed et al. [37]. ARBERT was trained using 61 GB (6.5 B tokens) of modern standard Arabic text that was collected from Wikipedia, crawled data, books, and news articles [49]. MARBERT [49] has been trained on a total of 128 GB of Tweets from various Arabic dialects, with at least three Arabic words each. The Tweets were kept in their original state with minimal preprocessing. The architecture of the AraBERT model for AA tasks is demonstrated below in Figure 2. Other used models in this study have similar architecture.

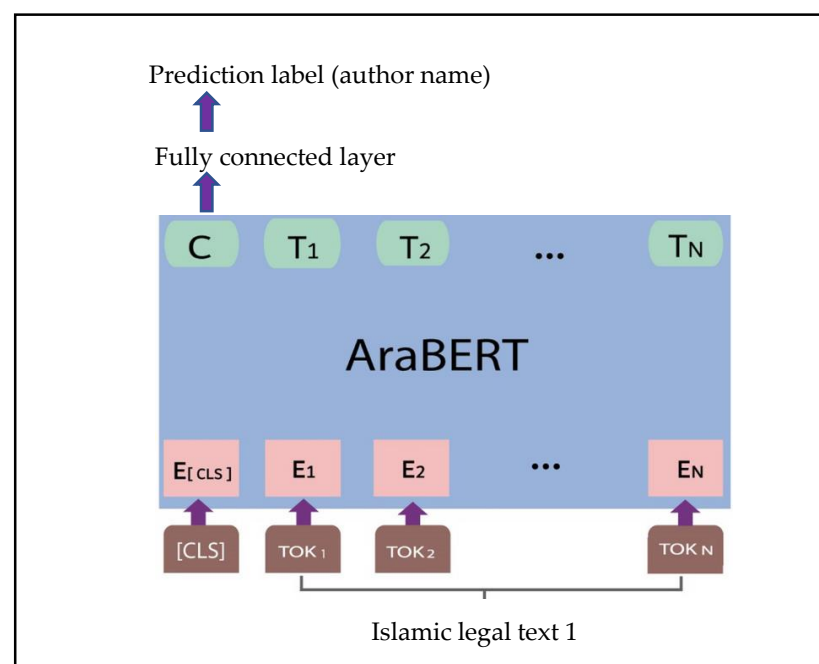


Figure 2. The AraBERT model architecture [50] for AA tasks.

3.4. Baseline Model

As indicated in the literature review, to the best of our knowledge, the current study is the first attempt at utilizing experiments in Arabic transformer-based models for the AA task. To further the research goal, we developed a baseline model to compare our results and improve the resulting AA model. We chose the transformer model AraELECTRA [35], because its results were satisfactory on various downstream Arabic NLP tasks, such as sentiment analysis, reading comprehension, and named-entity recognition. Hence, we decided to use AraELECTRA with the default hyperparameters (Table 1) as our baseline. Table 2 shows the performance of the baseline.

Table 1. Default AraELECTRA hyperparameters.

Hyperparameter	Value
Optimizer: adam_epsilon	1×10^{-8}
learning_rate	2×10^{-5}
num_train_epochs	2
per_device_train_batch_size	16
per_device_eval_batch_size	128
Max_len (input text length)	128

Table 2. Performance of AraELECTRA baseline.

	Dataset A		Dataset B		Dataset C		Dataset D	
	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
Baseline (AraELECTRA)	0.67	0.66	0.48	0.42	0.43	0.39	0.33	0.27

3.5. Experiments

We experimented with the transformer-based models by fine-tuning them using different parameter settings on the four datasets. We further fine-tuned the models using two main hyperparameters that have been shown to impact the model performance: the number of epochs [46] and the maximum length of tokens [51].

We performed the main experiments with three optimizations and an 80:20 split of the development set (80% training, 20% testing) [52,53]. In addition, we conducted a second set of experiments using five-fold cross-validation with an ensemble across all models (from the k-folds). The cross-validation splits the data randomly into K-fold, where K = 5 in our case, and then the model is trained on the K – 1 folds, while one-fold is left to test the model, as demonstrated in [54].

4. Results and Discussion

We revisit the research question established in the introduction of this article and answer it based on the results and observations from the experiments. Our research question was: To what extent can Arabic pretrained transformer models provide satisfactory performance for the task of Arabic authorship attribution in Islamic legal texts?

The conducted experiments utilized Google Colaboratory, with specifications illustrated in Table 3. Table 4 shows the results for the initial experiment, with default hyperparameters before any optimization. It is evident that ARBERT outperforms all other models on all metrics, achieving 60% in terms of accuracy and 58% in terms of F1-score. The nearest competitor was AraBERT, which achieved a 51% accuracy and a 49% F1-score. Moreover, all other models showed improvement on the baseline. The results also suggest that the performance is lower when the number of authors increases. This finding supports the results reported by Chadoulis [55].

Table 3. Hardware specifications for Google Colaboratory.

Parameters	Specifications
Hardware Accelerator	GPU
GPU Model Name	NVIDIA A100
GPU Memory	40 GB
CPU Model Name	Intel(R) Core (TM) i7-8550U
CPU Frequency	1.99 GHz
Available RAM	12 GB

Table 4. Performance results of initial experiment in authorship attribution for Islamic texts before optimization processes.

	Dataset A		Dataset B		Dataset C		Dataset D	
	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
MARBERT	0.73	0.72	0.58	0.56	0.56	0.54	0.48	0.44
AraBERT	0.75	0.74	0.65	0.63	0.55	0.53	0.51	0.49
ARBERT	0.86	0.85	0.72	0.71	0.70	0.69	0.60	0.58
Baseline	0.67	0.66	0.48	0.42	0.43	0.39	0.33	0.27

To answer this question, we present the results we obtained before optimization (Table 4) and after the three optimizations (Tables 5–7). For generalization, and because dataset D is the full dataset, we highlight in bold the best performance for this dataset. We started the experiments using the default hyperparameters in Table 1. Then, we increased the maximum length of the text from 128 to 512 in the first optimization. After that, we changed the number of epochs to five in the second optimization. Finally, we changed the number of epochs to 10 in the last optimization.

Table 5. Performance results for the experiment in authorship attribution for Islamic texts after first optimization (input text = 512).

	Dataset A		Dataset B		Dataset C		Dataset D	
	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
MARBERT	0.87	0.87	0.79	0.78	0.73	0.71	0.72	0.70
AraBERT	0.93	0.93	0.91	0.90	0.82	0.81	0.78	0.77
ARBERT	0.93	0.93	0.91	0.91	0.87	0.86	0.84	0.84
AraELECTRA	0.87	0.86	0.72	0.70	0.71	0.68	0.65	0.62
Baseline	0.67	0.66	0.48	0.42	0.43	0.39	0.33	0.27

Table 6. Performance results for the experiment in authorship attribution for Islamic texts with second optimization (no. of epochs = 5).

	Dataset A		Dataset B		Dataset C		Dataset D	
	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
MARBERT	0.95	0.95	0.94	0.94	0.94	0.94	0.93	0.93
AraBERT	0.97	0.97	0.98	0.97	0.93	0.93	0.93	0.93
ARBERT	0.96	0.96	0.96	0.96	0.94	0.94	0.94	0.94
AraELECTRA	0.97	0.97	0.96	0.96	0.93	0.93	0.90	0.90
Baseline	0.67	0.66	0.48	0.42	0.43	0.39	0.33	0.27

Table 7. Performance results for the experiment in authorship attribution for Islamic texts with third optimization (no. of epochs = 10).

	Dataset A		Dataset B		Dataset C		Dataset D	
	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score	Accuracy	F1-Score
MARBERT	0.98	0.98	0.97	0.97	0.95	0.95	0.95	0.95
AraBERT	0.97	0.97	0.98	0.98	0.96	0.96	0.96	0.96
ARBERT	0.97	0.97	0.97	0.97	0.95	0.95	0.96	0.96
AraELECTRA	0.99	0.99	0.99	0.99	0.96	0.97	0.97	0.97
Baseline	0.67	0.66	0.48	0.42	0.43	0.39	0.33	0.27

To improve the results, we performed three different optimizations on the dataset. Because this study is focused on the AA task that is relevant to each author's writing style, including such aspects as word selection, word placement, and punctuation marks, we considered the maximum input lengths that impact the authors' styles. Thus, we increased the maximum length of the text from 128 to 512 in the first optimization, as the maximum token length accepted by the transformer-based model is 512 [44,56]. Table 5 presents a comparison of performance results after this optimization. From the results, we can observe that performance for all models improved significantly. ARBERT still outperforms all other models by achieving 84% in terms of accuracy and 84% in terms of F1-score. However, AraBERT achieved 78% accuracy and 77% for F1-score, whereas its nearest competitor was MARBERT, which achieved 72% accuracy and 70% F1-score. AraELECTRA still presents a lower performance level in comparison with all other models. Moreover, Table 5 illustrates a decrease in performance when adding more authors, similar to the effect shown in the initial experiment (Table 4).

The results obtained so far are not comparable to the highest attained result for the AA task using transformer-based models in previously published studies on low-resource languages, such as the study performed for the Bengali language [57]. Thus, we further fine-tuned our models by increasing the number of epochs. According to [46], increasing the number of epochs can improve performance for the AA task. All the models' performance results improved after changing the maximum input tokens to 512 and the number of epochs from two to five. As Table 6 shows, ARBERT continued to outperform all other models by achieving 94% in terms of accuracy and 94% in terms of F1-score, whereas its nearest competitors MARBERT and AraBERT achieved 93% accuracy and 93% for the F1-score. In addition, AraELECTRA continued to show the lowest performance in comparison to all other models, achieving 90% accuracy and 90% for the F1-score. Our conclusion thus remained the same to that of previous experiments: performance decreases when more authors are added.

Although the results improved when increasing the number of epochs, as Table 6 shows, possibilities for improvement remained unchanged as compared to the literature, in which bnBERT [57] achieved 98% accuracy. Therefore, we further increased the number of epochs to 10, according to [46], to enhance the performance of the models. Table 7 shows the results of this final optimization; surprisingly, AraELECTRA achieved 97% in both accuracy and F1-score results, which are higher than previous experiments. In addition, MARBERT showed the lowest performance, with 95% in both accuracy and F1-scores. ARBERT and AraBERT obtained 96% in terms of accuracy and F1-scores, which is an increase compared to the previous experiment. To conclude, and based on our experiments, AraELECTRA (fine-tuned on 10 epochs with a maximum token length of 512) achieved the best performance for the AA of Islamic legal texts.

By comparing our final best results with those reported in the literature for low-resource languages, as Table 8 shows, our best model is found to be comparable to the best model reported. This answers our research question, in that Arabic pretrained transformer models can provide satisfactory performances for the task of Arabic authorship attribution in Islamic legal texts.

Table 8. Performance comparison with the existing literature using transformer-based approaches to authorship attribution for non-English texts.

Existing Literature	Base Model Name	Accuracy
Imran and Amin [57]	bnBERT (Bengali)	98%
Romanov et al. [58]	Transformer (Russian)	81%
Our Model	AraELECTRA (Arabic)	97%

The results in Tables 4–7 answer the first research question, and highlight the most important hyperparameters that affect the AA for Islamic legal texts. Based on this evidence, we conclude that the best hyperparameters are 10 epochs and a maximum input length of 512 tokens.

To further validate the performance and test for generalizability, we conducted an experiment using five-fold cross-validation with an ensemble of all models (from the K-folds) on dataset D, using the best optimization. Table 9 shows the results of this experiment. However, performance decreased by 1% for AraELECTRA and AraBERT. The five-fold cross-validation had no significant effect on the results for either MARBERT or ARBERT. Despite these lower results, AraELECTRA and ARBERT remain the best-performing models.

Table 9. Performance results for the experiment in authorship attribution for Islamic texts with 5-fold validation.

	Dataset D	
	Accuracy	F1-Score
MARBERT	0.95	0.95
AraBERT	0.95	0.95
ARBERT	0.96	0.96
AraELECTRA	0.96	0.96
Baseline	0.33	0.27

To investigate our results further, we performed a model testing on an unseen independent dataset. We randomly selected 100 texts from a sample of 13 authors from Al-Maktaba Al-Shamela, and evaluated the predictions using the AraELECTRA and ARBERT base models, which are the two that specifically outperformed the other models after the cross-validation process. AraELECTRA and ARBERT gained 96% for the task of AA as Table 9 shows. However, there was a slight performance decrease when predicting and labeling unseen data: both models misclassified seven texts, and classified 93 texts correctly.

5. Conclusions

This work addresses the task of Arabic AA for Islamic legal texts using Arabic transformer models. We conduct several experiments using SOTA Arabic BERT-based transformer models: AraELECTRA, AraBERT, ARBERT, and MARBERT. Since there were no available datasets for Islamic legal texts with corresponding authors, we developed our own dataset for this research. The results of this study show that ARBERT and AraELECTRA outperformed other models, reaching 96% in terms of both accuracy and F1-score. Finally, we can conclude that pretrained transformers, specifically ARBERT and AraELECTRA models fine-tuned using the Islamic legal text dataset, have shown significant results for the main task addressed in this research. A significant area of future work involves improving on the models' performance by using ensemble methods over the transformers. Another venue for future work involves investigating AA using scanned document images to recognize an author's written calligraphy script. Moreover, we can investigate with generative transformers such as GPT models.

Author Contributions: Conceptualization, F.M.A. and M.A.-Y.; methodology, F.M.A. and M.A.-Y.; software, F.M.A.; validation, F.M.A. and M.A.-Y.; formal analysis, F.M.A. and M.A.-Y.; investigation, F.M.A.; resources, F.M.A. and M.A.-Y.; data curation, F.M.A. and M.A.-Y.; writing—original draft preparation, F.M.A.; writing—review and editing, M.A.-Y.; visualization, F.M.A.; supervision, M.A.-Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research project was supported by a grant from the “Research Center of the Female Scientific and Medical Colleges”, Deanship of Scientific Research, King Saud University.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is available upon request.

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

References

1. Sarwar, R.; Nutanong, S. The Key Factors and Their Influence in Authorship Attribution. *Res. Comput. Sci.* **2016**, *110*, 139–150. [\[CrossRef\]](#)
2. Mendenhall, T.C. The Characteristic Curves of Composition. *Science* **1887**, *9*, 237–246. [\[CrossRef\]](#)
3. Zipf, G. *Selected Studies of the Principle of Relative Frequency in Language*; Harvard University Press: Cambridge, MA, USA, 1932. [\[CrossRef\]](#)
4. Yule, G.U. On Sentence-Length as a Statistical Characteristic of Style in Prose: With Application to Two Cases of Disputed Authorship. *Biometrika* **1939**, *30*, 363. [\[CrossRef\]](#)
5. Neal, T.; Sundararajan, K.; Fatima, A.; Yan, Y.; Xiang, Y.; Woodard, D. Surveying Stylometry Techniques and Applications. *ACM Comput. Surv.* **2018**, *50*, 1–36. [\[CrossRef\]](#)
6. Sayoud, H. Automatic authorship classification of two ancient books: Quran and Hadith. In Proceedings of the 2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA), Doha, Qatar, 10–14 November 2014; pp. 666–671. [\[CrossRef\]](#)
7. Bakly, A.H.E.; Darwish, N.R.; Hefny, H.A. A Survey on Authorship Attribution Issues of Arabic Text. *CiiT Int. J. Artif. Intell. Syst. Mach. Learn.* **2020**, *12*, 8.
8. Al-Sarem, M.; Cherif, W.; Wahab, A.A.; Emara, A.H.; Kissi, M. Combination of stylo-based features and frequency-based features for identifying the author of short Arabic text. In Proceedings of the 12th International Conference on Intelligent Systems: Theories and Applications, Rabat, Morocco, 24–25 October 2018; pp. 1–6. [\[CrossRef\]](#)
9. Swain, S.; Mishra, G.; Sindhu, C. Recent approaches on authorship attribution techniques—An overview. In Proceedings of the 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 20–22 April 2017; pp. 557–566. [\[CrossRef\]](#)
10. Custódio, J.E.; Paraboni, I. Stacked authorship attribution of digital texts. *Expert Syst. Appl.* **2021**, *176*, 114866. [\[CrossRef\]](#)
11. Juola, P. Authorship Attribution. *Found. Trends®Inf. Retr.* **2007**, *1*, 233–334. [\[CrossRef\]](#)
12. Zhao, Y.; Zobel, J.; Vines, P. Using Relative Entropy for Authorship Attribution. In *Information Retrieval Technology*; Ng, H.T., Leong, M.-K., Kan, M.-Y., Ji, D., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2006; Volume 4182, pp. 92–105. [\[CrossRef\]](#)
13. Hadjadj, H.; Sayoud, H. Arabic Authorship Attribution Using Synthetic Minority Over-Sampling Technique and Principal Components Analysis for Imbalanced Documents. *Int. J. Cogn. Informatics Nat. Intell.* **2021**, *15*, 1–17. [\[CrossRef\]](#)
14. Altheneyan, A.S.; Menai, M.E.B. Naïve Bayes classifiers for authorship attribution of Arabic texts. *J. King Saud Univ. Comput. Inf. Sci.* **2014**, *26*, 473–484. [\[CrossRef\]](#)
15. Hajja, M.; Yahya, A.; Yahya, A. Authorship Attribution of Arabic Articles. In *Arabic Language Processing: From Theory to Practice*; Smaïli, K., Ed.; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2009; Volume 1108, pp. 194–208. [\[CrossRef\]](#)
16. Rabab’ah, A.; Al-Ayyoub, M.; Jararweh, Y.; Aldwairi, M. Authorship attribution of Arabic Tweets. In Proceedings of the 2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA), Agadir, Morocco, 29 November–2 December 2016; pp. 1–6. [\[CrossRef\]](#)
17. Altakrori, M.H.; Iqbal, F.; Fung, B.C.M.; Ding, S.H.H.; Tubaishat, A. Arabic Authorship Attribution. *ACM Trans. Asian Low-Resource Lang. Inf. Process.* **2019**, *18*, 1–51. [\[CrossRef\]](#)
18. Abuhammad, Y. Authorship Attribution of Modern Standard Arabic Short Texts. In Proceedings of the 2021 Arab Women in Computing Conference (ArabWIC’21), Sharjah, United Arab Emirates, 25–26 August 2021; p. 7.
19. Jambi, K.M.; Khan, I.H.; Siddiqui, M.A.; Alhaj, S.O. Towards Authorship Attribution in Arabic Short-Microblog Text. *IEEE Access* **2021**, *9*, 128506–128520. [\[CrossRef\]](#)
20. Ouamour, S.; Sayoud, H. A Comparative Survey of Authorship Attribution on Short Arabic Texts. In *Speech and Computer*; Karpov, A., Jokisch, O., Potapova, R., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2018; Volume 11096, pp. 479–489. [\[CrossRef\]](#)

21. Bourib, S.; Sayoud, H. Author Identification on Noise Arabic Documents. In Proceedings of the 2018 5th International Conference on Control, Decision and Information Technologies (CoDIT), Thessaloniki, Greece, 10–13 April 2018; pp. 216–221. [\[CrossRef\]](#)
22. Boukhaled, M.-A. A Machine Learning based Study on Classical Arabic Authorship Identification. In *Proceedings of the 14th International Conference on Agents and Artificial Intelligence, Vienna, Austria, 3–5 February 2022*; SCITEPRESS–Science and Technology Publications: Setúbal, Portugal, 2022; pp. 489–495. [\[CrossRef\]](#)
23. Al-Sarem, M.; Emara, A.-H.; Wahab, A.A. Performance of authorship attribution classifiers with short texts: Application of religious Arabic fatwas. *Int. J. Data Min. Model. Manag.* **2020**, *12*, 350–364.
24. Al-Sarem, M.; Saeed, F.; Alsaedi, A.; Boulila, W.; Al-Hadhrami, T. Ensemble Methods for Instance-Based Arabic Language Authorship Attribution. *IEEE Access* **2020**, *8*, 17331–17345. [\[CrossRef\]](#)
25. Al-Yahya, M. Towards Automated Fiqh School Authorship Attribution. In Proceedings of the 19th International Conference on Computational Linguistics and Intelligent Text Processing CICLing 2018, Hanoi, Vietnam, 18–24 March 2018.
26. Kalyan, K.S.; Rajasekharan, A.; Sangeetha, S. AMMUS: A Survey of Transformer-based Pretrained Models in Natural Language Processing. *arXiv* **2021**, arXiv:2108.05542.
27. Xin, Y.; Kong, L.; Liu, Z.; Chen, Y.; Li, Y.; Zhu, H.; Gao, M.; Hou, H.; Wang, C. Machine Learning and Deep Learning Methods for Cybersecurity. *IEEE Access* **2018**, *6*, 35365–35381. [\[CrossRef\]](#)
28. Ahmed, A.-F.; Mohamed, R.; Mostafa, B. Arabic Poetry Authorship Attribution using Machine Learning Techniques. *J. Comput. Sci.* **2019**, *15*, 1012–1021. [\[CrossRef\]](#)
29. Kalchbrenner, N.; Grefenstette, E.; Blunsom, P. A Convolutional Neural Network for Modelling Sentences. *arXiv* **2014**, arXiv:1404.2188.
30. Liu, P.; Qiu, X.; Huang, X. Recurrent Neural Network for Text Classification with Multi-Task Learning. *arXiv* **2016**, arXiv:1605.05101.
31. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to Sequence Learning with Neural Networks. *arXiv* **2014**, arXiv:1409.3215.
32. Apoorva, K.A.; Sangeetha, S. Deep neural network and model-based clustering technique for forensic electronic mail author attribution. *SN Appl. Sci.* **2021**, *3*, 348. [\[CrossRef\]](#)
33. Modupe, A.; Celik, T.; Marivate, V.; Olugbara, O. Post-Authorship Attribution Using Regularized Deep Neural Network. *Appl. Sci.* **2022**, *12*, 7518. [\[CrossRef\]](#)
34. Shah, F.M.; Humaira, M.; Jim, A.R.K.; Ami, A.S.; Paul, S. Bornon: Bengali Image Captioning with Transformer-Based Deep Learning Approach. *SN Comput. Sci.* **2022**, *3*, 1–16. [\[CrossRef\]](#)
35. Antoun, W.; Baly, F.; Hajj, H. Araelectra: Pre-Training Text Discriminators for Arabic Language Understanding. *arXiv* **2020**, arXiv:2012.15516.
36. Antoun, W.; Baly, F.; Hajj, H. AraBERT: Transformer-based Model for Arabic Language Understanding. *arXiv* **2021**, arXiv:2003.00104.
37. Abdul-Mageed, M.; Elmadany, A.; Nagoudi, E.M.B. ARBERT & MARBERT: Deep Bidirectional Transformers for Arabic. *arXiv* **2021**, arXiv:2101.01785.
38. Omar, A.; Ibrahim, W. The Effectiveness of Stemming in the Stylometric Authorship Attribution in Arabic. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, 110114. [\[CrossRef\]](#)
39. Sayoud, H.; Hadjadj, H. Fusion Based Authorship Attribution-Application of Comparison Between the Quran and Hadith. In *Arabic Language Processing: From Theory to Practice*; Lachkar, A., Bouzoubaa, K., Mazroui, A., Hamdani, A., Lekhouaja, A., Eds.; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2018; Volume 782, pp. 191–200. [\[CrossRef\]](#)
40. Al-Sarem, M.; Alsaedi, A.; Saeed, F. A Deep Learning-based Artificial Neural Network Method for Instance-based Arabic Language Authorship Attribution. *Int. J. Adv. Soft Comput. Its Appl.* **2020**, *12*, 1–15.
41. El Bakly, A.H.; Darwish, N.R.; Hefny, H.A. Using Ontology for Revealing Authorship Attribution of Arabic Text. *Int. J. Eng. Adv. Technol.* **2020**, *9*, 143–151. [\[CrossRef\]](#)
42. Juola, P.; Milička, J.; Zemánek, P. Authorship and Time Attribution of Arabic Texts Using JGAAP. In *Intelligent Natural Language Processing: Trends and Applications*; Shaalan, K., Hassanien, A.E., Tolba, F., Eds.; Studies in Computational Intelligence; Springer: Cham, Switzerland, 2018; Volume 740, pp. 325–349. [\[CrossRef\]](#)
43. Huertas-Tato, J.; Huertas-Garcia, A.; Martin, A.; Camacho, D. PART: Pre-trained Authorship Representation Transformer. *arXiv* **2022**, arXiv:2209.15373v1.
44. Devlin, J.; Chang, M.-W.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv* **2019**, arXiv:1810.04805.
45. Rocha, A.; Scheirer, W.J.; Forstall, C.W.; Cavalcante, T.; Theophilo, A.; Shen, B.; Carvalho, A.R.B.; Stamatatos, E. Authorship Attribution for Social Media Forensics. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 5–33. [\[CrossRef\]](#)
46. Dipongkor, A.K.; Islam, S.; Kayesh, H.; Hossain, S.; Anwar, A.; Rahman, K.A.; Razzak, I. DAAB: Deep Authorship Attribution in Bengali. In Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 18–21 July 2021; pp. 1–9. [\[CrossRef\]](#)
47. Alrowili, S.; Shanker, V. ArabicTransformer: Efficient Large Arabic Language Model with Funnel Transformer and ELECTRA Objective. In *Findings of the Association for Computational Linguistics: EMNLP 2021*; Association for Computational Linguistics: Punta Cana, Dominican Republic, 2021; pp. 1255–1261. [\[CrossRef\]](#)

48. Wadhawan, A. Dialect Identification in Nuanced Arabic Tweets Using Farasa Segmentation and AraBERT. *arXiv* **2021**, arXiv:2102.09749v2.
49. Taboubi, B.; Nessir, M.A.B.; Haddad, H. iCompass at CheckThat! 2022: ARBERT and AraBERT for Arabic Checkworthy Tweet Identification. In Proceedings of the CLEF 2022: Conference and Labs of the Evaluation Forum, Bologna, Italy, 5–8 September 2022.
50. Bensoltane, R.; Zaki, T. Towards Arabic aspect-based sentiment analysis: A transfer learning-based approach. *Soc. Netw. Anal. Min.* **2021**, *12*, 7. [[CrossRef](#)]
51. Fabien, M.; Villatoro-Tello, E.; Motlicek, P.; Parida, S. BertAA: BERT Fine-Tuning for Authorship Attribution. In Proceedings of the 17th International Conference on Natural Language Processing, Patna, India, 18–21 December 2020.
52. Luyckx, K.; Daelemans, W. Authorship Attribution and Verification with Many Authors and Limited Data. In Proceedings of the 22nd International Conference on Computational Linguistics—COLING '08, Manchester, United Kingdom, 18–22 August; Association for Computational Linguistics: Cambridge, MA, USA, 2008; pp. 513–520. [[CrossRef](#)]
53. van Tussenbroek, T.; Viering, T.; Makrodimitris, S.; Naseri, A.; Tax, D.; Loog, M. Who said that? Comparing performance of TF-IDF and fastText to identify authorship of short sentences. Bachelor's Thesis, Delft University of Technology, Delft, The Netherlands, 2020.
54. Baturynska, I.; Martinsen, K. Prediction of geometry deviations in additive manufactured parts: Comparison of linear regression with machine learning algorithms. *J. Intell. Manuf.* **2021**, *32*, 179–200. [[CrossRef](#)]
55. Chadoulis, R.-T.; Nikolaou, A.; Kotropoulos, C. Authorship Attribution in Greek Literature Using Word Adjacencies. In Proceedings of the 12th Hellenic Conference on Artificial Intelligence, Corfu, Greece, 7–9 September 2022; pp. 1–9. [[CrossRef](#)]
56. Das, K.A.; Baruah, A.; Barbhuiya, F.A.; Dey, K. Ensemble of ELECTRA for Profiling Fake News Spreaders. In Proceedings of the CLEF 2020, Thessaloniki, Greece, 22–25 September 2020.
57. Imran, A.A.; Amin, M.N. Deep Bangla Authorship Attribution Using Transformer Models. In *Computational Data and Social Networks*; Mohaisen, D., Jin, R., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2021; pp. 118–128. [[CrossRef](#)]
58. Romanov, A.; Kurtukova, A.; Shelupanov, A.; Fedotova, A.; Goncharov, V. Authorship Identification of a Russian-Language Text Using Support Vector Machine and Deep Neural Networks. *Futur. Internet* **2020**, *13*, 3. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.