Journal of Theoretical and Applied Information Technology
31st January 2021. Vol.99. No 2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

QCM: QURANIC CITATIONS MANAGER BASED ON NATURAL LANGUAGE PROCESSING TECHNIQUES

[1]KHALID M.O. NAHAR

[1]Department of Computer Sciences,

Faculty of Information Technology and Computer Sciences, Yarmouk University, Irbid-21163, Jordan

ABSTRACT

Quran is the holy book for all Muslims. It is containing only truths, no suppositions, or uncertain information. In a variety of scientific areas, writers introduce witnesses for their claims or explain their dialogs by citing verses of the Holy Quran. A novel Quranic Citation Manager (QCM) is introduced in this research paper for automatically managing citations from the Holy Quran. The adopted scientific approach is based on Natural Language Processing techniques. The adopted QCM is based on a sliding window technique, and for more accurate Quranic search a context-sensitive error detection and correction are utilized. The citation manager for searching and error correction techniques were employed for referencing unreferenced in-text quotations from the Holy Quran and for preserving the style of a Quranic citation. QCM checks if the text current text is a Holy Quran text, then checks if it is correct or not, and finally if not correct, it will correct it. Experiments reveal that our adopted approach or QCM achieves a 93.95% accuracy rate.

Keywords: *Quranic Citations Manager; Context-Sensitive Error; Quran Searching; Lexically Neighboring Words.*

1. INTRODUCTION

The sacred text is that the holy book for quite one billion Muslims around the world speaking totally different languages.

Arabic is that the linguistic communication of the sacred text. Although, some people will scan and perceive the sacred text within the native language, they are unable to translate it properly. Although the number of the Quran's looking and translation websites is increasing [1],[20], and [21], a number of these websites square measure either incorrect or incomplete.

Thus, the requirement for a reliable translation service has become Associate in Nursing imperative necessity, particularly with the variety of Muslims' languages round the world. However, the interpretation of sacred text is translated from the initial Arabic text. Thus, if the original texts square measure corrupted then its translation are going to be incorrect. Consequently, the Quranic quote verification service became more necessary. Muslims rely on Quranic verses to support several decisions in their lifestyle. Decision-making is very dependent on the credibleness of the verses quoted from the Quran. Some Muslims use Quranic verses to deduce solutions for their social and non-secular issues or to research some events. On the opposite hand, most of the Muslim authors quote from these verses as Associate in Nursing proof to support their conclusions and their analysis of events. This methodology is standard in the monotheism societies normally and Arabs above all [2]. Thus, Muslims' reliance on sacred text makes it imperative for researchers to develop verification mechanisms to support Muslim's users. Ordinarily, readers with moderate spiritual datum cannot observe the credibleness of the quotes employed by the author. However, to verify the quote credibleness, they need to make an in-depth analysis within the original copy of the sacred text. Obviously, this can be a tedious task, particularly once that the verses enumeration is missing for the quotation. scrutiny the quote with the initial text and confirming the match emphasizes additional the credibleness of the Quranic quote. This procedure will increase the user's confidence on those Quranic quotes. Previous connected work includes the Quranic met codification performed at the International monotheism University in Malaysia [3]. The authors have mentioned the Quran's structure in terms of the quantity of verses within the chapter, the number of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

characters within the verses, and so on, using "Atomization Structure" and Unicode codification, which aims to protect the digital type of the sacred text from corruption. Previous search strategies square measure centered on the text info retrieval [1], linguistics search [4], [5], synchronic linguistics dependency [3]. However, none of the previous connected works have taken into consideration the Quranic basic text credibleness. This work aims to verify the basic text for specific Quranic quotes that ends up in improve the users' confidence in the digital content. during this paper, we are going to illustrate our authentication model that consists of 2 stages to validate Quranic quotes in digital forms.

Holy Quran is the literal word of Allah (God) for all people throughout the world, its revelation to the Prophet Mohammad (Peace Be Upon Him). Quran is the holy book for Muslims which covers all aspects of life and regulates all human actions through a wide set of commandments and prohibitions. In other words, the Quran forms a life map. This book contains only truths, no place for suppositions or uncertainty.

The language of the Holy Quran is Arabic. Holy Quran is divided into 114 chapters (Surah), which are then divided into verses (Ayah). Quranic words are precise and subtle, even interchanging a word by one of its synonyms will change the contextual meaning of the verse. In a variety of scientific fields, writers cite verses of the Holy Quran. There is a specific conventional style for Quranic citation, it should begin with Allah says (*قال تعالى*) followed by quotation marks which contain Quranic words and ends with chapter (Surah) name and verse (Ayah) number as follow:

قال تعالى *"ذَلِكَ الكِتَابُ لَا رَيْبَ فِيهِ هُدًى لِلْمُتَّقِينَ"* **(سورة البقرة ، الآية 2)**

Allah says: "*This is the Scripture whereof there is no doubt, a guidance unto those who ward off (evil)*" **(Chapter Al-Baqarah, Verse 2)**

Every natural language is governed by a set of rules to ensure the integrity of a written or typed sentence. Such rules cover two aspects. The first one is the GRAMMAR of language (Morphology and Syntax) which governs the use of the right morpheme at the right place. The second one is MEANING (Semantic and Pragmatic) in which assuring the use of a meaningful structure [1].

In regular text, syntax errors are known as nonword errors, while semantic errors are known as real-world errors or context-sensitive errors [2]. The task of detecting and correcting such errors is known as spell-checking. The problem of checking nonword errors is not in the scope of this work, it is an embedded task in most text editors such as MS_WORD.

The semantic error could be a semantic inconstancy or a semantic incompleteness. Semantic inconstancies error occurs when a word in a sentence is replaced by another word that has a close pronunciation or a close orthographic form, while semantic incompleteness error is caused by the mislaid of any particle or coordinating conjunction [3].

Context-sensitive errors detection and correction is the task of identifying grammatically valid words but out of the context and replacing erroneous word by another word in which it fits the context more than the erroneous word. Reference [4] classified the causes of context-sensitive errors into three classes, **typos** such as substituting (mate for made), **homonym confusions** such as substituting (desert for dessert), and **usage** errors such as substituting (into for through).

The Arabic language is considered a semantically rich language; it contains more than 12 million vocabularies. Subject, verb, and object may be included in one word such as "سيطعمونكم" (They will feed you), the previous Arabic word turned into a complete sentence when translated into English. Consequently, the Arabic language is considered a complex script [5].

All Muslims consider the Quran as a holy book, so any Quranic quotation should be treated carefully, mistakes are not allowed. Such citation mistakes may be made by writers as including Quranic text without quotation marks, without starting the quotation by (Allah Says), or without ending the quotation by Chapter and Verse reference. Also, context-sensitive error may affect the whole contextual meaning.

In this work, we propose a model for Quranic Citation Management, Quranic Citations Manager (**QCM1**) version 1.0. To preserve the style and content of the citations from the Holy Quran we adopt techniques for searching the holy Quran and techniques for context-sensitive errors detection and correction. The proposed model is composed of three main modules: Search Module, Error Correction Module, and Citation Module. The proposed model combining numerous methods; the search module employs a window sliding technique and exact text-matching techniques for matching the cited text with the Quran database. The error correction module utilizes a hybrid lexical disambiguation method. This method employs the context word method and lexical neighboring words technique. The citation module has two main tasks,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

detecting & correcting wrong citations and inserting citations it is not included.

The rest of the paper will be organized as follow: Section II reviews the literature. We present the architecture of our adopted model in section III. After that, the experimental results are presented in section IV. Finally, section V presents conclusions and future directions.

2. RELATED WORKS

Most of the reviewed research are partially related to our work, we classify them into two categories according to their relationship to our research. First, research on the techniques for searching Quran, and second, research on Arabic context-sensitive spell errors detection and correction.

A lot of research were approached techniques for facilitating the search in the Holy Quran. Reference [6] showed a classification of the techniques for searching the Holy Quran into Semantic-based search and Text-based search. Semantic-based search techniques were adopted to improve the search efficiency and related verses retrieval (i.e.) word matching is not a necessary condition in such techniques, which is contrary to our searching technique which involves a text-based search to detect or correct a cited verse (Ayah "آية").

Some approaches utilize information retrieval methods for searching the Holy Quran, for instance, reference [7] provides an approach of searching the Quran based on query expansion techniques such as the use of a thesaurus and stemming. Reference [8] introduced a Quran Search System which is mainly based on a keyword matching approach, it is utilizing stemming-based, text-based, and synonyms-based searching techniques in analyzing both Quran text and query text.

Searching the Holy Quran using semantic methods is more accurate than text-based (or keyword-based) searching techniques since the user's query doesn't always have an exact expression of what s/he inquires for. Reference [9] proposed a Quranic semantic search system based on ontology technique, the proposed system retrieved 95% of the related verses to a user's query. A semantic framework for the Quranic search engine was proposed in [10], NLP techniques were used to improve the retrieval efficiency such as part-of-speech tagging and morphological analysis beside Holy Quran concept extraction based on ontology. Another approach in searching the Holy Quran based on ontology technique and knowledge-extraction was proposed in [11], [22], [23], and [24].

The scope of semantic error detection and correction in our work is narrower than in the reviewed literature since the Quran is a limited corpus and the sequence of its words is already known. So, the prediction of a context-dependent word doesn't require complex statistical computations.

Approaches of Arabic context-sensitive spell errors detection and correction were mainly based on two perceptions. The first, lexical disambiguation that is uses predefined sets known as dictionaries, confusion sets, or lexical neighboring words. The second perception encompasses approaches that treat the problem based on language models in which using semantic and statistical information extracted from the surrounding context of the erroneous word [3]. Some approaches treat the problem by combining both perceptions such as [12], [13], [14], and [15].

Reference [16] tackle the problem through the collection of confusion sets from the dictionary (words of one letter disparity) besides collecting homophonic (same pronunciation and different meaning) words. For real-word error detection and correction, a combination of Context Word Method (CWM) and N-gram Language model had been used.

Reference [17] illustrates the development of a tool for spell detection and correction named El-Mossahih, the proposed approach combined dictionary, language model, permutation, and neighborhood.

3. METHODOLOGY

A. General Architecture

QCM is based on a multitask model for automatically handling Quranic citations. Those tasks are Searching in the Quranic text, Detection & Correction of context-sensitive errors in a citation text and insert a citation reference. The proposed model is composed of two components, QCM Resources, and QCM. The first component encompasses constructed and organized Quranic database files to facilitate the work of the second component. QCM component is structured into three modules Search Module, Error Correction Module, and Citation Module, which are combined in which performing the management role for Quranic citations. Figure 1 illustrates the adopted model design, which represents the QCM architecture.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

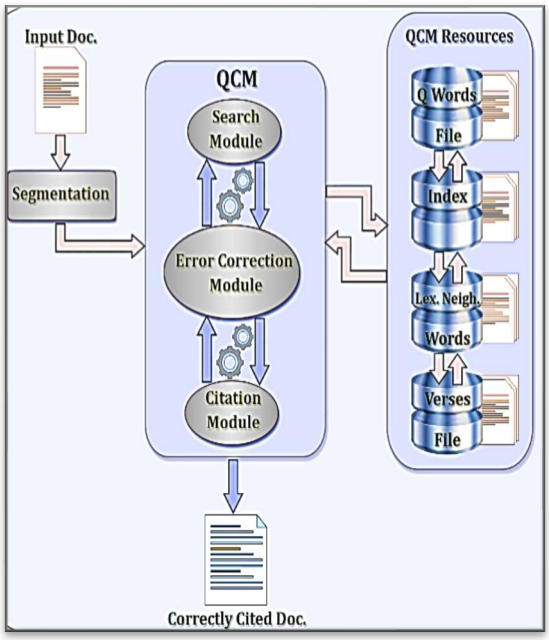

Fig. 1. Qcm General Framework

B. QCM Resources

The Quran corpus was obtained from the TANZIL website [18]. The corpus consists of one text file which includes all Quran verses. A preprocessing phase was performed over the Quran text to construct and organize the resource files required by the system.

QCM Resources were constructed and organized to facilitate and expedite the tasks of QCM modules. Resources are structured as follow:

- Quranic Words file: it contains all words of the Quran after removing duplications. Each record includes chapter numbers and verses' numbers that the word appears in. Also includes a reference to preceding and succeeding words as mentioned in the Quran.
- Index file: the Quranic Words file is indexed depending on two criteria, word length and alphabetical order. Since the longest word in the Quran consists of 11 letters, and the shortest could be of one letter. So, the index file is layered into 11 layers, where the number of layers indicates words length in such layer. Words in each layer are sorted in alphabetical order.
- Lexically neighboring words file: this file is utilized by the QCM component for context-sensitive errors detection and correction. It consists of all lexically neighboring words for each unique Quranic word (if there is any). Lexically neighboring words are words in which they have a close orthographic form to each other. To build up this file, we adopt a method for automatically generating lexically neighboring words for each unique word in the Quran corpus (unique Quranic words are obtained by removing duplicates). The adopted method was proposed in [3], it involves applying four editing operation on a Quranic word, appending one letter to the word, replacing one letter by another, removing a letter, and/or switching two contiguous letters. Resulting words are filtered by excluding nonwords (words with erroneous orthographic form).
- Verse file: contains all Quran verses with diacritic marks and verses references. This file is used to retrieve suggested verses by the system with its references.

To facilitate search in those files, a searching technique is adopted which is based on the exact string-matching approach and window sliding technique.

C. QCM Components

The task of this component could be abridged by detecting Quranic quotations in an input file, checking & correcting context-sensitive errors in detected quotations, then inserting or modifying a reference citation for a Quranic quotation. The input of this component is an Arabic text file which is segmented into sentences.

QCM component consists of three modules, Search Module, Error Correction Module, and Citation Module.

1) Search Module

This module is responsible for slicing each sentence in an input text file into search-patterns using the window sliding technique. Depending on the length of the search word, the module matching the search word throughout the correspondence layer the index file.

Window sliding or sliding window, interchangeably, the technique was originally used for packet transmission in computer networks [19]. This technique is adopted to facilitate matching between an input text and the Quran database. Given an input sentence of size (number of words) n, a window of size 3 would run over it to slice the sentence into n-3+1 matching-pattern. Where n-3+1 is the number of all possible patterns (3 contiguous words) that could be extracted from a sentence of n words. The sentence may be consisting of only two

words; in this case, the whole sentence will be treated as on match-pattern.

For instance, given a sentence of 7 words, 5 matching-patterns would be extracted when using a sliding window of size 3 such as the following:

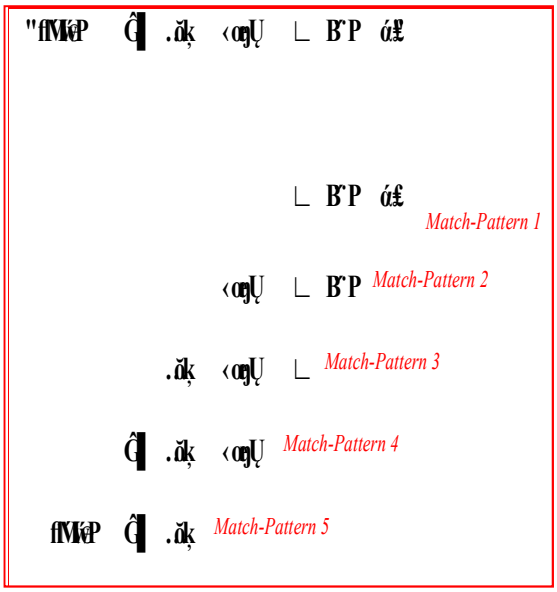

Fig. 2. Sliding Window For Extracting Match-Patterns

The proposed search module finds a Quranic quotation within an input text by matching the sliced patterns with the Quran text. The adopted searching technique guaranteed to find any Quranic quotation in a regular text regardless the quotation is cited or not.

In the case of multi-matching, where the Quranic quotation matched in more than one verse in the Quran text, a list of suggestions consisted of all matched verses will be retrieved and the user chooses the intended verse.

2) Error Correction Module

As discussed earlier in this work, the occurrence of context-sensitive errors may affect the whole meaning and also, it will prevent the search module from retrieving correct results.

If the search module found a partial matching, where two words within the same match-pattern are matched with a Quranic text but the third word is not matched, then the search module assumes the third word as a context-sensitive erroneous word and calls error correction module.

A list of lexically neighboring words for Quranic words is used to figure out and correct a context-sensitive error. Error correction module finds out if there is a word in lexically neighboring words file in which fitting the context more than the erroneous word. Once a more appropriate word is found, it replaced the erroneous word and recalling the search module to resume searching.

3) Citation Module

Where Quranic quotations in an input file are identified by the search module and context-sensitive errors –if any- corrected by the error correction module, the citation module verifies the Quranic citation style. If the citation reference is not included at all or not placed correctly by the user, the module inserts the correct citation reference by adding the correspondence chapter name and verse number.

D. Overall Process Flow

This subsection demonstrates the flow of the proposed methodology. Figure 3 represents the QCM flowchart.

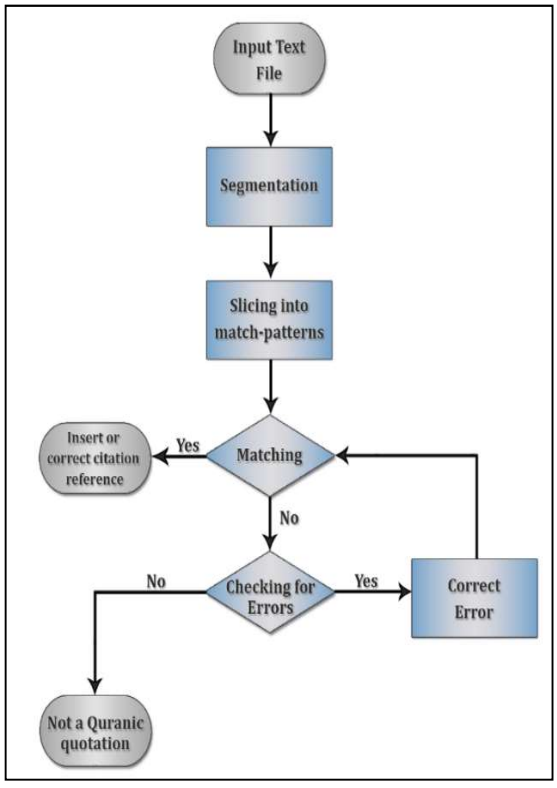

Fig. 3. Qcm Flowchart

The proposed Quran citation management algorithm has two phases, the preprocessing phase that is performed one time and the citation management phase.

QCM Algorithm:

Preprocessing Phase

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Input: *Quran corpus*
Appling text processing procedures
Output: *constructed and organized resources files which are:*

- *Quran words file*
- *Index file*
- *Lexically Neighboring words file*
- *Verses file*

Citation Management Phase
Input: *text file*
Output: *a text file with referenced Quranic citations*

1. ***Segmentation***: *segmenting input text into a sentence*
2. ***Slicing***:
   for *each sentence* ***do*** *the following:*
   if *the sentence length = 2* ***then***
   consider the whole sentence as one match-pattern.
   else if *the sentence length >=3* ***then***
   Let n = sentence length
   X_p = n-3 +1 ' number of sliced patterns
   Window_sliding (Sentence, window_size=3)
   Return *X_p match-patterns*
3. ***Search Module***:
   for *every X_i match-pattern,* ***do*** *the following:*
   3.1 Matching:
   if *Match(X_i and $Verse_{ij}$ in Quran Text)* ***then***
   Add $Verse_{ij}$ to list_of_suggistions
   else
   checks if there is any context-sensitive error.
   if *any 2 contiguous words in Xi match-pattern Matched a Quranic text* ***then***
   consider the third world as error_world
   Call ***Error Correction*** *(match-pattern X_i, _error_word)*
   else
   go to *step 5*
   3.2 Verifying Matching results:
   for *every $Verse_{ij}$ in the list_of_suggistions* ***do*** *the following:*
   if *$Verse_{ij}$ has exact or partial matching with a sentence* ***then***
   Keep $Verse_{ij}$ and discard other suggestions
   else
   Return *all verses in list_of_suggestions and let _user choose the intended verse*

3.3 Citation Module: *Check citation*
if *is it correctly placed* ***then***
finish
else
Insert the chosen suggested verse citation _reference.

4. ***Error Correction module:***
   if *the Lexical_neighbor_words file contains any words _which are close to error_word* ***then***
   replace *the error_word by the most probable _neighboring word that matches the context of _the match-pattern X_i*

4. EXPERIMENTAL SETUP AND EVALUATION

A. Experimentation Design

In experimentation design, quires are classified into four categories based on the component undertaken for testing. Queries in the first three categories were designed for testing each module individually, which are Search, Error Correction, and Citation category respectively. Quires belong to the fourth category were designed to evaluate the overall performance of QCM.

First Category: **Search**

Quires that fall into this category were designed to evaluate the ability of the search module for handling different possible causes of a Quranic quotation in a regular sentence. Evaluation aspects are as follow:

- The ability of the search module to find a verse that is relevant to the query of a single Quranic quotation without any other text.
- The ability of the search module to find a verse that is relevant to the query of a single Quranic quotation with other text.
- The ability of a search module to find multiple relevant verses where the query contains Quranic quotation that is relevant to more than one verse.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- The ability of the search module to rank relevant verses based on their similarity to a query and identifying the most similar verse.

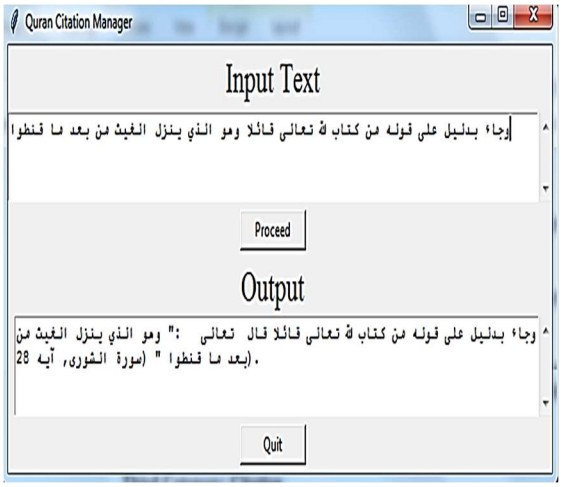

Fig. 4. Sample Result (1)

Second Category**: Error Detection and Correction**

Quires in this category were designed to measure the ability of the error correction module of identifying and correcting context-sensitive error words within a Quranic quotation.

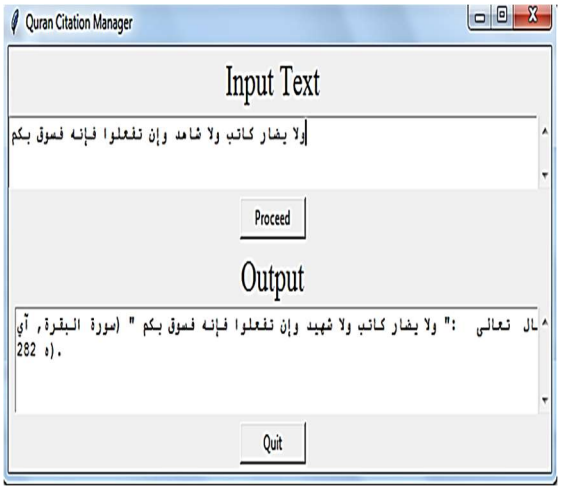

Fig. 5. Sample Result (2)

Third Category: **Citation**

The third category Quires' were designed to evaluate citation module from three aspects, identifying existed citation of a Quranic quotation, citing a nonreferenced Quranic quotation, and correcting the wrong citation.

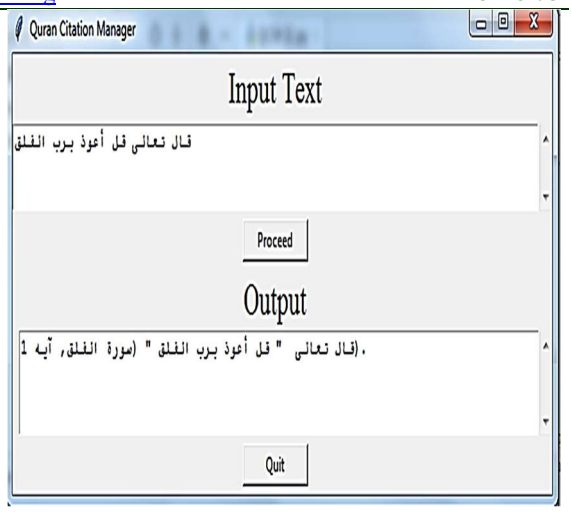

Fig. 6. Sample Result (3)

Forth Category: **Overall performance**

Overall performance evaluated throughout evaluating the collaboration between different modules. The design of these category quires covers aspects of different modules evaluation.

Fig. 7. Sample Result (4)

B. Data Set

Data for evaluating QCM were constructed manually. Data are obtained from Arabic texts that contain quotations from the Holy Quran. Quires were designed based on evaluation aspects discussed in the previous subsection. 40 different queries were extracted from the collected data.

C. Evaluation

F-measure, precision, and recall metrics are used to measure the accuracy of QCM performance. In searching evaluation, recall metric measures the number of the relevant verses were retrieved; while

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

precision metric measures the number of the retrieved verses were relevant.

$$\text{Precision} = \frac{|\{Relevant\ Verses\} \cap \{Retrieved\ Verses\}|}{|\{Relevant\ Verses\}|} \quad (1)$$

$$\text{Recall} = \frac{|\{Relevant\ Verses\} \cap \{Retrieved\ Verses\}|}{|\{total\ number\ of\ Acutual\ Relevant\ Verses\}|} \quad (2)$$

In Error correction evaluation, the recall metric measures the number of the context-sensitive errors detected were corrected; while the precision metric measures, the number of the corrected errors that were detected correctly.

$$\text{Precision} = \frac{|\{Errors\ Detected\} \cap \{Errors\ Corrected\}|}{|\{Errors\ Detected\}|} \quad (3)$$

$$\text{Recall} = \frac{|\{Errors\ Detected\} \cap \{Errors\ Corrected\}|}{|\{total\ number\ of\ Acutual\ Errors\}|} \quad (4)$$

In Citation evaluation, recall metric measures the number of the correctly cited verses among all cited verses; while precision metric measures the number of the cited verses were correct.

$$\text{Precision} = \frac{|\{Correctly\ Cited\ Verses\} \cap \{Cited\ Verses\}|}{|\{Correctly\ Cited\ Verses\}|} \quad (5)$$

$$\text{Recall} = \frac{|\{Correctly\ Cited\ Verses\} \cap \{Cited\ Verses\}|}{|\{total\ number\ of\ Acutual\ Correct\ Citations\}|} \quad (6)$$

Overall system precision and recall are represented by average precision and recall individually for all modules. F-measure represents an average metric of precision and recall; thus, system accuracy can be represented by only on a metric.

$$\text{F-measure} = \frac{2 . precision . recall}{precision + recall} \quad (7)$$

Experiments showed that the recall metric for overall system is 97.2%, precision is 90.87%, and overall accuracy represented by f-measure is 93.95%. Low precision corresponds to the high ratio of retrieved verses which is in turn corresponded to adopted matching technique. A sliding window of size 3 is guaranteed to find a relevant verse for a query, but it causes a high retrieval rate.

Figure 8, shows accuracy measures for QCM and its modules.

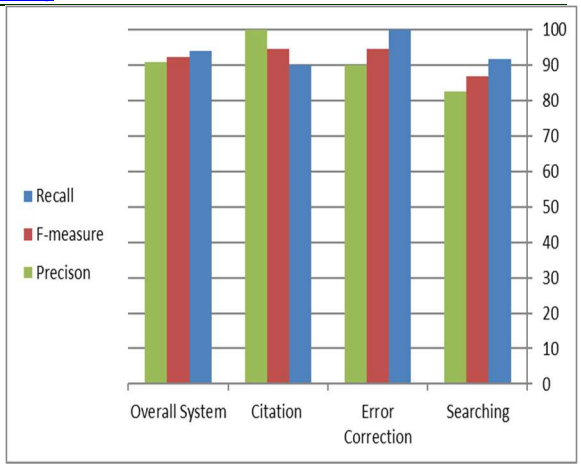

Fig. 8. Accuracy Measures

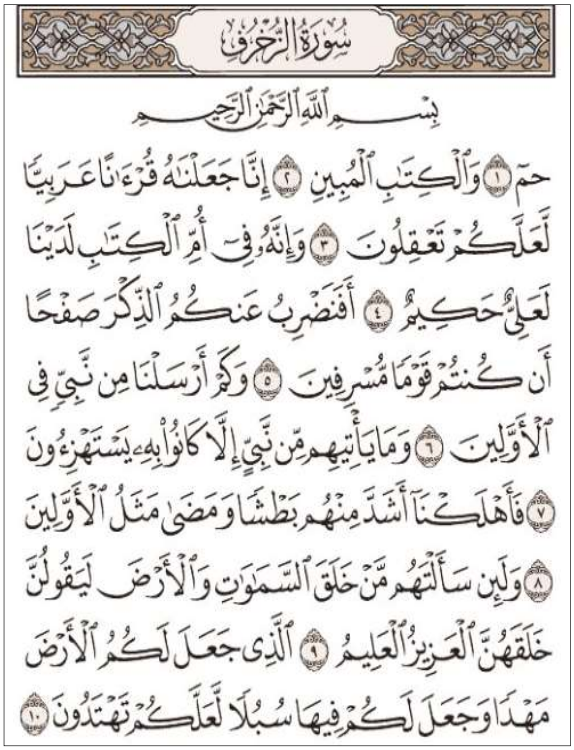

Fig. 9. Holly Quran Text with Full Diacritization

5. CONCLUSIONS AND FUTURE WORK

In this research paper, a novel Quranic Citations Manager was introduced. For handling Quranic quotations in a regular text, a Quranic citations management model is designed and implemented which is constructed of three modules, Search module, Error Correction module, and Citation module. To ensure fetching the relevant verses for a query, the sliding window technique was utilized in search matching. A regular Arabic text may be

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

containing context-sensitive errors that are unrecognizable by most of the text-editing software. To avoid the influence of context-sensitive errors on matching input text with Quranic verses, lexically neighboring words of Quranic words (words of a similar orthographical form) are used for detection and correction of context-sensitive errors. The rule of the Citation module is to cite uncited Quranic quotations and verifying the Quranic citation style.

Experiments revealed a guaranteed retrieval of query's relevant verse/s but with a high retrieval rate, that is due to limited sliding window size used for matching. Furthermore, Context-sensitive error detection and correction revealed precise results, but it is also limited for a specific number of lexically neighboring words stored in QCM resources files. Also, experiments show that handling in-text citations achieving 94.7% accuracy measure, the remaining 5.3% as lost accuracy corresponded to nature of holy Quran verses in which some verses share the same number of consecutive words, and to the nature of Quranic quotation in which some quotations consists of three or four words that is could be found in more than one verse.

To the best of our knowledge, this is the first work of automatically managing citations from the Holy Quran. Thus, comparing results to other research findings is not systematically applicable.

One of the limitations that may affects the results of thois research is the Arabic Diacritics. The Holly Quran uses the Arabic diacritics (Harakat Al-Tashkeel) (حركات التشكيل) in addition to its unique special symbols that is placed as a marker for the readers. Based on these markers , the readers may stop for a period of time, may prohibited from stopping, or may be allowed to do both actions upon desired Quranic citation. Fig. 9 shows an example of a standard Quranic text with full diacritics and unique symbols placement. Unfortunately, those symbols and diacritics limit the capabilities of the traditional search engines to provide acceptable and accurate results to the users.

As a future direction of this work, we aim to build a statistical language model for the Holy Quran, this will lead to more accurate detection and correction of context-sensitive errors and also minimizing retrieval rate.

ACKNOWLEDGMENT

The author would like to acknowledge the Deanship of Scientific Research (DSR) at Yarmouk University-Jordan since it funds this publication under grant number 22/2019, titled "Quranic Citations Managements Based on Context-Sensitive Error Detection and Correction".

REFERENCES:

[1] J. Allen, Natural language understanding, 2nd ed. CA: Benjamin Cummings, Menlo Park, 1995.

[2] S. M. S. Dashti, "Real-word error correction with trigrams: correcting multiple errors in a sentence" Lang. Resour. Eval., vol. 52, no. 2, pp. 485–502, 2018.

[3] C. Ben Othmane Zribi and M. Ben Ahmed, "Detection of semantic errors in Arabic texts," Artif. Intell., vol. 195, pp. 249–264, 2013.

[4] A. R. Golding, "A Bayesian hybrid method for context-sensitive spelling correction," Proc. Third Work. Very Large Corpora, vol. 3, p. 15, 1996.

[5] M. Boudchiche, A. Mazroui, M. Ould Abdallahi Ould Bebah, A. Lakhouaja, and A. Boudlal, "AlKhalil Morpho Sys 2: A robust Arabic morpho-syntactic analyzer," J. King Saud Univ. - Comput. Inf. Sci., vol. 29, no. 2, pp. 141–146, 2017.

[6] M. M. Alqahtani, E. Atwell, and M. Alqahtani, "Evaluation criteria for computational Quran search," Int. J. Islam. Appl. Comput. Sci. Technol., vol. 5, no. April, pp. 12–22, 2017.

[7] M. Hammo, B., Sleit, A., El-Haj, "Effectiveness of query expansion in searching the holy Quran," Colloq. Int. Trait. Autom. la Lang. Arab. CITALA, vol. 7, no. January, pp. 18–19, 2007.

[8] A. Al-Gharaibeh, A. Al-Taani, and I. Alsmadi, "The usage of formal methods in Quran search system," Icics 2011, no. January 2011, pp. 22–24, 2011.

[9] A. R. Yauri, R. A. Kadir, A. Azman, and M. A. A. Murad, "Quranic verse extraction base on concepts using OWL-DL ontology," Res. J. Appl. Sci. Eng. Technol., vol. 6, no. 23, pp. 4492–4498, 2013.

[10] S. Raza, M. Rehan, A. Farooq, and S. Ahsan, "An essential framework for concept-based evolutionary Quranic search engine (CEQSE)," Sci. Int., vol. 26, no. 1, pp. 181–184, 2014.

[11] P. E. D. Love, J. Smith, I. Simpson, M. Regan, and O. Olatunji, "Understanding the landscape of overruns in transport infrastructure projects," Environ. Plan. B Plan. Des., vol. 42, no. 3, pp. 490–509, 2015.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[12] N. AlShenaifi, R. AlNefie, M. Al-Yahya, and H. Al-Khalifa, "Arib @ QALB-2015 Shared Task: A hybrid cascade model for arabic spelling error detection and correction," Proc. Second Work. Arab. Nat. Lang. Processing, August 2017, pp. 127–132, 2015.

[13] M. Attia, P. Pecina, Y. Samih, K. Shaalan, and J. Van Genabith, "Arabic spelling error detection and correction," Nat. Lang. Eng., vol. 22, no. 5, pp. 751–773, 2016.

[14] H. M. Noaman, S. S. Sarhan, and M. A. A. Rashwan, "Automatic Arabic spelling errors detection and correction based on confusion matrix- noisy channel hybrid system," J. Theor. Appl. Inf. Technol., vol. 40, no. 2, pp. 54–64, 2016.

[15] M. Attia, P. Pecina, and Y. Samih, "Improved spelling error detection and correction for Arabic.," COLING (Posters), vol. 4, no. December, pp. 103–112, 2012.

[16] M. M. Al-Jefri and S. A. Mahmoud, "Context-sensitive Arabic spell checker using context words and N-Gram language models," Proc. - 2013 Taibah Univ. Int. Conf. Adv. Inf. Technol. Holy Quran Its Sci. NOORIC 2013, no. August 2016, pp. 258–263, 2015.

[17] A. Zahui, W. Elhor, and M. A. Cheragui, "EL-Mossahih V1.0: A hybrid approach for detection and correction of typographical and phonetic transcription errors in Arabic texts," ICIT 2017 - 8th Int. Conf. Inf. Technol. Proc., pp. 774–779, 2017.

[18] H. Zarrabi-Zadeh, "TANZIL," 2007. [Online]. Available: http://tanzil.net/download/. [Accessed: 16-May-2019].

[19] Larry L. Peterson and Bruce S. Davie, Computer networks: a system approach, 5th ed. Elsevier, Inc., 2012.

[20] NAHAR, K., Khatib, R., Shannaq, M., & Barhoush, M. (2020). AN EFFICIENT HOLY QURAN RECITATION RECOGNIZER BASED ON SVM LEARNING MODEL. Jordanian Journal of Computers and Information Technology, 1. https://doi.org/10.5455/jjcit.71-1593380662.

[21] Al-Kabi, M. N., Kanaan, G., Al-Shalabi, R., Nahar, M. O. K., & Bani-Ismail, M. B. (2005). Statistical Classifier of the Holy Quran Verses (Fatiha and Yaseen Chapters). Journal of Applied Sciences, 5(3), 580–583.

[22] Yousfi, B., & Zeki, A. M. (2016). Holy Qur'an speech recognition system distinguishing the type of recitation. In 2016 7th International Conference on Computer Science and Information Technology (CSIT) (pp. 1–6). IEEE. https://doi.org/10.1109/CSIT.2016.7549454.

[23] Nahar, K. M., Al-Shannaq, M., Manasrah, A., Alshorman, R., & Alazzam, I. (2019). A Holy Quran Reader/Reciter Identification System Using Support Vector Machine. International Journal of Machine Learning and Computing, 9(4).

[24] Al-Shenak, M., Nahar, K., Halwani, H.: AQAS: Arabic question answering system based on SVM, SVD, and LSI. J. Theor. Appl. Inf. Technol. 97(2), 681–691 (2019).