

## A systematic review of text classification research based on deep learning models in Arabic language

Ahlam Wahdan<sup>1</sup>, Sendeyah Hantoobi<sup>2</sup>, Said A. Salloum<sup>3</sup>, Khaled Shaalan<sup>4</sup>

<sup>1,2,3,4</sup>Faculty of Engineering and IT, The British University in Dubai, United Arab Emirates

<sup>3</sup>Research Institute of Sciences and Engineering, University of Sharjah, United Arab Emirates

---

### Article Info

#### Article history:

Received Apr 10, 2020

Revised Jun 6, 2020

Accepted Jun 17, 2020

---

#### Keywords:

Learning management system

Neural network algorithms

Systematic review

Text classification

Virtual learning environment

---

### ABSTRACT

Classifying or categorizing texts is the process by which documents are classified into groups by subject, title, author, etc. This paper undertakes a systematic review of the latest research in the field of the classification of Arabic texts. Several machine learning techniques can be used for text classification, but we have focused only on the recent trend of neural network algorithms. In this paper, the concept of classifying texts and classification processes are reviewed. Deep learning techniques in classification and its type are discussed in this paper as well. Neural networks of various types, namely, RNN, CNN, FFNN, and LSTM, are identified as the subject of study. Through systematic study, 12 research papers related to the field of the classification of Arabic texts using neural networks are obtained: for each paper the methodology for each type of neural network and the accuracy ration for each type is determined. The evaluation criteria used in the algorithms of different neural network types and how they play a large role in the highly accurate classification of Arabic texts are discussed. Our results provide some findings regarding how deep learning models can be used to improve text classification research in Arabic language.

Copyright © 2020 Institute of Advanced Engineering and Science.

All rights reserved.

---

### Corresponding Author:

Said A. Salloum,

Research Institute of Sciences and Engineering,

University of Sharjah,

Academic city Road, P. O.Box 27272 Sharjah, United Arab Emirates.

Tel: +971 6 5585000, Fax: +971 6 5585099

Email: ssalloum@sharjah.ac.ae

---

## 1. INTRODUCTION

The classification of texts is a method of searching for data and exploring the data among large data and classifying them into groups for easy reference [1-5]. Internet pages, books [6], magazines, and social media [7-20], have become a rich source of information that needs to be categorized and organized for easy reference [21-25]. There has been a lot of research in this field, but most of it involves the classification of English and Spanish texts and texts in other languages. There is a lack of research in Arabic text classification, and the techniques and algorithms used on English texts do not fit the Arabic language texts. This is because the Arabic language has certain characteristics where the structure of the word is concerned [26]. The main aim of this study was to make reviewing Arabic text classification based on neural networks. So, this research summarized the list of neural network techniques that were used to classify Arabic texts and determined which ones were more efficient and accurate. Finding the gap in the current literature was another important aim. In this research, we focused more on the use of neural networks in the classification of Arabic texts. We discussed the concept of the classification of texts first alongside the types of classification then we presented the topic of neural networks, their types, and the classification processes for the different types. Through systematic review, the research questions were initially identified, then the research strategy was developed to obtain the topics more accurately. Subsequently, quality

assessment criteria were developed to assess the quality of the research papers in order to extract the important ones. In the discussion section, all research questions were answered in detail.

In order to achieve the main aim of this study, we defined our research questions as follows:

- RQ1: Which corpora were mainly used for Arabic text classification? Were they open source or created by the author?
- RQ2: Which countries focused on publishing research on ATC?
- RQ3: Which databases had more publications?
- RQ4: During which years was there more focus on ATC?
- RQ5: What were the types of neural network algorithms used to classify Arabic texts? And which one was most used?
- RQ6: What were the efficiency measures used?
- RQ7: Did the author compare NN with other techniques? And which ones were the most frequently compared with NN?
- RQ8: Which type of neural network algorithm proved most efficient?

## **2. LITERATURE REVIEW**

### **2.1. Definition of TC**

The text classification process is the process of automating a set of documents into specific groups based on the content of the text itself through the use of certain technologies and algorithms [24, 27-30]. Elhassan and Ahmed [30] also defined the classification of texts as a method of searching for data and exploring them among large data and classifying them into groups for easy reference. According to [31], the classification of texts was defined as the process of organizing documents according to a known and pre-existing structure of specific categories that suited this type of text. Others mentioned that there was a slight difference between text categorization and classification. Whereas the categorization of texts entailed sorting them according to their content, their classification placed them in certain groups that suited their content according to author, subject, title, language, and other classifications [32]. Specialized research on the classification of texts has increased significantly due to the enormous data available from many sources, including Internet pages, e-mail messages, news pages, texts circulated through social media, reports, and journal articles. Therefore, this research focused attention in order to make the best possible use of all these data and classify them [33].

### **2.2. Arabic text classification**

Arabic is one of the six primary languages worldwide. Its use is spread across many countries in the world, and it is the basis of the Arab world, but some Islamic countries also use the Arabic language as their primary language because it is the language of the Qur'an. The Arabic language consists of 28 letters, including hamzas and elmodod. Each letter has certain structures that depend on the position of the letter in the whole, that is, whether it is at the beginning of the word, in the middle of the word, or at the end of the word [34]. There are no upper- and lower-case letters in Arabic unlike English or other languages [31]. The morphology of the Arabic language is not easy; it may be complicated as it has radical words, prefixes, and suffixes; moreover, unlike those in other languages, its words do not consist of sequential forms as its word structures are completely different and depend on position in the sentence and meaning [31]. As mentioned above, the classification of texts has become widespread, and most of the research includes the classification of other languages, such as English. Little of the research on the classification of texts has focused on Arabic. Many techniques and algorithms have been created to classify English texts, and they give excellent results and high accuracy due to the nature of the language, letters, and words. But when it comes to the Arabic language, the application of these algorithms does not give similar accuracy and validity in classification. Therefore, appropriate algorithms have been created to synthesize words in the Arabic language, but they have not yet demonstrated fully reliable merit for the classification process [34].

### **2.3. Classification technique**

There are two basic methods of classifying texts. One is the grammar or linguistic based method, which depends on preparing certain rules and operating them in some expert systems in order to classify the texts. The second method is the process of feeding the text extrapolation process with a set of documents called training documents that have previously been classified according to specific categories [30]. Below are the explanations of the two methods.

### 2.3.1. Manual and statistical techniques

Manual text classification (TC) is done by writing specific queries manually for each category, for example sports, nutrition, clothing, and health. Then the text entered within a specific search engine is categorized based on the predefined queries. However, this method works on small texts, not on large ones or on many documents. The accuracy of this approach to classification depends on the validity of queries and the skill behind the query design. Expert systems have been created that prepare rules and inquiries. An example is Construe-TIS, the one created for Reuters News, which was able to apply 674 categories and to get to know more than 170,000 companies [32]. This classification was not based on individual words but on concepts graded from the actual text as shown in Figure 1. Techniques which rely on mathematical rules and principles are also called statistical text classification. These techniques are suitable for small data, such as “Frequentist procedures, Bayesian procedures, and the Binary and multiclass procedures” [34].

```
(if
test:
  (or [australian-dollar-concept]
    and [dollar-concept]
      [australia-concept]
      (not [us-dollar-concept])
      (not [singapore-dollar-concept])))
action: (assign australian-dollar-category))
```

Figure 1. The rule of the category Australian Dollar

### 2.3.2. Machine learning techniques

Due to the large amounts of information during the past two decades, there emerged a need for technical methods of classifying huge texts as statistical and manual methods had never become useless in this field [35-48]. Therefore, machine learning classification techniques appeared which aim to classify unstructured texts and documents based on certain algorithms designed for this purpose. Machine learning techniques can be divided into groups as shown in Figure 2 [34]. In the Supervised Learning technique, different data are recognized, whether visual, audio, or text. The data are compared with the expected ideal data or results through a process called backpropagation in which the data are directed from the output layers to the input layers and errors are corrected and minimized to achieve better accuracy. In Unsupervised Learning, the learning process takes place during the processing. There are no previous data for comparison as the network analyzes the data and processes them; then, builds a function to determine the error rate and reduce it to obtain high accuracy. The Semi-Supervised learning technique is a combination of the previous two techniques. Researchers saw that there was little unnamed data that could contribute to improving the accuracy of the learning process that took place during the data classification stage. So, a small percentage of the data was used in Semi-Supervised learning.

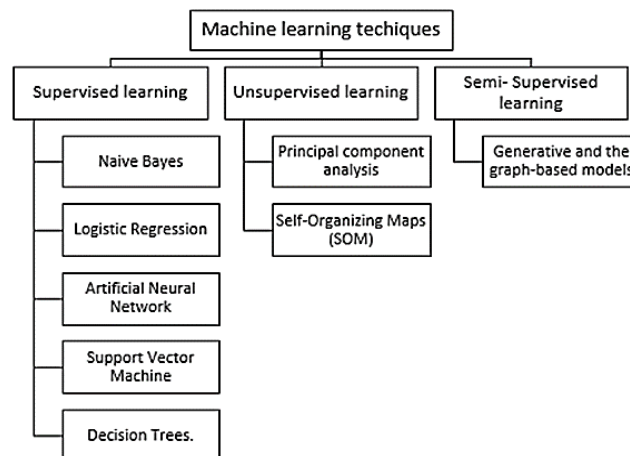


Figure 2. Machine learning techniques

## 2.4. Neural network

Algorithms of neural networks give very accurate results in the area of NLP and support deep learning very significantly as they solve some errors and problems related to the variance of data resulting from the process of deep learning [36-39]. Neural networks are used in the classification of texts to address linear and non-linear problems. The backpropagation model classifies texts using neural networks by means of a group of nodes that form and represent a mathematical model of biological nerves. No interference and self-learning affect these neural networks as they are used to identify systems and patterns and to classify text and image processing [40]. Many consider neural networks to be units that are connected to each other and intertwined similarly to the dynamicity of the neuron. In addition, through this process of entanglement, the inputs are made and the other neuron allows the exit of the transmitted outputs. The design of ANN is similar in structure to that of the brain and its neural network. All types of sabotage of the neural network have a basic principle, which is that every neuron in the network receives inputs, processes them, and sends outputs. Each neuron is linked to at least one neuron, and each connection or association has a specific digital weight called the weight factor, which works to reflect the importance of the communication between neurons in the network [41]. Some popular primary types of neural networks follow.

### 2.4.1. Feedforward neural network (FNN)

It is a neural network whose cells are linearly connected to each other and do not represent a cycle like any other neural network. The single layer perceptron (SLP) is the simplest form of this type of neural network since the inputs are directly related to the output. In the SLP, the inputs go through several layers of transformation, so they are considered very suitable for categorizing texts [42]. The structure of the feedforward neural network is shown in Figure 3.

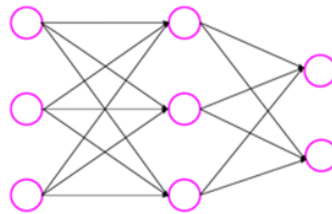


Figure 3. FNN

### 2.4.2. Convolutional neural network (CNN)

Convolutional neural networks are inspired by the multi-layer perceptron (MLP) and are designed to extract the spatial structure in image data and the positions of objects that can be used in the image [43]. The same principle and idea were used to classify texts and one-dimensional words, and this was done through interaction with neighboring neurons. Its name is “convolutional” due to folding process that occurs between neurons during the classification process. The structure of the convolutional neural network is shown in Figure 4.

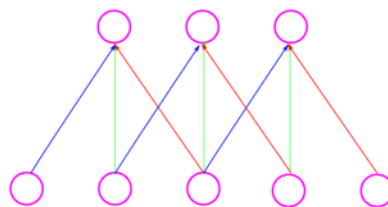


Figure 4. CNN

### 2.4.3. Recurrent neural network (RNN)

The structure in this type of network is done continuously and sequentially since the outputs of the previous stage are the inputs of the next stage. By contrast, with traditional neural networks, the inputs and outputs are separate units from each other. Sometimes there is a need to predict the desired word,

and therefore, the need develops to refer to the words prior to the review, and this is what recurrent neural networks do. A hidden layer has been used to help with this type of network and to complete the process [44]. An example of a frequent neural network is shown in Figure 5.

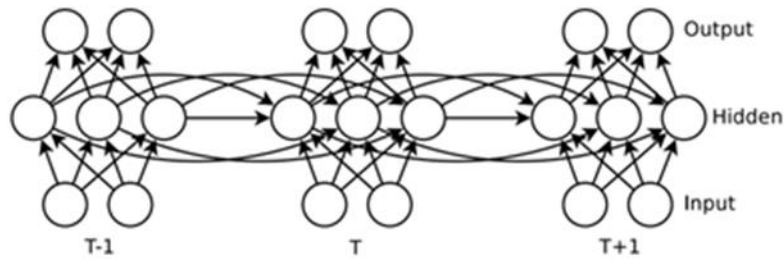


Figure 5. RNN

**2.4.4. Long short-term memory (LSTM)**

LSTM are long short-term memory networks as they are a specific recurrent neural network structure (RNN) designed for use in modeling long-range time series that are observed to operate more precisely than RNN. The neurons of this network are made up of units that are gates; in this way, the network can control the flow of the inputs that lead to the final output. Thus, only a few inputs may participate in the output, so the error rate is reduced [45]. LSTM contains units called memory blocks, and they are located in the hidden layer. These blocks have self-connections in memory cells that record the time status of the network during the work. They also contain units called gates that contribute to controlling the flow of inputs and outputs. Their input gates control the activation process for entering information into the memory. As for the output port, it controls the output after the activation process, which takes place at the entry gate. A portal called the forget gate, which addresses the weakness of LSTM in determining flows for specific units, has also been added [46]. Figure 6 shows the structure of LSTM RNN.

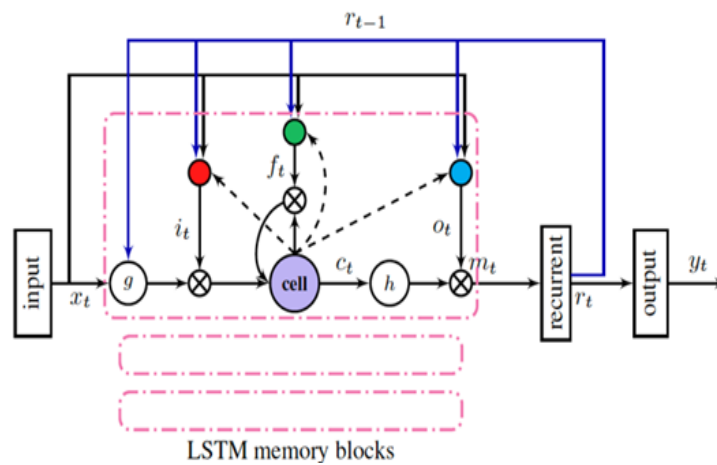


Figure 6. LSTM

**2.5. Arabic Text classification process**

There are three main stages in the classification of texts, which based on the type of neural network used as shown in Figure 7.



Figure 7. Arabic text classification process

### 2.5.1. Data pre-processing

At this stage, the word is filtered and returned to its original root by removing all the changing polices related to the whole, for instance, hamzah, Ta Marbouta ( “ة” ), tashkeel from all the words, numbers, and punctuation marks. Also, the words that link words, such as ( “بالنسبة ل ، لذلك” ), are removed. Prefixes and suffixes are also removed, and all words of the same root are grouped. This stage aims to give precision to the classification process and save time [34]. After removing all of the appendages, the word is returned to the root by three methods: “the root-based stemmer, the light stemmer, and the statistical stemmer.” Then the document is classified based on the vectors for each document [30].

### 2.5.2. Text classification

The training phase takes place here; a specific algorithm is applied to the words obtained from the previous stage in order to complete the classification process. Several comparisons can also be made here by using a different algorithm to classify the text and then choosing the best in terms of performance and the result of accuracy in classification [30].

### 2.5.3. Evaluation

Afterwards, and at this stage, the effectiveness of classification is evaluated; several techniques are used, but the most famous and frequently used one in the classification of texts is “F1, precision and recall” [30].

## 3. RESEARCH METHOD

Figure 8 demonstrates the key phases we followed. Generally, research questions were identified then we identified the search strategy and the keywords we used in our research. Subsequently, we identified the quality assessment questions. After that, extraction of the papers occurred. In the last step, we did critical analysis of the chosen papers.

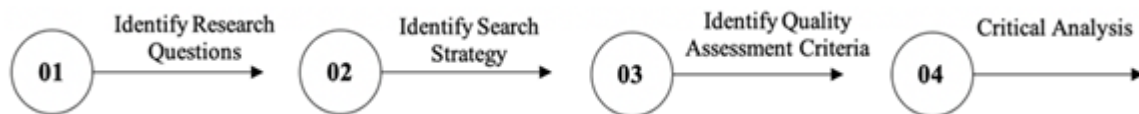


Figure 8. Systematic review methododloy

### 3.1. Data source

This study has begun in January 2020, and we included most researches from 2016 to 2019 and some researches before that period. The databases we used follow: IEEE, Science Direct, Springer, ACM. The Figure 9 shows the data source and the number of articles that were used.

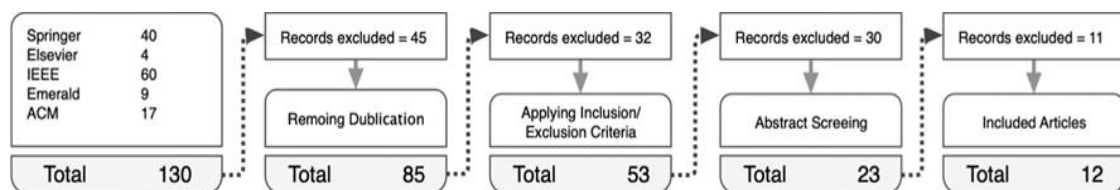


Figure 9. Data source

### 3.2. Search strategy

The keywords we used to collect all the previous studies follow:

- “Arabic text” and “classification” and “neural networks”
- “Arabic text” and “classification” and “deep learning”
- “Arabic script” and “classification” and “deep learning”
- “Arabic script” and “tagging” and “deep learning”
- “Arabic script” and “categorization” and “deep learning”

### 3.3. Inclusion criteria

In the selection criteria step, inclusion and exclusion criteria were set to ensure that the research included in this study was valuable and relevant and would lead us to our main aim.

#### 3.3.1. Exclusion criteria

- Must be a text classification study
- Must be for Arabic text classification only
- Must use neural network for text classification
- Must report the model and its performance measures, for instance, accuracy or another metric
- Must address summary for the corpus used.

#### 3.3.2. Selection criteria

- Papers not for Arabic text classification
- Paper didn't address the accuracy
- Paper had not been published in journal or conference
- Paper did not use NN

### 3.4. Quality assessment

In this part, we designed quality assessment questions to make a checklist for the research and ensure that it would satisfy the aim of this systematic review. Below are the questions.

Q1: Was the corpus identified and described well?

Q2: Was the corpus identified well regarding the extent of training and testing?

Q3: Were the text classification (model / framework) steps described clearly?

Q4: Did the author make comparisons with techniques other than the NN he used?

Q5: Was the performance of the model identified clearly?

Scale: A three-point scale was used in this assessment. If the paper addressed the question exactly, it would be graded 1, and if it did not address it, it would be graded 0, but if it answered the question partially, it would be graded 0.5. Table 1 shows the result of evaluating the research using the designed assessment questions.

Table 1. Quality assessment evaluation

Study	Q1	Q2	Q3	Q4	Q5	Total	Percentage
S1	1	1	1	1	0	4	80%
S2	1	1	1	1	0.4	4.5	90%
S3	1	1	1	1	1	5	100%
S4	1	1	1	1	1	5	100%
S5	1	1	1	1	1	5	100%
S6	1	1	0.5	1	1	4.5	90%
S7	1	1	1	0	1	4	80%
S8	1	1	1	1	1	5	100%
S9	1	1	1	1	1	5	100%
S10	1	0.5	1	0	1	3.5	70%
S11	1	0.5	1	1	0	3.5	70%
S12	1	1	0.5	1	1	4.5	90%

## 4. RESULTS AND DISCUSSION

- a. RQ1: Which corpora were mainly used for Arabic text classification? Were they open source or created by the author?

From the graph shown in Figure 10, it is clear that the documents that were used in the process of classifying the Arabic texts were mostly open source and news sites with 4 sources in each case. Some also used books: two research papers applied the classification process to used books. One research paper classified guest reviews of a hotel, which could be considered to not be an open source. That is, the number of open source resources in the research papers under study was 11, and the number of non-open source ones was only one as shown in Table 2.

- b. RQ2: Which countries focusing on publishing researches in ATC?

Many researchers from Arab countries were interested in studying and publishing research related to the classification of Arabic texts. Jordan was at the forefront of the countries that were classified. After the liquidation, we obtained three research papers of researchers from Jordan, then some from Morocco, the United Arab Emirates, and the United States of America at a rate of two research papers from each country. As for the least published countries in the period identified in this research, Algeria,

Saudi Arabia, and Tunisia qualified, with one research paper from each country. It is clear that the Arab countries have started to study and publish on the subject of the classification of Arabic texts in order to use the language in almost all fields. Also, these countries have started to use Artificial Intelligence techniques in almost all fields. Therefore, the Arabic language should make a contribution in the fields of Artificial Intelligence and Machine Learning. Figure 11 illustrates the number of studied articles in each country.

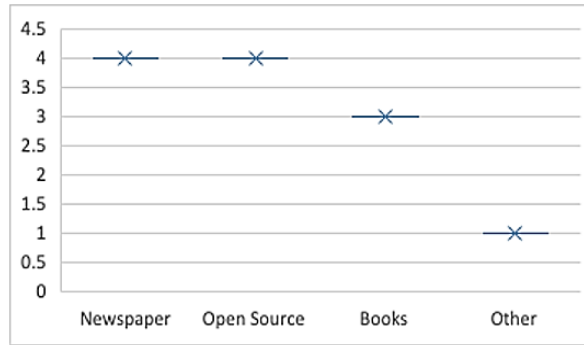


Figure 10. corpora were mainly used in the selected papers

Table 2. Open sources corpora used and private sources

Study	Number of documents used	Source
S1	319,254,124 words and 111,728 documents	Three Newspaper Online Article
S2	200K articles	SANAD (Alarabiya.net , alkhaleej.ae and akhbarona.com)
S3	693k books	BRAD (Arabic Corpus in Github)
S4	5070 documents	CNN Arabic news
S5	16K documents	Aljazeera.net and Alharbi Saudi press agency
S6	1400 documents	Aljazeera news website, Saudi Press Agency.
S7	15 K tweets	AraSenTi
S8	453 Hadith	Collection of Prophet be upon him (ﷺ) hadith from the Nine books encyclopedia
S9	1065 Hadith	Collection of Prophet be upon him hadith (ﷺ)
S10	22,429 documents	Open Source Arabic Corpora (OSAC)
S11	11 books	10 Tashkila books as a collection from Islamic heritage books And Holy Quran
S12	2291 reviews texts 6029 annotated sentences	Arabic Hotels' reviews

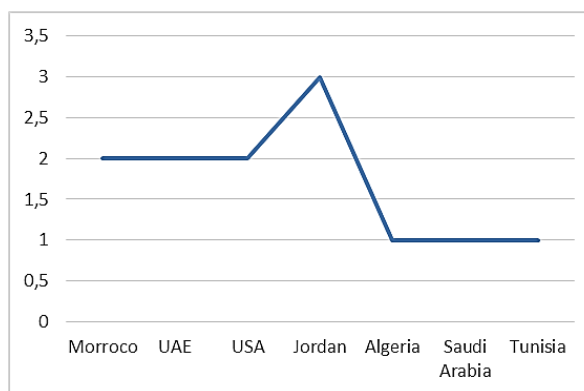


Figure 11. Countries publishing research in Arabic text classification

c. RQ3: Which database have more publications?

According to the chart below in Figure 12, a number of well-known scientific journals published research related to the classification of Arabic texts. IEEE was at the forefront. 60 research papers were obtained from IEEE, 40 were obtained from Springer, and 17 were obtained from ACM. As for the scientific



journals in which the number of published papers on the topic was small, they were Emerald and Elsevier with 8 and 4 research papers on average, respectively. The reason for IEEE's position at the top was the holding of a number of conferences with this focus, with many researchers participating in this field with their research papers. In all cases, the subject of the classification of Arabic texts is still under development when compared to that of other languages, such as English and Spanish.

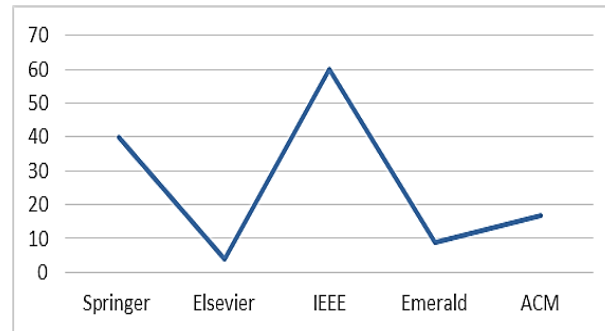


Figure 12. Number of Articles based on science journals

d. RQ4: During which years was there more focus on ATC?

Based on the graph illustrated in Figure 13, the period from 2017 to 2019 witnessed a significant development in the field of Arabic text classification research. Six research papers were distributed between 2018 and 2019, that is, three research papers for each year. In the period between 2013 and 2016, only four research papers were studied; they were distributed as follows: one article in 2015 and three in 2016. Only two research papers were obtained between 2009 and 2012. From the data we have, it is clear that research interest in the field of the classification of Arabic texts is starting to increase due to the importance of the Arabic language, its use in more than one field, and its use in more Arab Islamic countries.

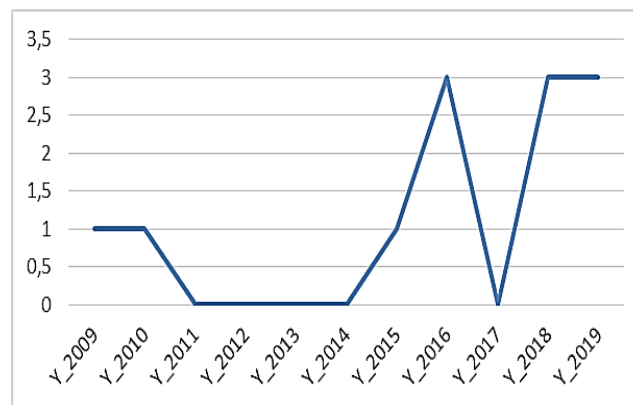


Figure 13. Number of articles selected per year

e. RQ5: What were the types of neural network algorithms used to classify Arabic texts? And which one was most used?

Many types of neural networks were used in previous research. They included convolutional neural networks (CNN), recurrent neural networks (RNN), BiRNN, LSTM, the deep belief network (DBN), backpropagation autoencoder (BPNN), the multilayer perceptron (MLP), and feed forward neural networks (FFNN). LSTM and CNN were the most used neural networks. Both were used in four studies. LSTM were used in [47-50]. Convolutional neural networks (CNN) were used to classify Arabic text in [50-53]. Recurrent neural networks (RNN) were used in only one study [52]. In this study, the author also tried to build ensemble methods, which combined RNN with CNN, in order to improve efficiency. In the same study, another model was built using a special type of RNN: BiRNN. Three-layer feed-forward NN have been used twice in [40] and in [54]. But, in [55], the multilayer perceptron (MLP) was used as a part of 3-layer

feed-forward NN. The deep belief network (DBN) was only used in one study [56]. Backpropagation autoencoder was used just once in [57]. Figure 14 shows the types of NN and how many times it has been used in all reviewed papers. Table 3 shows the references for each type of neural network founded in each paper. The papers did not mention the reason behind choosing these types of NN. We believe that the reason should be stated.

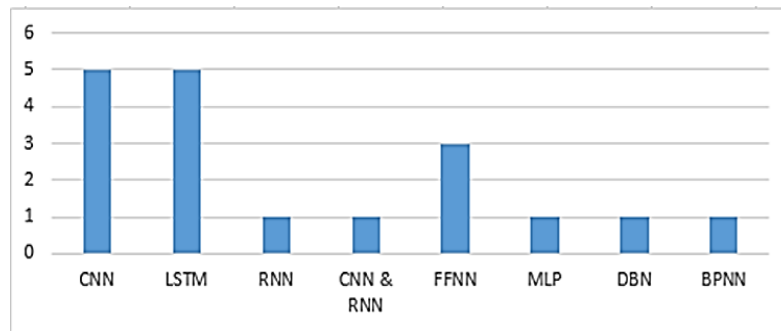


Figure 14. Types of NN

Table 3. Types of NN

Types of NN	References
CNN	[47-50]
LSTM	[50-53]
RNN	[52]
CNN & RNN	[52].
Feed forward NN (FFNN)	[40, 54]
MLP	[55]
DBN	[56]
Backpropagation Auto_encoder 3 layers (BPNN)	[57]

f. RQ6: What were the efficiency measures used?

The efficiency measure mainly used in these studies was primarily accuracy; the others used included Recall, F-measure, and Precision as shown in Figure 15. Accuracy was used in 8 studies in our systematic review as shown in Table 4, but not all of these studies identified it well. Accuracy was identified well and the researchers cleared the definition or the equation in only four studies: [48-50, 57].

The Recall efficiency measure came in second position: it was used in five studies. Precision and F-measure were the least used; they were used in only 4 studies. Table 4 shows the efficiency measure that used in each paper.

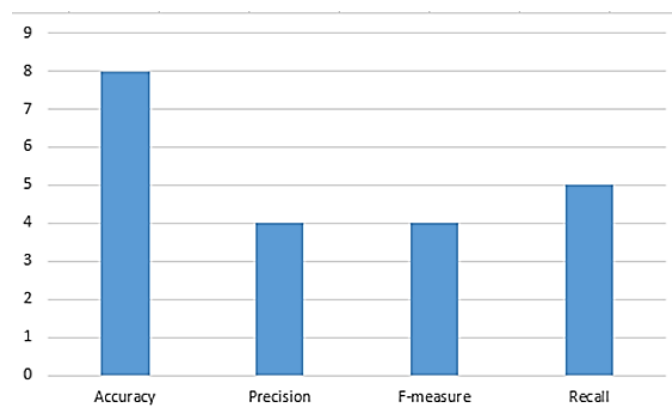


Figure 15. Efficiency measures

Table 4. Efficiency measures

Efficiency measures	References
Accuracy	[47]-[53], [57]
Precision	[40], [54], [56], [57]
F-measure	[54], [55], [56], [57]
Recall	[32], [40], [54], [56], [57]

- g. RQ7: did the author used other techniques to compare NN with? And what were the most one used to compare with NN?

Many machine learning algorithms have been used to compare it with the efficiency of neural network in classifying Arabic text. Some of them are supervised machine learning algorithms and others are unsupervised machine learning algorithms as shown in Figure 16. The algorithms that have been used in Arabic text classification based on our systematic review are Support Vector Machine, Logistic Regression, Naïve Bayes, KNN, Decision Tree, Random Forest and XGBoost. As it is clear from Table 5, SVM and Naïve Bayes were the most algorithms used in the past researches to compare it with NN. Followed by Decision Trees where it has been used twice and other techniques only one time.

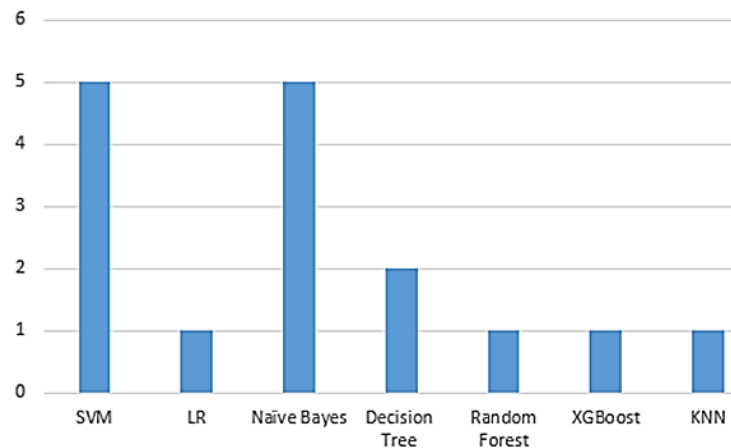


Figure 16. Machine learning algorithm used

Table 5. Machine learning algorithms used

Machine Learning Algorithms	References
SVM	[50, 51, 55, 56, 57]
LR	[51]
Naïve Bayes	[40, 50, 55, 56, 57]
Decision Tree	[50, 57]
Random Forest	[50]
XGBoost	[50]
KNN	[56]

- h. RQ8: Which type of neural network algorithm proved most efficient?

After the deep systematic review, answering this question was not easy, nor was the answer comprehensive. Hence, we cannot generalize one type of NN as the best in Arabic Text Classification for many reasons we have discovered in this systematic review. We found that the databases used in all these studies were different, and no two similar corpora were used. Some were open source, and others had been created by the authors. Another reason was that the neural networks used were different in each study. In the cases where the researchers used the same type of NN, there was other missing information. The researchers did not indicate in detail the parameters used in these networks. Hence, usually these algorithms would be tuned by changing parameters. This made it difficult to compare or to make sharp decisions about which neural networks were the best. The only clear finding would be arrived at if more than one neural network was used in the same research. The result of this comparison already stated in each study had done this. In general, this is a comparison, but as discussed before, we cannot state it. In [55] paper, MLP did not revert to high accuracy in two studies, mostly around 50%. The CNN result was better than

the MLP once but lower than DBN and LSTM in the following papers [47-49], respectively. DBN reverted to high results, but it was only tested for one study [56]. LSTM in all studies reverted to high efficiency measures [50-53]. Because, LSTM can control the flow of the inputs that lead to the final output. Thus, only a few inputs may participate in the output, so the error rate is reduced [45].

## 5. CONCLUSION

Arabic text classification is the process of classifying texts into groups by subject, title, or author. This study was a systematic review that focused on using neural networks as a basis for classifying Arabic text. It included critical analysis, for the papers that were reviewed in many senses. It criticized the corpora used in classification and the methods used to build the model; whether they entailed neural networks or other machine learning techniques. It also listed the types of neural networks used: namely, CNN, LSTM, FFNN, and BPNN. Moreover, it focused on the year of publication and the databases that published these papers. In addition, it criticized the efficiency measures used to compare the models built.

After the deep systematic review, we found that the corpora used in all these studies were different, and no two similar corpora were used. Some were open source, and others have been created by the authors. The Arabic natural language community is encouraged to create many Arabic open resources, like Kaggle, so that researchers can find many open source corpora with which to test their models. Also, researchers are encouraged to publish their results, accompanied by detailed efficiency measures that produced these results. Unfortunately, some of the papers reviewed did not mention the measures used clearly and precisely, and the reason behind choosing this particular evaluation measure. In other words, the definition and the equation used to measure the efficiency of the model must be declared such that if another researcher decided to use the same dataset to evaluate their methodology, both studies can be compared. The papers did not mention the reason behind choosing the types of NN. We believe that the reason should be stated.

Moreover, we cannot generalize one type of NN as the best in Arabic Text Classification for many reasons we have discovered in this systematic review because the neural networks used were different in each study. In the cases where the researchers used the same type of NN, there was other missing information. The researchers did not indicate in detail the parameters used in these networks and how they are tuned. Usually, the machine learning algorithms are tuned by changing parameters and re-running the experiments to get significant results. This made it difficult to compare or to make sharp decisions about which neural networks were the best. Arabic text classification still has room for a lot of improvement. Hence, the challenges are many: there is no valid open source dataset, pre-processing steps for the dataset are not easy, the number of publications in this area are not enough, and a lot of work is still necessary to reach the English text classification level. We hope that this systematic review will contribute to the research field and will constitute a foundation for other researchers to focus their research on.

## ACKNOWLEDGEMENTS

This work is a part of a project undertaken at the British University in Dubai.

## REFERENCES

- [1] S. H. P. Oo, et al., "Sentence Sentiment Classification Using Convolutional Neural Network in Myanmar Texts," in *Proceedings of the 2020 2nd International Conference on Image, Video and Signal Processing*, pp. 144-149, 2020.
- [2] P. Li, et al., "User-Guided Aspect Classification for Domain-Specific Texts," *arXiv Prepr. arXiv2004.14555*, 2020.
- [3] S. A. Salloum, et al., "The Impact of Knowledge Sharing on Information Systems: A Review," in *International Conference on Knowledge Management in Organizations*, pp. 94-106, 2018.
- [4] M. Alshurideh, et al., "Factors affecting the Social Networks Acceptance: An Empirical Study using PLS-SEM Approach," in *8th International Conference on Software and Computer Applications*, pp. 414-418, 2019.
- [5] S. A. Salloum, et al., "Analyzing the Arab Gulf Newspapers Using Text Mining Techniques," in *International Conference on Advanced Intelligent Systems and Informatics*, pp. 396-405, 2017.
- [6] S. A. Salloum and K. Shaalan, "Adoption of E-Book for University Students," in *International Conference on Advanced Intelligent Systems and Informatics*, pp. 481-494, 2018.
- [7] S. A. Salloum, et al., "Factors affecting the Adoption and Meaningful Use of Social Media: A Structural Equation Modeling Approach," *International Journal of Information Technology and Language Studies*, vol. 2, no. 3, pp. 96-109, 2018.
- [8] S. A. Salloum, et al., "Studying the Social Media Adoption by university students in the United Arab Emirates," *International Journal of Information Technology and Language Studies*, vol. 2, no. 3, pp. 83-95, 2018.
- [9] M. Alghizzawi, et al., "The role of social media in tourism marketing in Jordan," *International Journal of Information Technology and Language Studies*, vol. 2, no. 3, pp. 59-70, 2018.

- [10] M. Habes, et al., "The role of modern media technology in improving collaborative learning of students in Jordanian universities," *International Journal of Information Technology and Language Studies*, vol. 2, no. 3, pp. 71-82, 2018.
- [11] M. T. Alshurideh, et al., "Understanding the Quality Determinants that Influence the Intention to Use the Mobile Learning Platforms: A Practical Study," *International Journal of Interactive Mobile Technologies*, vol. 13, no. 11, pp. 157-183, 2019.
- [12] S. A. Salloum, et al., "Understanding the Impact of Social Media Practices on E-Learning Systems Acceptance," in *International Conference on Advanced Intelligent Systems and Informatics*, pp. 360-369, 2019.
- [13] M. Alshurideh, et al., "Examining the Main Mobile Learning System Drivers' Effects: A Mix Empirical Examination of Both the Expectation-Confirmation Model (ECM) and the Technology Acceptance Model (TAM)," *International Conference on Advanced Intelligent Systems and Informatics*, pp. 406-417, 2020.
- [14] R. S. Al-Marouf, et al., "A Unified Model for the Use and Acceptance of Stickers in Social Media Messaging," *International Conference on Advanced Intelligent Systems and Informatics*, pp. 370-381, 2020.
- [15] M. Habes, et al., "The Relation Between Social Media and Students' Academic Performance in Jordan: YouTube Perspective," *International Conference on Advanced Intelligent Systems and Informatics*, pp. 382-392, 2020.
- [16] M. Alghizzawi, et al., "The effect of social media usage on students'e-learning acceptance in higher education: A case study from the United Arab Emirates," *International Journal of Information Technology and Language Studies*, vol. 3, no. 3, pp. 13-26, 2019.
- [17] M. Alghizzawi, et al., "The Relationship Between Digital Media and Marketing Medical Tourism Destinations in Jordan: Facebook Perspective," in *International Conference on Advanced Intelligent Systems and Informatics*, pp. 438-448, 2019.
- [18] R. S. Al-Marouf, et al., "Understanding an Extension Technology Acceptance Model of Google Translation: A Multi-Cultural Study in United Arab Emirates," *International Journal of Interactive Mobile Technologies*, vol. 14, no. 03, pp. 157-178, 2020.
- [19] A. Y. Zainal, et al., "Dimensions of Agility Capabilities Organizational Competitiveness in Sustaining," in *The International Conference on Artificial Intelligence and Computer Vision*, pp. 762-772, 2020.
- [20] B. Al Kurdi, et al., "An Empirical Investigation into Examination of Factors Influencing University Students' Behavior towards Elearning Acceptance Using SEM Approach," *International Journal of Interactive Mobile Technologies*, vol. 14, no. 02, pp. 19-41, 2020.
- [21] S. A. Salloum, et al., "A survey of text mining in social media: facebook and twitter perspectives," *Advances in Science, Technology and Engineering Systems Journal*, vol. 2, no. 1, pp. 127-133, 2017.
- [22] C. Mhamdi, et al., "Text mining and analytics: A case study from news channels posts on Facebook," *Intelligent Natural Language Processing: Trends and Applications*, pp. 399-415, 2018.
- [23] S. A. Salloum, et al., "Analysis and Classification of Arabic Newspapers' Facebook Pages using Text Mining Techniques," *International Journal of Information Technology and Language Studies*, vol. 1, no. 2, pp. 8-17, 2017.
- [24] S. A. Salloum, et al., "Mining Text in News Channels: A Case Study from Facebook," *International Journal of Information Technology and Language Studies*, vol. 1, no. 1, pp. 1-9, 2017.
- [25] M. Habes, et al., "The Relationship between Social Media and Academic Performance: Facebook Perspective," *International Journal of Information Technology and Language Studies*, vol. 2, no. 1, pp. 12-18, 2018.
- [26] S. A. Salloum, et al., "A Survey of Lexical Functional Grammar in the Arabic Context," *International Journal of Computing and Network Technology*, vol. 4, no. 3, pp. 141-147, 2016.
- [27] W. Alabbas, et al., "Arabic text classification methods: Systematic literature review of primary studies," in *2016 4th IEEE International Colloquium on Information Science and Technology (CiSt)*, pp. 361-367, 2016.
- [28] S. A. Salloum, et al., "Using text mining techniques for extracting information from research articles," *Intelligent Natural Language Processing: Trends and Applications*, pp. 373-397, 2018.
- [29] S. A. Salloum, et al., "Mining Social Media Text: Extracting Knowledge from Facebook," *International Journal of Computing and Digital Systems*, vol. 6, no. 2, pp. 73-81, 2017.
- [30] R. Elhassan and M. Ahmed, "Arabic Text Classification review," *International Journal of Computer Science and Software Engineering*, vol. 4, no. 1, pp. 1-5, 2015.
- [31] L. Khreisat, "A machine learning approach for Arabic text classification using N-gram frequency statistics," *Journal of Informetrics*, vol. 3, no. 1, pp. 72-77, 2009.
- [32] K. B. Cohen, "Natural Language Processing for Online Applications: Text Retrieval, Extraction and Categorization (review)," *Language*, vol. 80, no. 1, pp. 178-178, 2004.
- [33] R. M. Duwairi, "Machine Learning for Arabic Text Categorization," *Journal of the American Society for Information Science and Technology*, vol. 57, no. 8, 2006.
- [34] M. A. R. Abdeen, et al., "A closer look at arabic text classification," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 11, pp. 677-688, 2019.
- [35] B. Zhong, et al., "Convolutional neural network : Deep learning-based classification of building quality problems," *Advanced Engineering Informatics*, vol. 40, pp. 46-57, 2019.
- [36] S. A. Salloum, et al., "Machine Learning and Deep Learning Techniques for Cybersecurity: A Review," in *The International Conference on Artificial Intelligence and Computer Vision*, pp. 50-57, 2020.
- [37] S. A. Salloum, et al., "A Survey of Semantic Analysis Approaches," in *The International Conference on Artificial Intelligence and Computer Vision*, pp. 61-70, 2020.
- [38] S. A. Salloum, et al., "Mining in Educational Data: Review and Future Directions," in *The International Conference on Artificial Intelligence and Computer Vision*, pp. 92-102, 2020.

- [39] S. F. S. Alhashmi, et al., "A Systematic Review of the Factors Affecting the Artificial Intelligence Implementation in the Health Care Sector," in *The International Conference on Artificial Intelligence and Computer Vision*, pp. 37-49, 2020.
- [40] F. Harrag, "Neural Network for Arabic Text Classification," *2009 Second International Conference on the Applications of Digital Information and Web Technologies*, pp. 778-783, 2009.
- [41] D. Svozil, "Introduction to multi-layer feed-forward neural networks," *Chemometrics and Intelligent Laboratory Systems*, vol. 39, no. 1, pp. 43-63, 1997.
- [42] D. E. Rumelhart, et al., "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533-536, 1986.
- [43] T. N. Sainath, et al., "Deep Convolutional Neural Networks for LVCSR," *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 8614-8618, 2013.
- [44] I. Sutskever, et al., "Generating text with recurrent neural networks," in *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, pp. 1017-1024, 2011.
- [45] S. Hochreiter and J.U. Schmidhuber, "Long shortterm memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997.
- [46] D. Lee, et al., "Long short-term memory recurrent neural network-based acoustic model using connectionist temporal classification on a large-scale training corpus," *China Communication*, vol. 14, no. 9, pp. 23-31, 2017.
- [47] M. Al-Smadi, et al., "Using long short-term memory deep neural networks for aspect-based sentiment analysis of Arabic reviews," *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 8, pp. 2163-2175, 2019.
- [48] G. A. Abandah, et al., "Automatic diacritization of Arabic text using recurrent neural networks," *Document Analysis and Recognition*, vol. 18, no. 2, pp. 183-197, 2015.
- [49] A. Alwehaibi and K. Roy, "Comparison of Pre-trained Word Vectors for Arabic Text Classification using Deep Learning Approach," *2018 17th IEEE International Conference on Machine Learning and Applications*, pp. 1471-1474, 2018.
- [50] A. Elnagar, et al., "An Annotated Huge Dataset for Standard and Colloquial Arabic Reviews for Subjective Sentiment Analysis," *Procedia Computer Science*, vol. 142, pp. 182-189, 2018.
- [51] S. Boukil, et al., "Arabic text classification using deep learning technics," *International Journal of Grid and Distributed Computing*, vol. 11, no. 9, pp. 103-114, 2018.
- [52] A. Elnagar, et al., "Automatic Text Tagging of Arabic News Articles Using Ensemble Deep Learning Models," *Proceedings of the 3rd International Conference on Neural Language and Speech Processing*, pp. 59-66, 2019.
- [53] A. Mahmoud and M. Zrigui, "Deep Neural Network Models for Paraphrased - Text Classification in the Arabic Language," *International Conference on Applications of Natural Language to Information Systems*, pp. 3-16, 2019.
- [54] F. Harrag, et al., "A Comparative Study of Neural Networks Architectures on Arabic Text Categorization Using ature Extraction," *International Conference on Machine and Web Intelligence*, pp. 102-107, 2010.
- [55] A. H. Mohammad, et al., "Arabic Text Categorization Using Support vector machine, Naïve Bayes and Neural Network," *GSTF Journal on Computing*, vol. 5, no. 1, pp. 108-115, 2016.
- [56] V. Jindal, "A personalized markov clustering and deep learning approach for Arabic text categorization," *Proceedings of the ACL 2016 Student Research Workshop*, pp. 145-151, 2016.
- [57] F. Z. El-Alami and S. O. El Alaoui, "An Efficient Method based on Deep Learning Approach for Arabic Text Categorization," *International Arab Conference on Information Technology*, pp. 1-7, 2016.

## BIOGRAPHIES OF AUTHORS



**Ahlam Wahdan** had graduated from University of Wollongon in Dubai with a distinction with MSc in Information Technology Managemet. She got her Bachelor's degree in Computer Engineering from AlBalqaa Applied University. Currently, she is working at the British University of Dubai Instructor in Computer Science Department. Her research Area is in Text mining, data mining, Machine learning algorithm and AI.



**Sendeyah Hantoobi**, has graduated from University of Wollongon in Dubai with Master in (Information Technology management). She got his Bachelor's from United Arab Emirates University in Computer Sciences. She is working in Ministry of Education in UAE as curriculum and learning management system specialists. Her research area is Text classification, organization, knowledge management, Big Data, AI and Machine Learning.



**Said A. Salloum** had graduated from The British University in Dubai with a distinction with MSc in Informatics (Knowledge and Data Management). He got his Bachelor's degree in Computer Science from Yarmouk University. Currently, He is working at the University of Sharjah "Research Institute of Sciences and Engineering (RISE)" as a researcher on different research areas in Computer Science such as data analysis, machine learning, knowledge management, and Arabic Language Processing. Salloum is an Oracle expert since 2013, along with various recognized international certificates that are issued by Oracle. You can contact. Said Salloum via emails at: [ssalloum@sharjah.ac.ae](mailto:ssalloum@sharjah.ac.ae) or [salloum78@live.com](mailto:salloum78@live.com) sa s11



**Dr. Khaled Shaalan** is a full professor of Computer Science/Artificial Intelligence at the British University in Dubai (BUiD), UAE. He is an Honorary Fellow at the School of Informatics, University of Edinburgh (UoE), UK. Over the last two decades, Prof Khaled has been contributing to a wide range of research topics in AI, Arabic NLP, Knowledge management, health informatics, and educational technology. Prof Khaled has published 200+ referred publications. Prof Khaled's research work is cited extensively worldwide and the impact of his research using GoogleScholar's H-index metric is 35+. Prof Khaled has been actively and extensively supporting the local and international academic community. He acts as the chair of international Conferences, journals and books editor, keynote speaker, external member of promotions committees, among others.