Quran Companion – A helping tool for Huffaz

Muhammad Rafi, Behraj Khan, Abdul Wahab Usmani, Zulqarnain Qasim, Shuja Ali, Syed Osama Hasan

***Abstract*— This paper presents work on a novel and innovative mobile application for Huffaz who find it difficult to maintain their Hifz in professional life. It is because they require a partner (Hafiz) who can listen to their recitation of Quran and correct it (if required). An application that listens Quran recitation and indicate mistakes (if any) with a feedback mechanism for their correction can solve the problem. A variety of methods related to Speech Recognition and Supervised Learning have been explored to develop such an invention. The way is to compare the input (speech) with standard (speech or text), outputting the status. However, the accuracy level obtained this way is not very reliable. Even so, it is inferable that this extraordinary application is a major step towards effective human-computer interaction. It is important to note that this experimentation is not concerned with teaching people how to recite Quran but with helping people (who already know how to recite Quran) to memorize and revise the Quran in a better way.**

***Keywords*— Speech Recognition; Mobile Application; Huffaz; Supervised Learning.**

I. INTRODUCTION

Speech is the most convenient form of communication and Learning is a parameter for intelligent machines. Speech is the ability and Learning is a method to gain an ability. They are part of any normal human's behavior. They help people understand each other better. But it is not the same with machines as they are unable to speak and learn. This has been an area of great interest for scientist that how improve human's interaction with machines.

In this research, an application is being discussed that can listen the recitation of Holy Quran, detect mistakes (if any) and help humans in correcting their mistakes. In other words, the application has learned the Quran and it is now able to tell people about their mistakes. In order to develop such an advanced application, there are two methods that have been discussed.

Mohummad Rafi is with Computer Science Department FAST-National University Karachi, Pakistan. (Email: muhammad.rafi@nu.edu.pk)
Behraj Khan is with Computer Science Department FAST-National University Karachi, Pakistan. (Email: behraj.khan@nu.edu.pk)
Abdul Wahab Usmani is with Computer Science Department FAST-National University Karachi, Pakistan. (Email: k152159@nu.edu.pk)
Zulqarnain Qasim is with Computer Science Department FAST-National University Karachi, Pakistan. (Email: k152160@nu.edu.pk)
Shuja Ali is with Computer Science Department FAST-National University Karachi, Pakistan. (Email: k152424@nu.edu.pk)
Syed Osama Hqsan is with Computer Science Department FAST-National University Karachi, Pakistan. (Email: k153632@nu.edu.pk)

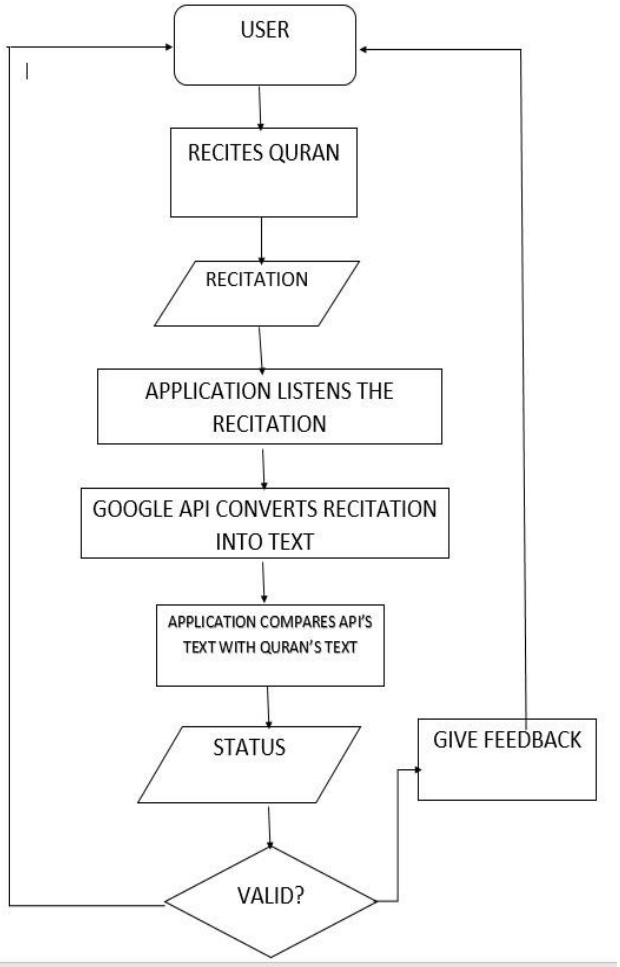

Fig 1 Model of the system.

First method is Speech Recognition. An application listens the recitation of Quran and converts it into text and then compares the text with the original text of Quran to discover its validity. It also gives feedback via text and voice and correct people's recitation of Holy Quran.

Second approach is Supervised Learning. In this case, application compares the audio of recitation with the audios of correct recitation and calculate similarity. A threshold is set to label the recordings 'incorrect' with similarity less than threshold and 'correct' otherwise.

It is always a good idea to investigate about the related work. The upcoming part of the paper provides a detailed analysis of the study that is concerned with the similar idea or approach.

Table 1 Collected Data

No	Applications Name	Features
1	Analyze Quran	Recitation, Translation, Search, Listen Quran, Guides, Sharing,
2	Quran.com	Recitation, Translation, Search, Listen Quran, Salah Times, Sunnah
3	Alim Mobile App	Recitation, Translation, Tafsir, Guides,Themes Salah Times, Sunnah
4	Zekr	Recitation, Translation, Search, Listen, Quran, Tafsir, Navigation, Multiple, languages, Themes, Bookmark
5	Ayat	Recitation, Translation, Search, Listen, Quran, Navigation, Multiple languages, Memorizing Test, Guides, Repetition
6	Tanzil	Recitation, Translation, Search, Listen, Quran, Tafsir, Navigation, Multiple, languages, Sharing, Text Setting
7	UQuran (iQuran)	Recitation, Translation, Listen Quran,Navigation
8	QuranRecitor	Recitation, Translation, Listen Quran,Multiple languages, Repetition
9	Openburhan	Recitation, Translation, Search, Listen, Quran, Tafsir, Navigation
10	Al Quran MP3-	Recitation, Translation, Search, Listen, Quran Reading Quran, Navigation, Multiple languages,Themes, Sharing, Repetition, Sunnah
11	Quran Majeed	Recitation, Translation, Search, Listen, Quran, Navigation, Multiple languages, Memorizing Test, Guides, Themes, Sharing, Text Setting, Repetition, Sunnah
12	Learn Quran	Recitation, Translation, Search, Navigation, Tajweed Multiple languages, Memorizing Test, Guides, Themes, Text Setting, Repetition
13	Deeniyat	Recitation, Translation, Search, Listen, Quran, Tafsir, Navigation, Themes, Salah, Times, Qibla direction, Bookmark, Text, Setting, Repetition, Sunnah
14	Quran for Android	Recitation, Translation, Search, Listen, Quran, Multiple languages, Themes, Salah, Times, Qibla direction, Text Setting, Repetition
15	AL Quran MP3	Recitation, Translation, Search, Listen, Quran, Tafsir, Repetition, Sunnah
16	Islam360	Recitation, Translation, Search, Listen, Quran, Tafsir, Navigation, Multiple, languages, Memorizing Test, Guides, Themes, Sharing Salah Times, Qibla, direction, Bookmark, Text Setting,, Repetition, Sunnah
17	Quran tutor	Recitation, Translation, Listen Quran, Tafsir, Memorizing Test, Guides, Repetition, Sunnah, Tutoring
18	Quran Companion	Recitation, Translation, Listen Quran, Multiple languages, Memorizing Test
19	Iquran Lite	Recitation, Translation, Search, Navigation,Bookmark, Repetition
20	Quranic-Learn	Recitation, Translation, Guides, Bookmark Quran, Learn Islam
21	Holy Quran	Recitation, Search, Bookmark
22	Holy Quran (Read Free)	Recitation, Bookmark
23	Al Quran-Ul-Kareem	Recitation, Search, Listen Quran, Bookmark
24	Listen Quran	Recitation, Translation, Search, Listen Quran Offline
25	Color Quran 30	Recitation, Search, Tafsir Juz Offline Read

II. BACKGROUND

1. Survey

zIn order to solve the problem regarding the need of partner to memorize and revise Quran, the 23 Quran applications have been checked to see if they provide a reasonable solution for this problem. The data collected in the survey is shown in Table 1.

It can be seen in the table above that there is not a single feature in any application that fills the gap of a partner for Hafiz (plural:Huffaz). Therefore, this research was initiated.

II. Literature Review

2.1. Speech Recognition

Speech recognition is the associative subfield of computational linguistics that develops methodologies and technologies that enables the recognition and translation of spoken language into text by computers [1]. It is also referred to as Automatic Speech Recognition (ASR) or simply Speech-to-text (STT) [1]. Model of the system is shown in fig 1.

ASR (as shown in Figure 2) is the ability of a computer to convert the words spoken by a person into text [1]. ASR represents the function of modulating a speech signal to a series of words with the assistance of an algorithm performed by a computer program [2].

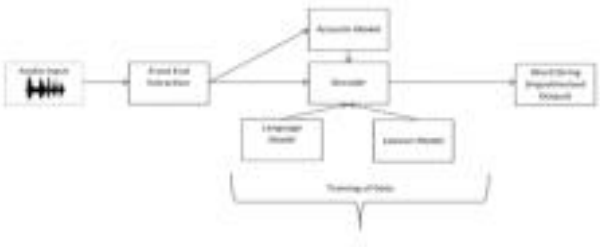

Figure 2. Block diagram for Automatic Speech Recognition [1]

2.1.1. Integrating Speech Recognition with Android

2.1.1.1 Text to Speech (TTS)

TTS is a Google API that is used in making of the application. With the help of this the textual data can be converted into sound waves. There are many existing languages that can be

converted from text to speech such as Arabic, English, French, German, Spanish etc.

2.1.1.2 Speech to Text (STT)

STT is another name for automatic speech recognition. This API is used for the more difficult part of converting the audio word into a textual string. Large sets of data need to be trained so that the system can recognize the words being spoken.

2.1.2. Holy Quran Speech recognition

The act of learning and teaching of the Holy Quran has become a scientific practice to Muslims around. Several efforts were made by previous systems on the development of feasible guiding techniques to the act of Tajweed [3]. A research in represents a speech recognition system that distinguishes the types of Madd (elongated tone) or prolongation and the type of Qira'at (method of recitation) related to Madd [3].

2.1.3. Neural Network for Arabic Speech Recognition

Boltzmann machine neural network has been used to recognize the Arabic Speech in which Fast Fourier transformation algorithm has been used to extract spectral features from an acoustic signal [4].

2.2. Supervised Learning

Supervised learning is the Data mining task of inferring a function from labeled training data. Training data consist of a set of training examples. In supervised learning, each example is a pair consisting of an input object (typically a vector) and a desired output value (also called the supervisory signal).

Supervised learning entails learning a mapping between a set of input variables X and an output variable Y and applying this mapping to predict the outputs for unseen data

[5]. Semi supervised learning used in Big Social Data Analysis to provide a learning model based on focus projection and vector machines for emotion recognition and polarity detection [6].

2.2.1. Mobile application to detect accidents

A mobile application was created that can distinguish between normal activities and accidents by using Supervised Learning and if detects an accident then it informs appropriate authorities [7].

2.3. Holy Quran digitization

A survey was conducted that includes 17 questions related to topic "Holy Quran Digitization: Challenges and Concern" and about 668 responses were recorded from different parts of the world and was found that the youth uses more smart technologies than any other age group related to Holy Quran [8]. This leads to make secure Quran applications as understanding the threats and vulnerabilities of mobile Quran applications and the way to manage them is crucial. Any attempt that could compromise the originality of Quran must be identified [9].

It is required to continuously update all the Quran applications in order to minimize the security risks that may affect the originality of the Quran available in the applications. Security can be achieved by following the procedures as it is for the all mobile applications [10]. But there lies a gap between the Quran mobile applications and other mobile applications which can be covered by updating the Quran applications rapidly in order to make them be compete with other application available on Play Store [11].

A potential mobile application has been identified for Huffaz with which they can easily revise the Quran which they have hifz a period ago. They would also be notified by the application if they make any mistake while reciting the Holy Quran. At present there is no such mobile application available which can listen the recitation of Holy Quran and point out the mistakes if any in the recitation made by the reciter and this is the main reason for which we have designed an android mobile application which overcomes the gap that in between the Quran mobile applications and the users' requirement.

III. RESEARCH METHODOLOGY

In first approach, an android mobile application was designed to record the audio, play the recorded audio, and pause the played recording. It also had a text field to note down the mistakes in the recording of recitation of Holy Quran. Filing was used to save mistakes against the name of recording in which they were made. There are different kinds of mistakes that may occur in recitation of Quran but only "Skipping of verse/word" was dealt as error in this approach i.e. you recite first verse of a chapter in Quran, skip the second verse and start reciting the third verse. Thirty recordings of first two parts (rukus) of chapter "Surah Yaseen" were collected for this purpose.

In time, it was realized that this approach is not suitable as there are various types of mistakes made by people in recitation of Quran and one more research will just be needed to identify and categorize these mistakes. This approach was to be accomplished by Supervised Learning.

Secondly Speech Recognition was used to achieve the goal. For this an android mobile application was developed which uses the Google Speech to Text API, converting recited audio by into Arabic text. Then this text was compared with original Quran text to detect mistakes, if any. The problem occurred in this approach was that the Arabic text of API was different from the Arabic text of Quran in Indo-Pak. The example of this is the Arabic text of a word "Malik" provided by API is "كلام" but in Quran Arabic text it is written as "كٰلِـم". In order to solve issue the API's Arabic text was mapped to Quran

Arabic text to display the text the way it is written in Quran.

We revisited Supervised Learning to apply number of techniques to compare two audios. More than thirty recordings of "Surah Kawsar" were collected and each verse was separated from recordings. Please note that these recordings were noise-free. Then recordings of same verse and of different verses were compared in following ways:

1. Energy Spectral Density to compare two audi
2. Percentage Difference between two audios
3. Cosine Similarity
4. Matching pattern of positive and negative values of two audios
5. Euclidean Distance
6. Root Mean Square Error

None of the above techniques' results were according to

our expectations as we were expecting that the audios with the same voice should have more similarity as compared to the audios with different voice, when compared with each other.

Word by word comparison was also tried by featuring out in how many ways we can recite a word of any Surah.

For this we have only gathered the first word of 'Surah Takasur' and tried to find out the sound properties after

pronouncing in different correct versions using MATLAB Signal processing toolbox. But this study lacks because we will not be aware about the feelings of the reciter and reciter may recite it in a much different way then we have expected.

Thus, this Supervised Learning approach was not successful with respect to the experimentations performed using above techniques.

1. Concepts and Tools

As far as the concepts of project are concerned, there are three main concepts: Supervised Learning, Signal Processing and Speech Recognition.

Supervised Learning is a machine learning technique of learning a method or function that maps an input to an output based on example input-output pairs. It basically derives a function from labeled training data (containing a set of training examples). Similarly, Speech Recognition is another concept that deals with converting speech into text. It basically records our voice (speech) and after analyzing it, shows in the form of text whatever we have said.

Signal Processing is just simply a way of dealing with signals and manipulating them. Since we have use of voice (sound) in our project so we deal with sound signals. Signal Processing includes many functions like reading a signal, plotting a signal, filtering the signal, changing the frequency of signal and many other operations can be performed on signals.

Now the concepts have been explained, but the question remains that how are we going to apply these concepts in our project. For that tools are needed. The tools used are MATLAB and Google Speech to Text API and Android Studio. The tool used for Signal Processing and Supervised Learning is MATLAB and Google API plays its role in Speech Recognition.

Android Studio is used to build the android mobile application that can record the recitation of reciter and correct the mistakes in recitation.

IV. IMPLEMENTATION

There are two approaches which are under consideration for this project. First approach uses Speech Recognition and String Matching/Searching Algorithm. The second approach is concerned with Signal Processing and Supervised Learning.

Following the first approach an android application was created that compared the converted text with the original text of Quran. No string matching/search algorithm was applied two strings were just simply compared using an equal function available in java. If the match was success then it will tell that it was successful and if not then it would tell that it was not successful. Later, an algorithm provided by libraries in JAVA like "java.util.regex.Matcher" and "java.util.regex.Pattern" was used to obtain better results.

In case of second approach, data (30 recordings of first two rukus of Surah Yaseen) was collected, an android app was created that has the voice recorder and could save the recorded voice in a file. There was an "editText" also available to note down mistakes during the recording and save them and retrieve them.

In different models that were applied during experimentation, there was a training data of mostly fifty-five recordings. The models (except model VII) have been discussed below with test data of five recordings out of which three are correct and two are incorrect and they are of different people.

1. Model I

There is a 'musicg' library in Java which gives similarity between two audio files on the basis of

Energy Spectral Density. Therefore, experiment was performed and following results were obtained.

1.1. Results

Serial #	Test Recording	Similarity	Expected

1	Correct 1	2.09	>5.63
2	Correct 2	2.94	>5.63
3	Correct 3	2.68	>5.63
4	Incorrect 1	5.63	<2.09
5	Incorrect 2	2.95	<2.09

1.2. Discussion

It can be seen that the similarity of incorrect recordings is more than correct recordings. It was deduced from the above results that speech-based comparison was required but voice-based comparison was being done. And by voice, it means that different characteristics (like frequency, amplitude, time period etc.) are being compared not the content(speech). Therefore, the similarity is being altered due to the involvement of characteristics other than speech characteristics.

2. Model II

In this case, two audio files were read and their matrix was generated. The Percentage difference of two obtained matrices was calculated which was the dissimilarity. And then dissimilarity was subtracted from total (100 or 1) and thus similarity was obtained.

2.1. Results

Serial #	Test Recording	Similarity	Expected
1	Correct 1	39.50	>72.64
2	Correct 2	29.40	>72.64
3	Correct 3	36.46	>72.64
4	Incorrect 1	72.64	<29.40
5	Incorrect 2	48.45	<29.40

2.2. Discussion

It can be seen that the similarity of incorrect recordings is more than correct recordings. It was deduced from the above results that speech-based comparison was required but voice-based comparison was being done. And by voice, it means that different characteristics (like frequency, amplitude, time period etc.) are being compared not the content(speech). Therefore, the similarity is being altered due to the involvement of characteristics other than speech characteristics.

3. Model III

In this case also, two audio files were read and their matrix was generated. Their matrices were converted into vectors (column matrix) and the sign (+ or -) of corresponding values in vectors was compared and similarity was calculated.

3.1. Results

Serial #	Test Recording	Similarity	Expected
1	Correct 1	34.98	>47.64
2	Correct 2	42.78	>47.64
3	Correct 3	22.39	>47.64
4	Incorrect 1	47.64	<22.39
5	Incorrect 2	45.77	<22.39

3.2. Discussion

It can be seen that the similarity of incorrect recordings is more than correct recordings. It was deduced from the above results that speech-based comparison was required but voice-based comparison was being done. And by voice, it means that different characteristics (like frequency, amplitude, time period etc.) are being compared not the content(speech). Therefore, the similarity is being altered due to the involvement of characteristics other than speech characteristics.

4. Model IV

In this case again, two audio files were read and their matrix was generated. Their matrices were converted into vectors (column matrix) and cosine similarity was calculated between the two.

4.1. Results

Serial #	Test Recording	Similarity	Expected
1	Correct 1	22.90	>60.56
2	Correct 2	51.53	>60.56
3	Correct 3	46.50	>60.56
4	Incorrect 1	60.56	<22.90
5	Incorrect 2	55.96	<22.90

4.2. Discussion

It can be seen that the similarity of incorrect recordings is more than correct recordings. It was deduced from the above results that speech-based comparison was required but voice-based comparison was being done. And by voice, it means that different characteristics (like frequency, amplitude, time period etc.) are being compared not the content(speech). Therefore, the similarity is being altered due to the involvement of characteristics other than speech characteristics.

5. Model V

In this case again, two audio files were read and their matrix was generated. The Euclidean distance between corresponding values of the two obtained matrices was calculated and thus dissimilarity was calculated.

5.1. Results

Serial #	Test Recording	Dissimilarity	Expected
1	Correct 1	4.93e+05	<1.72e+05
2	Correct 2	5.18e+04	<1.72e+05
3	Correct 3	2.12e+05	<1.72e+05

4	Incorrect 1	1.72e+05	>4.93e+05
5	Incorrect 2	2.84e+05	>4.93e+05

5.2. Discussion

It can be seen that the dissimilarity of correct recordings is more than incorrect recordings. It was deduced from the above results that speech-based comparison was required but voice-based comparison was being done. And by voice, it means that different characteristics (like frequency, amplitude, time period etc.) are being compared not the content(speech). Therefore, the similarity is being altered due to the involvement of characteristics other than speech characteristics.

6. Model VI

In this case, the two audio files were divided into two smaller files each containing a part of the original file. Then these smaller files were read in the form of matrix and were subtracted from each other and then similarity was calculated based on percentage difference. The threshold was set to 81. Greater than 81 was correct and vice versa.

6.1. Results

Serial #	Test Recording	Similarity	Expected
1	Correct 1	85.1953	>81
2	Correct 2	87.3703	>81
3	Correct 3	88.1865	>81
4	Incorrect 1	87.1849	<81
5	Incorrect 2	83.6515	<81

6.2. Discussion

It can be seen that the similarity of incorrect recordings is more than correct recordings. It was deduced from the above results that speech-based comparison was required but voice-based comparison was being done. And by voice, it means that different characteristics (like frequency, amplitude, time period etc.) are being compared not the content(speech). Therefore, the similarity is being altered due to the involvement of characteristics other than speech characteristics.

7. Model VII

Lastly, Google Speech Recognition API was tried. It converted Quran recitations (speech) into text, which is the content. Thus, the main problem of extracting content from voice was achieved. Now the text obtained from API was compared with text of Quran to indicate mistakes (if any). Feedback mechanism via text and voice was also implemented for correction of mistakes.

7.1. Results

Test Group	Group I	Group II	Group III
Members	100%	Above 50 %	Hardly 5 %
Accuracy	23	4	3

7.2. Discussion

In this model, there were three groups that were identified. Thirty people were approached to test this model. Twenty-three became the members of Group I with 100% accuracy. Four became the members of Group II with accuracy above 50%. Three became the members of Group III with accuracy hardly 5%. 100% means the model worked as it was supposed to (i.e. it labelled correct recitation correct and vice versa). 50% means the model extracted the content that was partially incorrect. But then it was tried again and again and soon accuracy was 100%. Similarly, in case of Group III, the content obtained from API was either not correct at all or very little correct. Thus, in this approach the main problem occurring is correct extraction of content or text from speech.

For comparison, pattern matching/searching algorithm was used, provided by libraries of Java like 'java.util.regex. Matcher' and 'java.util.regex.Pattern'. The algorithm can search for the right content wherever it is in the given text. It can find the pattern twice if it is occurring twice in the given text. This model overall has been a successful model to a great extent.

V. FINDINGS

The findings of research include:

1. Supervised Learning for this particular problem is not very effective, as it requires a lot of data and proper training mechanism.
2. Speech processing and recognition approach is also very subtle for this problem as most innovative and popular Google Speech API do not support it very well. It is very limited to find pronunciation and its variation.
3. Text comparison for speech-to-text and simple text in Arabic is also of limited use in this context of real time analysis.

VI. FUTURE WORK

It is recommended that in future, an approach of Deep Learning and Neural Networks be applied with a dataset that is at least in thousands for Quran Recognition that is free from the flaws mentioned in this paper.

VII. CONCLUSION

It is inferable that such an application will be a great innovation in this period of time and a remarkable contribution in the field of research. Not only it will help Huffaz community to revise Quran wherever they feel ease using this mobile application but also Non-Huffaz community to memorize Quran. It would not be wrong to say that this research and it's resulting mobile application have a huge market in entire world.

REFERENCES

[1] Viswanath Subramanian Ramesh*, P. Saranya and Ruchira Kapoor "Speech Recognition based Adaptive Examination Application for Visually Impaired Students." Indian Journal of Science and Technology, Vol 9(37), DOI: 10.17485/ijst/2016/v9i37/102102, October 2016

[2] Haridas, Arul Valiyavalappil, Ramalatha Marimuthu, and Vaazi Gangadharan Sivakumar. "A critical review and analysis on techniques of speech recognition: The road ahead." International Journal of Knowledge-based and Intelligent Engineering Systems 22.1 (2018): 39-57

[3] Yousfi, Bilal, and Akram M. Zeki. "Holy Qur'an speech recognition system distinguishing the type of recitation." 2016 7th International Conference on Computer Science and Information Technology (CSIT). IEEE, 2016

[4] Mohamed, H. R. "Arabic Speech Recognition.

[5] Cunningham, Pádraig, Matthieu Cord, and Sarah Jane Delany. "Supervised learning." Machine learning techniques for multimedia. Springer, Berlin, Heidelberg, 2008. 21-49

[6] Hussain, Amir, and Erik Cambria. "Semi-supervised learning for big social data analysis." Neurocomputing 275 (2018): 1662-1673

[7] Kansiz, A. Oguz, M. Amac Guvensan, and H. Irem Turkmen. "Selection of time-domain features for fall detection based on supervised learning." Proceedings of the World Congress on Engineering and Computer Science, San Francisco, CA, USA. Vol. 2325. 2013.

[8] Khan, Muhammad Khurram, and Yasser M. Alginahi. "The holy Quran digitization: Challenges and concerns." Life Science Journal 10.2 (2013): 156-164.

[9] Talib, Shuhaili, et al. "Mobile Quran app security vulnerabilities." Proc. 5th Int. Conf. Comput. Informatics, ICOCI. No. 198. 2015

[10] Mukherjee, Shubhankar, Jyoti Prakash, and Deepak Kumar. "Android application development & its security." Int. J. Comput. Sci. Mobile Comput. 4.3 (2015): 714-719

[11] Zahari, Nuril Ham Al Hafizah Binti, Sharifah Norshah Bani Binti Syed Bidin, and Syadiah Nor Binti Wan Syamsuddin. "Development of Al-Quran Android Application from Year 2013 To 2016: The Highlight." International Journal of Academic Research in Business and Social Sciences 7.6 (2017): 183-195.