

Exctracting Quaranic words using Kabeer Program

Dr. Noora Subian Aljohani
Associate Professor in Linguistics
University of Jeddah, Jeddah,
Saudi Arabia
nosbj@hotmail.com
<http://nosbj.com>

Abstract

In the Era of technology where computers become the main machines of any system, the development of any system is measured by the degree of using IT in various life domains. This device has entered into many fields in our lives such as social, economic and culture fields due to its speed processing and the huge ability to recognize, store information, and provide secure transactions, and accurate managing of the system .This research is interested in exploring the use of the vocabulary of the Holy Qur'an and replacing the words that a person may use with his daily life with the terms of the Holy Qur'an through diagnosing the alternatives that a person deals with and replacing them with the terms of the Holy Qur'an to revive the use of the Qur'an in our daily life. This research has merged the advantages of modern technology with the origins of vocabulary to increase the accuracy of extracting Quranic words.

Keywords: NLP, classification, Qur'an, The use of the Holy Quran

الملخص

في عصر التكنولوجيا حيث أصبحت أجهزة الكمبيوتر هي الآلات الرئيسية لأي نظام ، يقاس تطور أي نظام بدرجة استخدام تكنولوجيا المعلومات في مجالات الحياة المختلفة. دخل هذا الجهاز في العديد من المجالات في حياتنا مثل المجالات الاجتماعية والاقتصادية والثقافية نظراً لسرعة معالجته وقدرته الكبيرة على التعرف على المعلومات وتخزينها وتوفير معاملات آمنة وإدارة دقيقة للنظام. يهتم هذا البحث باكتشاف استخدام مفردات القرآن الكريم واستبدال الكلمات التي قد يستخدمها الإنسان في حياته اليومية بمصطلحات القرآن الكريم من خلال تشخيص البدائل التي يتعامل معها الشخص واستبدالها لإحياء استخدام القرآن في حياتنا اليومية. وقد دمج هذا البحث بين مزايا التقنية الحديثة وأصول المفردات لزيادة

دقة استخراج الكلمات القرآنية. (المعجم المفهرس لألفاظ القرآن الكريم)

الكلمات المفتاحية: اللغات الطبيعية , مفردات القرآن , تصنيف المفردات

Introduction

It is clear and known to human beings that the Arabic language is one of the most languages in the world that contains a lot of vocabulary and is considered one of the richest languages in the world where one thing has more than one name in this language. For example, the word honey has many names such as witness, presumption, diet, Melt, blame, feast, and offspring. Also, the fox has many names that are more than two hundred names. And other words that have many names (1318هـ)Also, one word has many meanings. Each of the words can relate and link to another vocabulary such as a book, reading, and story as shown in the figure 1. Therefore, all word meanings can be interlinked, interconnected and constitute a relation network. We used Natural Languages processing in our methodology to increase the accuracy of a classification.

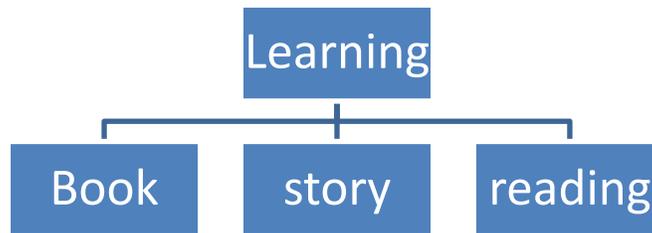


Figure 1 words related to learning

Natural Language Processing

It is a branch of artificial intelligence that includes the integration of the field of computer science and linguistics (Alisa Kongthon, 2018) to deal with texts. Natural Language Processing is the analysis of linguistic data. It is concerned with analyzing vocabulary through mathematical operations. The goal of processing natural languages is to design a model that adds structure to the natural, informal language by learning the language to analyze these vocabulary (Karin Maria Verspoor, 2013). In our methodology, we used the combination of the NLP with training dataset to export query in the database to get the word we need and our work can be described as the following steps:

- 1- Read Natural language input
- 2- Parsing syntax
- 3- Apply training dataset with knowledge base
- 4- Export Query

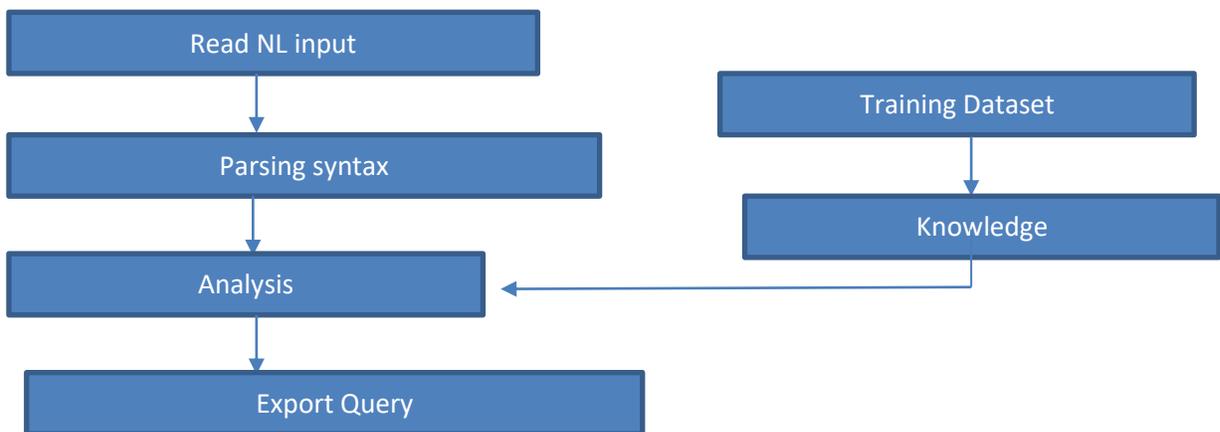


Figure 2 NL Processing

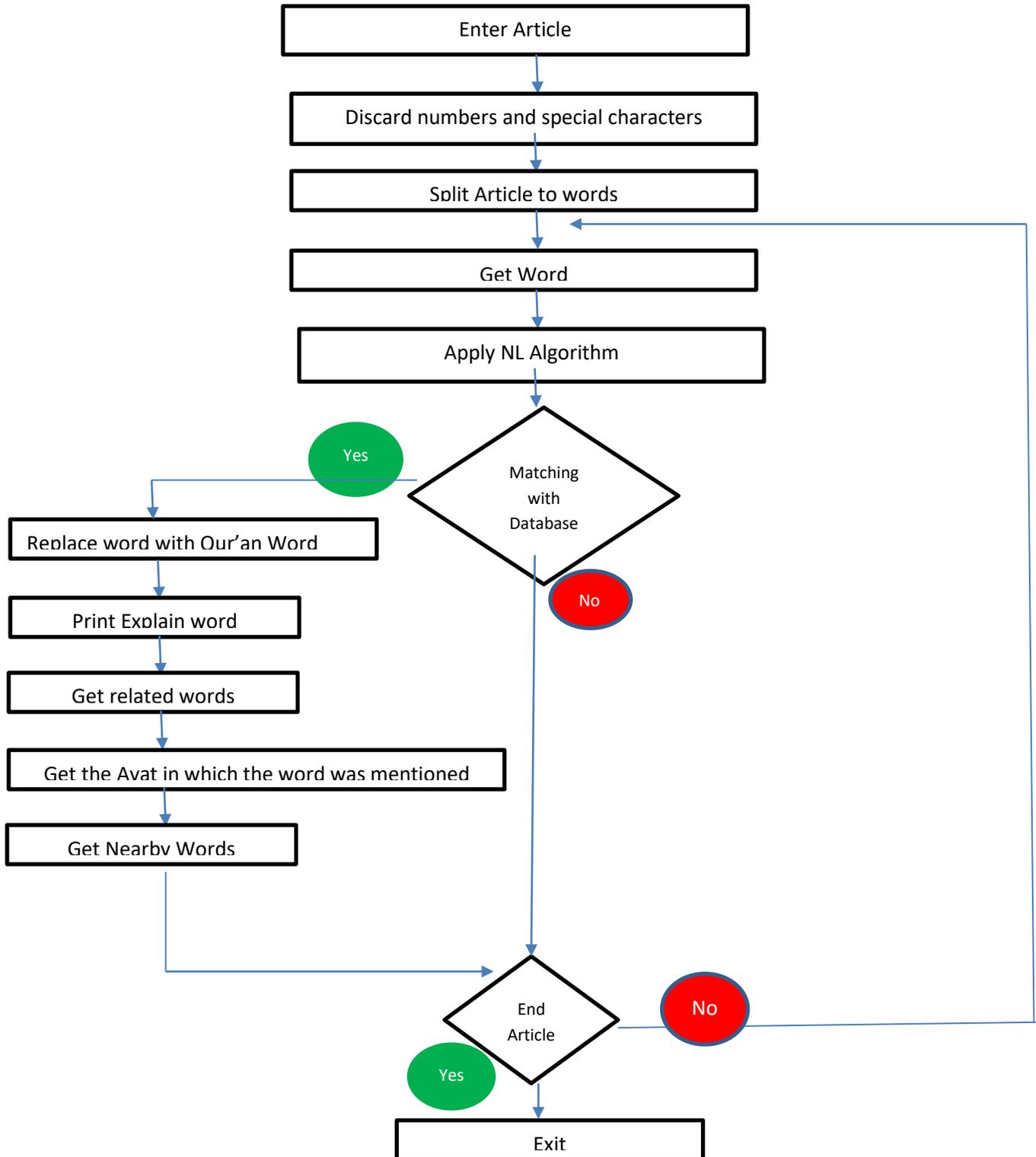
The Literature Review

In (Shalan, 2010) the author addresses efforts that involved a rule-based process for different Arabic natural language processing tasks. The characteristics of this method are that it is easy to combine domain knowledge into linguistic knowledge which increases accurate results. Domain rules have been used in generating Arabic sentences (K. Shalan, 2009) and the analysis of learning Data input (M. Magdy, 2007). Also, the linguistic knowledge acquired for one natural language processing system may be reused to build the knowledge required for a similar task in another system. In (El-Halees, 2007) the author focus on classifying Arabic data by using the maximum entropy method. The process flow in many steps, in the preprocessing step, the data processing using natural language processing techniques such as tokenizing, stemming and part-

of-speech. Then, we used a maximum entropy method to classify Arabic documents. The author experimented approach using real data, then compared the results with other existing systems. In (Alaa Al-Gharaibeh) the author explains how to use formal methods of natural language processing in the Qur'anic research system. Z coding is used to express the official specifications of the three text, text and synonym-based search systems used in QSS. QSS allows the user to search for keywords in the Holy Quran and retrieve related verses. The Z / EVES tool is used to examine and analyze Z specifications. In (Abidah Zainal, 2017) the author develops a mobile application that is taking a natural language approach to searching topics in the Quran based on keyword searching.

The Proposed Work

In this paper, we combine the Natural Language Processing with Training Dataset and apply a multi-filter operation on each parsing word to increase the accuracy of classification in the first step of our method is to filter the input article from any numeric or special character as(%, #,\$,&,!) then split this to words depended on the space character after saving these word in the array we parsing each word to discard any harakat or added characture as (-) then we pass the output word to NL algorithm to find the highest matching word in the dataset and if the result found we can get the Explain word, Related Words, Nearby words, and ayat, else we print the same word and repeat the above steps on the new word until the end of the article. We can show the flowchart in the following figure



Implementation and result

The next figures display the running application using ASP.NET and display the output of system

الرجاء ادخال النص المراد البحث فيه

لقد فاجأه ما فعلت معه، لم يكن يتوقع أن اعفو عنه،
وهنا كان للعفو طعم آخر، ولذلك فقد أقبل ابني نحوي
وقبلني وقال: "أحبك" ♥ واتفقت معه على رد
المظالم لأمه وإخوته، وفكرنا معاً كيف يصلح ما
أفسده.
وبعدها بأيام بدأت أفكر معه كيف يكسب أمه
وإخوته، وكم فرح المسكين بذلك وتغيرت أحواله
للأفضل،
لقد اكتشفت أننا نعاقب أبناءنا عندما يسيئون، لكننا لا
نعلمهم كيف يحسنون!
من كتاب: (بالحب نربي أبناءنا)

بحث

نتائج الفحص

لقد فاجأه ما فعلت معه، لم يكن يتوقع أن اعفو عنه،
وهنا كان **لقد فاجأه ما فعلت معه، لم يكن يتوقع أن اعفو عنه،**
وهنا كان نُقِفَ للعفو طعم آخر،
وإخوته، وكم فرح المسكين بذلك وتغيرت أحواله
للأفضل، لقد اكتشفت أننا نعاقب أبناءنا

آيات ذكرت فيها

- 1 وَأَمْتَلُوهُمْ حَيْثُ ثَقِفْتُمُوهُمْ (البقرة ١٩١)
- 2 صُرِّبَتْ عَلَيْهِمُ الدَّلَّةُ أَيْنَ مَا تُوقَمُوا إِلَّا بِحِجْلِ
مِنَ اللّٰهِ وَحِجْلٍ مِّنَ النَّاسِ (آل عمران ١١٣)
- 3 فَحَدِّثُوهُمْ وَأَمْتَلُوهُمْ حَيْثُ ثَقِفْتُمُوهُمْ (٩١)
(النساء)

التفسير

المعاني الجامع : ثَقَّفَ فعل ثَقَّفْتُ أَنْثَقِّفُ
مصدر ثَقَّفَ ثَقَّفَ صَاحِبُهُ عَلِيَّةٌ فِي الدُّعَاةِ
وَالْفَهَازَةِ ثَقَّفَهُ بِالرِّمْحِ طَعَنَهُ بِهِ ثَقَّفَ اسم
صفة مشبهة تدلُّ على الثبوت من ثَقَّفَ
وَتَهَفَّ ثَقَّفَ اسم تَهَفَّ فاعل من ثَقَّفَ ثَقَّفَ
اسم ثَقَّفَ مصدر ثَقَّفَ ثَقَّفَ اسم ثَقَّفَ جمع
ثَقَافِ ثَقَّفَ فعل ثَقَّفَ يَثَقِّفُ ثَقَافَةً فهو
ثَقِّفٌ ثَقَّفَ الشَّخْصَ صارَ حَادِقًا فَطِنًا ثَقَّفَ
فعل ثَقَّفَ يَثَقِّفُ ثَقِّفًا فهو ثَقِّفٌ والمفعول

كلمات ذات صلة

أَثَقَفَةٌ تَثَقِّفُ تَثَقِّفُ تَثَقِّفُ تَثَقِّفُ تَثَقِّفُ
ثَقَافِيٌّ ثَقَافٍ ثَقَافٍ ثَقَافِيٌّ ثَقَافِيٌّ ثَقَافِيٌّ
مُتَثَقِّفٌ مُتَثَقِّفٌ مُتَثَقِّفٌ مُتَثَقِّفٌ مُتَثَقِّفٌ



Figure 4 Testing

Results

The system tested and publish as a website as shown in the above figures and can anyone test the result from the website: <http://nosbj.com/>, also all trained words from the system can retrieve successfully, so the accuracy of our system is 100%.

Recommendations

At the end of our research we would like to say that the field of research is endless and every day new research is developed. In our research, we recommend using the words of the Qur'an and preserving it and using it in our daily life to revive this language, as in any system or research on which continuous development is made, and for this, all the ayat of the Holy Qur'an and all the words and words related to it will be entered

References

- 1318هـ، ب. أ. (n.d.). معجم أسماء الأشياء المسمى اللطائف في اللغة. القاهرة-دار الفضيلة. 346.
- Abidah Zainal, N. A. (2017). Natural Language Processing Approach for Searching the Quran. *MATEC Web of Conferences* .
- Alaa Al-Gharaibeh, A. A.-T. (n.d.). The Usage of Formal Methods in Quran Search System. *Faculty of Information Technology, Yarmouk University, Jordan*.
- Alisa Kongthon, C. S. (2018). Implementing an online help desk system based on conversational agent . *Proceeding, MEDES '09 Proceedings of the International Conference on Management of Emergent Digital EcoSystems*.
- El-Halees, A. M. (2007). Arabic Text Classification Using Maximum Entropy. *The Islamic University Journal (Series of Natural Studies and Engineering)*, 15(1), 157-167.
- K. Shaalan, A. A. (2009). Syntactic Generation of Arabic in Interlingua-based Machine Translation Framework. *Third workshop on Computational Approaches to Arabic Script-based Languages (CAASL3)*.
- Karin Maria Verspoor, K. B. (2013). Natural Language Processing. *Research Gate*.
- M. Magdy, K. S. (2007). Lexical Error Diagnosis for Second Language Learners of Arabic. *Proceedings of The Seventh Conference on Language Engineering*.
- Shaalan, K. (2010). Rule-based Approach in Arabic Natural Language Processing . *International Journal on Information and Communication Technologies*, 3(3), 11-20.
- (n.d.). المعجم المفهرس لألفاظ القرآن الكريم