Towards a Deep Learning-based Approach for Hadith Classification

Moath Mustafa Ahmad Najeeb

Abstract — The exploration of Hadith sciences gains significant consideration over the most recent couple of years. Hadith is mostly the sayings of Prophet Mohammad. The Holy Quran represents the first origin of law in Islam then Hadith takes the second role. Many research efforts manage Hadith with respect to the “Isnad” and “Matn”; which are the main two pieces of Hadith. In this paper, we examine the chance of utilizing Deep Learning to process Isnad of Hadiths. Consequently, a definitive objective of our framework is to help in the systematic classification of Hadiths and differentiate among the correct (“Sahih”) Hadiths and the not accurate (“Da’ief”) Hadiths.

Index Terms — Deep Learning, Classification, Hadith, Arabic Language Processing, Isnad.

I. INTRODUCTION

There are a lot of concepts and techniques used by researchers and specialist to find the most ideal means that people can interface with PCs. Artificial Intelligence (AI) is one of these efforts. Nature Language Processing (NLP) considered one of the greatest AI strategies that support amazing correspondence with computers. Nature Language Processing intends to interchange information with PCs in many forms such as text and speech. Nature Language Processing deal with systems, hypotheses, and programming projects which may support in language detection, break down setting and create a communicated speech. Broadly, the operated pieces of the Nature Language Processing could be phonetic substances like sentences and words, these data sources controlled through Nature Language Processing methods which utilization specific principle, for example, syntactic guidelines, derivational standards, and inflectional guidelines. NLP utilize different kinds of assets, for example, corpus, and vocabulary and word reference [1].

Arabic Natural Language Processing takes an extraordinary consideration of studies over most recent couple of years, it pursues Nature Language Processing in numerous highlights, for example, labs and instruments. There are some Arabic Natural Language Processing research facilities started [2] and a few apparatuses created [3], yet these endeavors in labs and devices face a few impediments like essentialness diacritics of vowel and the solidity in Arabic language. Fig. 1 depicts an Arabic Natural Language Processing Model [4].

The sacred Quran Hadith is the subsequent central wellspring of Islamic enactment, then Hadith comes in the second rank. Every Hadith comprises of two sections: Isnad and Matn, Isnad is the storytellers’ chain through whom the portrayal will be transmit, in the interim Matn is simply the portrayal. Prior to over 1400 years, Muslims scholars are focusing of Isnad science since it assists in segregation among the correct (“Sahih”) Hadith, and the not accurate (“Da’ief”) Hadith. The researcher of Hadiths passes judgement depending on the portrayal Isnad (the chain) and the people (storytellers) who form the chain. Consider the following example (Hadith1), which is narrated in “Sahih Muslim” book [5]:

""ّحَدَّثَنَا غَيْبَٰبُنَّ مُوسَى، قَالَ أَخْبَرَنَا حَنْظَلَةُ بْنُ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي S

The Isnad of Hadith1 is:

""ّحَدَّثَنَا غَيْبَٰبُنَّ مُوسَى، قَالَ أَخْبَرَنَا حَنْظَلَةُ بْنُ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي سُفْيَانَ، عَنْ عِكْرِمَةَ بْنِ أَبِي S

While the Matn of this Hadith is:

""ّيَبْنِيِ الْإِسْلاَمِ عَلَى خَمْسِ شَهَادَةِ أَنْ لاَ إِلَهَ إِلَّاَ اللََُّّ وَأَنَّ مُحَمَّدًا رَسُولُ اللََِّّ، قَالَ رَسُولُ اللََِّّ صلى الله عليه وسل

Fig. 1. Arabic Natural Language Processing (ANLP) Model.

Scholars of Hadiths established seven criteria to classify Hadiths according to, these criteria are:
1. Number of storytellers (narrators) in each stage:
According to this criterion, Hadiths could be categorized to:
   A. “Mutawatir” (Consecutive)
   B. “Ahad” (Isolated)

2. Reference to authority:
According to this criterion, Hadiths could be categorized to:
   A. “Qudsi” (Sacred)
   B. “Marfu” (Elevated)
   C. “Mauqfi” (Stopped)
   D. “Maqu” (Severed)

3. Links of Isnad:
According to this criterion, Hadiths could be categorized to:
   A. “Musnad” (Supported)
   B. “Mursal” (Hurried)
   C. “Mutassil” (Continuous)
   D. “Munqati” (Broken)
   E. “Mu’dal” (Perplexing)
   F. “Mu’allaq” (Hanging)

4. Reliability and memory of the storytellers (narrators):
According to this criterion, Hadiths could be categorized to:
   A. “Sahih” (Correct)
   B. “Hasan” (Good)
   C. “Da’if” (Not accurate)
   D. “Maudu” (Forged)

5. Nature of Matn and Isnad:
According to this criterion, Hadiths could be categorized to:
   A. “Chad” (Irregular)
   B. “Munkar” (Denounced)
   C. “Mudraj” (Interpolated)

6. Receive approach:
According to this criterion, Hadiths could be categorized to:
   A. “Mudallas” (Concealed)
   B. “Musalsal” (Uniformly linked)

7. Exist of Hidden defect:
According to this criterion, Hadiths could be categorized to:
   A. “Mudtarib” (Shaky)
   B. “Maqlub” (Reversed)
   C. “Ma‘lul” (Defective)

There are a lot of details for each of the above norms which can be found in Hadiths books and dictionaries.

As for deep learning, and before investigating into how it works and its uses, it should be said that it is often confused by non-professionals in understanding the differences between deep learning and machine learning and their relationship to artificial intelligence. Therefore, it should be noted that deep learning is a form of machine learning, which in turn is a branch of artificial intelligence. Machine learning relies on algorithms that can be fed with structured data and analyzes it to reach conclusions, while deep learning is characterized by the presence of different levels of algorithms that form artificial neural networks (ANNs) that have the ability to understand unorganized data and complex patterns, such as languages, pictures and speech.

Therefore, deep learning is an approach that could be used in text classification, so, in this paper we will investigate the opportunity of using deep learning to classify Hadiths to the correct (“Sahih”) and the not accurate (“Da’ief”) Hadiths.

II. RELATED WORK

The examination of Hadith sciences gains significant consideration over the most recent couple of years. Hadith is mostly the sayings of Prophet Mohammad. Many research efforts manage Hadith with respect to the “Isnad” and “Matn”; which are the main two pieces of Hadith. Hadith Scholars give an extraordinary consideration on contemplating Isnad on the grounds that it discloses to us if Hadith is correct (“Sahih”) or not accurate (“Da’ief”); consequently, scholars established guidelines and systems for accomplishing that reason. Some of the vital books that collect the authentic Hadiths are: “Sahih Muslim”, “Sahih Al-Bukhari”, “Sunan Abu Dawod”, “Sunan AlTermizi”, “Suana Ibn Majah” and “Sunan Al-Nasa’i”. Some of the extreme popular Hadith researchers who examine Isnads, and investigate all the detail life story of all storytellers in the Isnad is “Ibn Hajar Al Asqalani”, he examined all data identified with storytellers like name, date of death, age, educators of storyteller, underestudies of storyteller, Rutba, Tabqa, and so forth, he recorded all these valuable data in his book “Taqreeb AlAtthabih”.

Numerous studies using computing methods presented in the literature that aim to help Arabic Natural Language Processing (ANLP); Najeeb et al. [6] show to their efforts in initiation an ANLP technical lab in Al-Qunfudah college, Umm Al-Qura University, Saudi Arabia, to assist ANLP and Islamic sciences. In [7], Najeeb proposed a novel Hadith processing approach based on Genetic Algorithms (GA), the approach aims to predict storytellers’ names and the other Part-Of-Isnads (POIs) for Hadiths. In [8], Najeeb presented an XML database for storytellers and Hadiths. In [9], Najeeb et al. constructed a corpus-based lexicon for Hadith science. A comparison study of classification and data mining techniques of Hadiths was conducted by Saloot et al. in [10].

A text mining tactic for knowledge extraction in Hadith presented by Harrag [11], the tactic is based on Finite State Transducers (FST). Alldhaln et al. [12] offered a classification system for Hadiths relied on decision trees; this system classifies Hadiths to four groups: “Sahih”, “Hasan”, “Daief” and “Maudo”. Alhawarat [13] prepared an approach that merges n-grams model with rule-based method to extract storytellers’ names from Hadiths. Mahmood et al. [14] showed a tactic that rely on regular expressions, this tactic process web pages that embrace Hadiths, they elicit the data from Hadiths as a conventional database. A lot of Hadiths duties presented in websites [15]–[20]; these duties include search tools, translation tools, e-books, etc.

One of the greatest accommodating frameworks in Hadith knowledge is Shamel library [15], this library examines the storytellers’ chain to provide a precise data regarding Matn and Isnad, it gives the researcher of immensely significant insights regarding storytellers, despite the fact that, it can't do a programmed judge on Hadith, consequently it cannot separate among correct (“Sahih”) Hadiths and the not accurate (“Da’ief”) Hadiths.

These days, many data frameworks proposed which may assist in contemplating Hadith science and perceiving
storytellers of Isnad, they may recover data like storytellers’ names, date of death, instructors, and so on, a portion of those frameworks may draw the storytellers tree that provide an overall view about storytellers’ chain, and explain which storyteller receive his story from which one [21], when all is said in done, these data frameworks need programmed judgment of Hadith and don't recognize correct ("Sahih") Hadiths and the not accurate ("Da'ief")

Other data frameworks attempt to utilize content mining strategies to extract Hadith data. AuthenTique [22] utilize the question arrived from the client and examine the grades of closeness to obtain the ideal Hadith, this framework dependent on vector space model and Term Frequency Inverse Document Frequency (TFIDF).

In [23], Kawther et al. classify Hadith by using supervised learning classification, they also investigate the Knowledge extraction from Hadiths, and they claim that the way of knowledge-extraction from Hadiths is determined depending on the Knowledge objects.

In [24], Alrazo examine the main characteristics of Hadiths by using data mining techniques, he took in consideration two books of Hadiths: "Alresalah" for "Alshafie" and "Almuamalat" for "Alshatibi". The researcher demonstrates that the "Hadith methodology" for "Alresalah" can be determined using data mining Techniques, and the main features of Hadiths in "Almuamalat" can also be determined using data mining Techniques. In [25], Alrazo constructed a specialist framework for Hadith science by setting information rules for Hadith dependent on Hadith books, these principles help Hadith researchers in following the storytellers of Hadith and passing judgment on the Isnad and Matta.

Ghazizadeh [26] utilize Fuzzy Logic to fabricate a specialist framework which may assist in deciding the accuracy of Hadith. In [27], Hyder and Ghazanfer speak to the storytellers’ chain as a diagram, and fabricated a Database to store the data of storytellers in a manner which may assist in following the recorded occasions and true to life data, they gave loads for circular segments among hubs and assess the different ways of storytellers, at that point they gave an evaluation for every one of them.

Harrage el al. [28] manufactured Information Retrieval framework to aid Hadith sciences utilizing limited state transducer-based element extractor; the researchers utilize the proposed framework to obtain the helpful data from storytellers’ chain. The writers get Hadiths from "Sahih Al-Bukhari” book and change these Hadiths from unstructured configuration to semi-structure one by separating every Hadith to nine sections, and afterward they constructed the corpus dependent on this organization.

Muhadith [29] is a cloud framework for Hadith characterization, it's a dispersed master framework use Service Oriented Architecture (SOA) to beat the communicational issue between the heritage online disseminated master frameworks. Muhadith mimics the job of Hadith researcher and attempt to recognize correct ("Sahih") Hadiths and the not accurate ("Da'ief") Hadiths with some clarification.

An HMM approach for indexing names from a collection of Hadiths in Indonesian language was proposed by Sari et al. [30], the indexed data comprises the Hadiths’ numbers, Hadiths’ collections, and person’s name. Hadiths’ numbers indicate Hadiths that cover the person’s name. In [31] Bounhas presented a significant examination of researches related to Hadiths in varied genres which are: Natural Language Processing (NLP), Knowledge Extraction (KE) and Information Retrieval (IR). The investigator explained the difficulty of comparing different researches in Hadiths as they use different collections of Hadiths, and the pre-processing stages are also different. In [32] Azmi et al. presented a notable survey on computational and natural language processing studies of Hadiths, they categorized studies into three arenas: General studies, studies relied on contents of Hadiths, and studies relied on storytellers of Hadiths.

As should be obvious from the above works, the enthusiasm of considering Hadith science was expanded, and the endeavors of using the new methods to serve it give idealistic outcomes. However, there are no studies examine the utilization of deep learning to handle Hadiths.

As for deep learning used as an approach that could be used in text classification, we will investigate in this paper the opportunity of using deep learning to classify Hadiths to correct ("Sahih") Hadiths and to not accurate ("Da'ief") Hadiths.

III. CLASSIFICATION USING DEEP LEARNING

To distinguish among the correct ("Sahih") Hadiths and the not accurate ("Da'ief") Hadiths, we must utilize one of the classification methods like Artificial Neural Networks (ANN), Decision tree, Support Vector Machine (SVM), Bayesian Classification, etc. Classification is a procedure that customs the training data that has labels defined in advance to construct a prototypical blueprint which may forecast the labels of novel data with unidentified labels.

Deep learning depends on processing a huge amount of data, by passing it through what are known as Deep Neural Networks (DNN), to train them to extract useful information from this data. Just as in the neural networks in humans, these artificial networks consist of neurons that contain several layers, starting with a layer for entering data and ending with a layer for outputting the result, and between them there are several layers for analysis and conclusion, called the hidden layers [33].

Each layer in this network captures specific information or a specific type of data, transfers it to another layer for training using learning algorithms, and then analyses it, to ultimately produce a model that can be used. And each neuron in the network has a weight, and this weight is what determines the importance of each element in relation to the rest of the elements, and with the increasing volume of the input data, the weights of the elements are adjusted repeatedly, in order to increase the accuracy of the final result. One of the most popular types of deep neural networks used is known as Convolution Neural Networks (CNNs), and it is this type of network that is often used in image processing. The depth and capabilities of the network increase with the increase in the number of layers, and the number of these layers sometimes reaches 150 layers [34].
A. Deep Learning Applications

Deep learning is not just used to recognize images and sounds, but is used in many scientific fields, including meteorological data analysis and biological research. One of the examples that can be cited here is that in 2016 a team at the University of California (Los Angeles) developed a device consisting of a microscope and a deep learning program [35]. As the microscope take a photo of the cells in the blood samples very quickly, without the need to damage the cells and make the samples unusable for future analysis. After that, it is analysed using deep learning software to identify cancer cells, and the accuracy of this analysis has reached more than 95%. It also recently began using a Convolution neural network called the Deep Image Prior to restore and colorize old or damaged photos and films. Many practical examples of this technology can be seen on the web developers page on the GitHub website [36].

In addition to the above, many deep learning-based technologies are used in the field of marketing and target audience selection in advertising; Amazon e-commerce company, for example, has a deep learning system based on a self-learning algorithm, called the “anticipatory shipping” system, which can determine customers’ buying patterns with great accuracy, and predict the expected price range and the product to be purchased. Thus, Amazon can ship products to distribution centres before customers place their orders [37].

Google also uses an advanced type of neural network in its machine translation service, called a Recurrent Neural Network (RNN), which relies on searching for complete sentences and not just individual words, as it expects words that should appear in the language that is being translated into [38].

B. Deep Learning Challenges

Despite the superiority of deep learning technologies in many areas, there are several challenges facing their widespread use. The first is the limited ability to collect what is known as big data in certain areas such as industrial fields, in contrast to the situation in consumer areas which is easy to collect big data.

In addition, it is becoming widely known that AI systems in general, and deep learning in particular are relatively easy to deceive, whether intentionally or automatically. For example, a research paper in the Nature journal [39] pointed to the possibility of deceiving a deep learning system in a self-driving car by placing certain stickers indicating to the car that a sign bearing the word “stop” means changing speed, and deceiving face recognition systems using patterns printed on glasses or hats, and deceiving systems that recognize sounds by using fake noises.

Thus, this technology, despite all its advantages, can lead to serious and unexpected errors when it encounters an unfamiliar situation or “believes” that it must interpret data that is not sufficiently clear or does not follow the general framework of the models on which it was trained.

Finally, accelerated attempts to overcome previous obstacles would allow us to assess the strengths and weaknesses of neural network architecture, and it might not take a long time before deep learning finds solutions to the major challenges that have eluded our human minds so far. However, this progress without restrictions leads some people to express their fear that humans will pay a heavy price if deep learning reaches capabilities that allow it to defeat the human mind in general. As this may lead to the disintegration of the social and economic ties that bind our societies, by pushing people into unemployment or even slavery.

C. Models of Deep Learning

Deep Neural Networks (DNN) have numerous and deep collection of hidden layers. Hidden layers are added into the network to give more information when a job is very complicated for a small network. The number of hidden layers could be more than one hundred. Deep Neural Networks have very good accuracy, and they have many types; some of them are used to handle images, and the others are used to handle text.

The numerous deep neural networks models have some differences in the way they are connected. The following discussion will explain the differences between the most popular models which used in text classification task.

1. The Recursive Neural Network (ReNN) model:

As exposed in Fig. 2, the construction of the ReNN model is built as syntax tree, the model could by itself learn the grammatical and logical rules concerning with meaning of text recursively. The words of Hadith’s Isnad are considered as the leaves nodes of the tree. At that time the model will use the matrix of weights to join the parent node with the leaves, it is good to notice that the leaves and the parent node have a similar dimension. So, to guess the label, all the Hadith’s Isnad is represented by doing a recursive combination of all nodes into a parent [40].

Fig. 2. Recursive Neural Network (ReNN) model.

2. The Multilayer Perceptron (MLP) model:

This model called also “Vanilla” neural network, which aims to discover and get the features in automatic way. The construction of the (MLP) model depicted in Fig. 3, it consists of three layers: input, output, and hidden layer. The hidden layer comprises the activation function for the nodes, these nodes links with some weight, all these words are considered as one bag, so, for one Isnad we will get one bag.

In literature, we can find some studies that proposes a specific multiyear perception models such as Paragraph Vector (Paragraph-Vec) [41] and Continuous Bag of Words (CBOW) [42].
3. The Recurrent Neural Network (RNN) model:
This model is largely utilized because of his capability of taking dependency in a wide range using recurrent method. The RNN linguistic method learns from old data, taking into a count the position of the data between all the words appropriate for text classification job. The construction of the (RNN) model depicted in Fig. 4, the words that considered as inputs are signified by a certain vector by utilizing a method for embedding words. After that, the vectors that represents these words are move gradually and steadily into RNN nodes one after one. The following hidden layer is getting the production of RNN nodes which is by the way have identical dimension as the input vector. The RNN model assign equal weights to every input word. At the end, the final hidden layer’s output used to guess the input’s label [43].

4. The Convolutional Neural Networks (CNNs) model:
This model used in various natural language processing jobs such as text classification. CNNs uses convolving filters to excerpt images’ features in the image classification systems, they use convolutions which settled by various kernels to several segments of sequence. As depicted in Fig. 5, the classification process of the text needs to form the text in a vector as the representation of the image. The text’s features may filter from many positions. The vectors of the input Isnad could be joined or connected with a matrix, which in turn, could move gradually and steadily to the convolutional layer, this layer comprises many dissimilar dimension filters. At the end, the outcomes that obtained from the convolutional layer sent to the pooling one. To get the concluding vector of the Isnad, the model links the outputs of pooling together in a chain, then the last vector used to guess the label [44].

5. The Attention-based model:
This model is effectively handling the classification tasks related texts because it is using the interpretability to enhance the performance. In [45] Bahdanau et al. introduced an attention-based model for machine translation. After that, Yang et al. [46] generated an attention-based model to get a good conception by engaging text components that have tremendously information, this model called hierarchical attention network (HAN).

6. The transformers models:
These models don not pay attention to the successive information appropriate for big data sets, so they are becoming common in handling many tasks in natural language processing such as classification of texts. These models which trained previously, could efficiently do learning for semantic description and greatly support the tasks of natural language processing.

They are automatically extract semantic information by employs unsupervised approaches, and by structure data sets that are trained previously, consequently, the system could recognize semantics by learning [47].

7. The Graph Neural Networks (GNNs) model:
This model gets an outstanding performance using encoding method for constructions of the sentence to accomplish the classification task of relation [48] and accomplish the label task for semantic role [49] and accomplish the tasks of translation carried out by a computer [50].

IV. CONCLUSION
The distinguish process among the correct (“Sahih”) Hadiths and the not accurate (“Da’ief”) Hadiths is a vital mission in “Hadith judgment” science. As this paper clarified; our examination shows a good chance to construct a computerized classifier to categorize Hadiths using the Deep Learning technique. In future work, numerous models of Deep Learning will be examined to form a preliminary model to classify Hadiths and then to judge Hadiths automatically. A comparative analysis will be conducted for these models.
ACKNOWLEDGMENT

The author would like to thank the Deanship of Scientific Research at Umm Al-Qura University for the financial support (project number: 43508026). As well, the author would like to express his sincere thanks and gratitude to Dr. Naser Younes Sabra, Ph.D. in Islamic Studies and Hadith sciences, for his participation and support, without him, this research would not have completed.

REFERENCES


Moath Mustafa Ahmad Najeeb received the B.Sc. degree in computer science from Al al-Bayt University, Jordan, in 2000, the M.Sc. degree in computer science from Amman Arab University, Jordan, in 2006, and the Ph.D. degree in computer information systems from the University of Banking and Financial Sciences, Jordan, in 2011. From 2000 to 2004, he worked as a Web Developer, Systems Analyst and Oracle Database Administrator in UAE. He worked as an Assistant Professor with the Computer College, Jadara University, Jordan, in 2011. He is currently an Assistant Professor with the Computer College, Al-Qunfudah, Umm Al-Qura University, Saudi Arabia. His research interests include Arabic natural language processing, pattern recognition, and data mining.