

PAPER • OPEN ACCESS

Answering Islamic Questions with a Chatbot using Fuzzy String-Matching Algorithm

To cite this article: M T Sihotang *et al* 2020 *J. Phys.: Conf. Ser.* **1566** 012007

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

Answering Islamic Questions with a Chatbot using Fuzzy String-Matching Algorithm

M T Sihotang¹, I Jaya^{1*}, A Hizriadi¹ and S M Hardi¹

¹ Faculty of Computer and Information Technology, Universitas Sumatera Utara, Jalan Universitas No. 9A, Medan, Indonesia

*Email: ivanjaya@usu.ac.id

Abstract. The guidance of Muslims in worship refers to the holy Quran and the hadith. Not all people understand the law related to a case in accordance with Islamic teachings. Questions relating to a matter based on Islamic law are widely circulated on the Internet with long and detailed answers. This is good but for some people, a short and direct answer to the core is the desired answer, of course in accordance with the majority of Muslim scholars. One of the many technologies that can be used to answer questions is chatbot. A chatbot is one of many implementations of Natural Language Preprocessing. In this study, the chatbot can find answers to questions in accordance with Islamic law using the fuzzy string-matching algorithm. The research test data were obtained from several people who used chatbot directly by looking at pairs of questions and answers whether they were appropriate or not. The accuracy of the test is 70.37% and the chatbot's performance is quite good.

1. Introduction

Every human being has a problem in his life and they are always looking for a solution to that problem. Some problems experienced by humans include economic problems, education, health and so forth.

For some religious people such as Muslims, worship can be a problem if it is not in accordance with the Qur'an and the hadith. Some people worship because they follow others or lack knowledge in religious understanding. For this reason, an approach is needed to overcome this problem.

With the development of current information technology, references related to the problems of worship of Muslims are widely spread and can be accessed online via the Internet, in detail, and on target. But the information provided on the Internet also needs to be checked whether it is in accordance with the Qur'an and hadith. Some websites also provide space for internet users to ask questions about things they want to know about Islam. Of course, the user must wait for an answer from the website manager. It could be that the answer is given quickly or it might take days. To deal with this, we need a system that can answer questions raised by users directly and quickly. This task can be done using Natural Language Processing (NLP).

NLP is one of the fields of computer science, part of the field of artificial intelligence. The purpose of NLP is to design and build applications that can provide facilities for human interaction with machines or with other devices through the use of natural language. One of the major areas of NLP is Question Answering Systems (QAS) [1]. With QAS, users can ask questions directly in natural languages such as English, Arabic or Indonesian. One of the implementations of QAS is chatbot.

A chatbot is a computer program that makes users feel as if they are able to communicate with humans using natural language. Chatbot is widely applied in various fields ranging from weather reports,



customer service, troubleshooting and so forth. A language model and pattern matching algorithm are part of the chatbot architecture that can be utilized to imitate informal conversations between humans and systems [2].

In this study, we created a system (chatbot) that can answer questions from users about issues related to Islamic religious worship using the natural language of Indonesia (Bahasa Indonesia).

2. Material and methods

The data used to answer user questions about Islam is named data.rive which is a collection of questions and answers summarized from several sites such as eramuslim, dakwatuna and muslim.or.id. The specifications of data.rive (Figure 2) contains questions and answers about prayer (shalat) and zakat. So at this research, the questions that can be answered by chatbot at this time only questions related to prayer and zakat.

The general architecture could be explained in Figure 1.

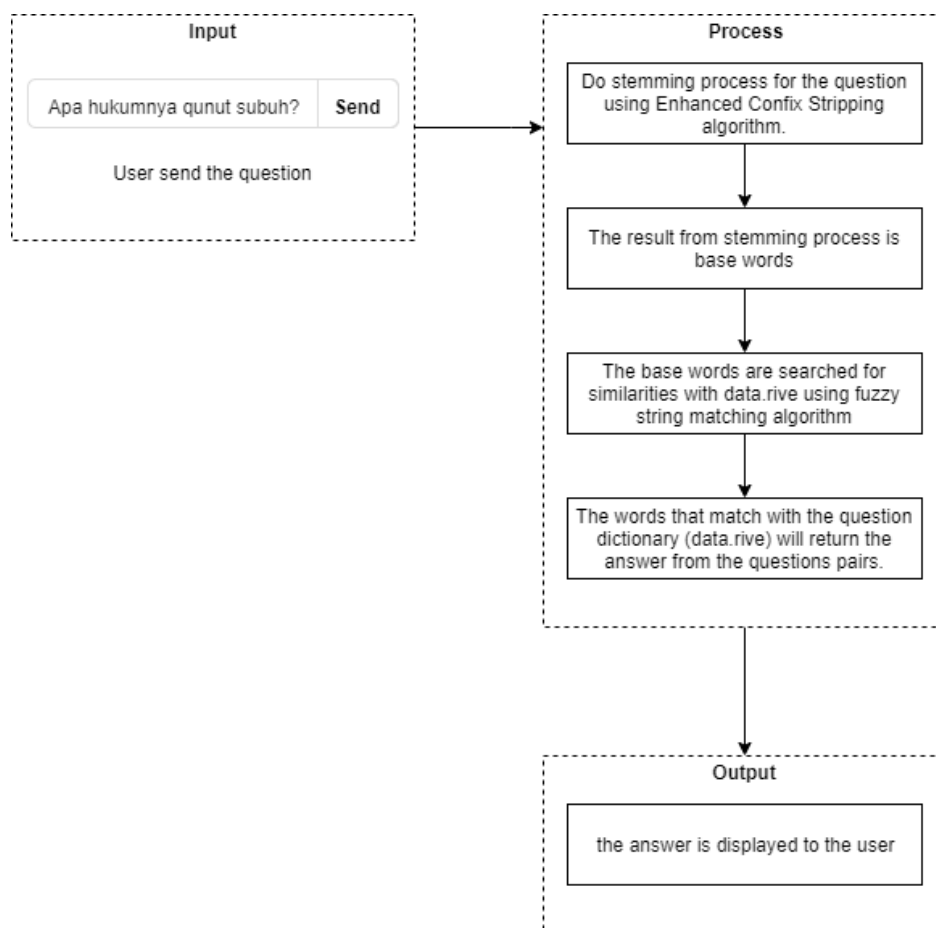


Figure 1. The general architecture of chatbot.

Line No.	Text	Status
4898	^\\n Demikian juga berdoa agar menjadi penghuni surga, ketika membaca ayat tentang surga. Hudzaifah Al Yaman r.a berkata,	✓
4899	^\\n "aku shalat bersama Nabi SAW. Ketika membaca ayat-ayat tasbih, ia bertasbih, ketika membaca surat pertanyaan, ia bertanya dan jika membaca	
4900	^\\n ia memohon perlindungan". (HR. Muslim)	
4901		
4902	+ [*] (bagaimana apa) hukum qunut subuh [*]	
4903	- Dari Anas bin Malik, ia berkata: "Rasulullah Saw terus menerus membaca Qunut pada shalat Shubuh hingga beliau meninggal dunia".	
4904	^\\n Hadits ini riwayat Imam Ahmad, ad-Daraquthni dan al-Baihaqi.	
4905	^\\n Qunut Shubuh Menurut Mazhab Syafi'i:	
4906	^\\n Adapun Qunut, maka dianjurkan pada I'tidal kedua dalam shalat Shubuh berdasarkan riwayat Anas, ia berkata: "Rasulullah Saw terus	
4907	^\\n meninggal dunia". Diriwayatkan oleh Imam Ahmad dan imam lainnya. Imam Ibnu ash-Shalah berkata,	
4908	^\\n "Banyak para al-Hafizh (ahli hadits) yang menyatakan hadits ini adalah hadits shahih. Diantara mereka adalah Imam al-Hakim, al-Baihaqi dan	
4909	^\\n Al-Baihaqi berkata, "Membaca doa Qunut pada shalat Shubuh ini berdasarkan tuntunan dari empat Khulafa' Rasyidin".	
4910		
4911	+ [*] (bagaimana apa) hukum orang yang sholat subuh pakai qunut [*]	
4912	- Dari Anas bin Malik, ia berkata: "Rasulullah Saw terus menerus membaca Qunut pada shalat Shubuh hingga beliau meninggal dunia".	
4913	^\\n Hadits ini riwayat Imam Ahmad, ad-Daraquthni dan al-Baihaqi.	
4914	^\\n Qunut Shubuh Menurut Mazhab Syafi'i:	
4915	^\\n Adapun Qunut, maka dianjurkan pada I'tidal kedua dalam shalat Shubuh berdasarkan riwayat Anas, ia berkata: "Rasulullah Saw terus	
4916	^\\n meninggal dunia". Diriwayatkan oleh Imam Ahmad dan imam lainnya. Imam Ibnu ash-Shalah berkata,	
4917	^\\n "Banyak para al-Hafizh (ahli hadits) yang menyatakan hadits ini adalah hadits shahih. Diantara mereka adalah Imam al-Hakim, al-Baihaqi dan	
4918	^\\n Al-Baihaqi berkata, "Membaca doa Qunut pada shalat Shubuh ini berdasarkan tuntunan dari empat Khulafa' Rasyidin".	
4919		

Figure 2. Data collection of questions and answers (data.rive).

2.1. Input

This section is the initial stage where a user submits a question relating to Islamic law. For example: "Apa hukumnya qunut subuh".

2.2. Process

The process stage is the stage after the input stage. There are two main steps at this stage, the explanation being as follows:

- Stemming

In this step, each word in the question sentence will be checked whether it is in the basic word dictionary or not. The process begins with tokenization, breaking the question sentence into word-to-word tokens and comparing each word to a dictionary (KBBI). If there is a match, then the stemming process will be stopped and the words become basic words. But if there isn't, then the stemming process will be run.

Stemming is a process to search root words from each word of the text by deleting affix in the word [3, 6]. For example: "Apa hukumnya qunut subuh". There are three basic words: "apa", "qunut" and "subuh". There is one word that needs to be removed from the "-nya" affix, which is the word "hukumnya".

In this system, there is a suffix that is part of the affix that needs to be removed from the question sentence namely inflectional suffixes. Inflectional suffixes are the set of suffixes that do not alter the root word, divided into a particle (P) and Possessive Pronoun (PP). P consist of "-lah", "-kah", "-tah", and "-pun". PP consists of "-ku", "-mu" and "-nya" [4].

From the example question, the word "hukumnya" will be stemmed from the removal of the affix "-nya" to the word "hukum". The result of the stemming process is "apa hokum qunut subuh". the next process is the search for answers by using string-matching to get rules based on the verses of the Qur'an or the corresponding hadith that have been collected.

- String-matching

In this system, we use a fuzzy string-matching algorithm. The model of the fuzzy string-matching algorithm can be seen in Figure 2. The first step in this process is that the string is matched with the string in the dictionary, if it has a match then the answer from the question pair can be displayed. If the string has not been found, enter the next step.

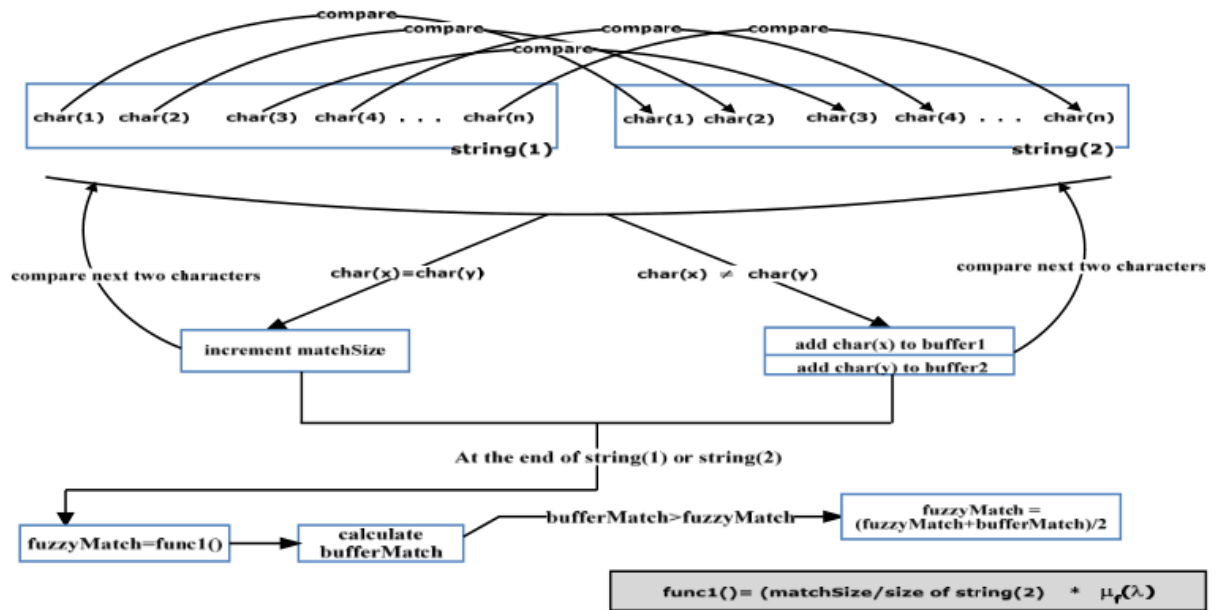


Figure 3. Fuzzy string-matching model [5].

In the next step, the initial letter and the ending letter and the length of the string will be checked for similarities with the dictionary. If a match is found then the pair of answers that correlated with the question will be displayed.

In the example, the string-matching process will start with the words "apa hukum" and "qunut subuh". After checking the dictionary, the system gets matched data: "(bagaimana|apa) hukum qunut subuh", so that user questions can be taken from questions in the matching dictionary, that is "(bagaimana|apa) hukum qunut subuh". After that, the pair of answers to the question will be returned to the user.

2.3. Output

The output is a valid answer obtained from the pair of questions contained in the dictionary displayed on the screen.

3. Result and discussion

Testing is carried out randomly, asking users to enter questions related to prayer and zakat. From 27 questions, we get accurate and inaccurate answers. Table 1. shows some of the questions answered by the chatbot. To calculate the percentage of success, the following formula is used:

$$\text{Percentage of success} = \frac{\text{success test responses}}{\text{total test}} \times 100\% = \frac{19}{27} \times 100\% = 70,37\%$$

From the results above, we found some answers that were not appropriate. This is because the fuzzy string-matching algorithm does a search by looking at a number of characters that are similar both at the beginning and at the end so that the full sentence of the user's question is pretty much ignored. For example when a user asks "apa saja yang membatalkan shalat?", the string-matching process will search starting from "apa saja" then continuing to "yang membatalkan" and the last "shalat".

The dictionary which contains a collection of answers about prayer and dzakat found that there is a word that fits in the first part: "apa saja rukun dalam shalat". The word "apa saja" becomes a reference in the string-matching process. The words "yang membatalkan shalat" from the user's question is found

have a similarity at the end which is "salat", so the user's question sentence is taken from the question in the dictionary that is "apa saja rukun dalam shalat" so that the pair of answers given is not right.

Table 1. Chatbot test results.

Question	Stemming Process	Answer	Result
Jelaskan manfaat shalat	Jelas manfaat shalat	Pewaris surga Firdaus dan kekal di dalamnya. Allah Ta'ala berfirman (yang artinya), "Sungguh beruntung orang-orang yang beriman ... dan orang yang memelihara shalat mereka itulah orang-orang yang akan mewarisi, (yakni) yang akan mewarisi surge Firdaus. Mereka kekal di dalamnya. (QS. Al Mu'minun: 1 – 11)	Matched
Kapan shalat subuh dilaksanakan	Kapan shalat subuh laksana	No Reply Matched	Not Matched
Sebutkan rukun shalat	Sebut rukun shalat	Hukumnya wajib bagi laki-laki karena Rasulullah selalu mencontohkan shalat di Masjid dan shalat berjamaah juga mendapat dua puluh tujuh tingkatan lebih banyak daripada di rumah sebagaimana dalam hadits, Dari Ibnu umar, sesungguhnya Rasulullah SAW bersabda: "Shalat berjamaah lebih baik daripada shalat dua puluh tujuh tingkatan" (HR. Muslim)	Not Matched

4. Conclusion

From the test results, it was concluded that chatbot which implements a fuzzy string-matching algorithm has an accuracy rate of 70.37% in the process of searching for user answers related to questions related to prayer and zakat. To improve the results of the answers, it is necessary to add more vocabulary questions to the dictionary. Additional NLP methods are also needed such as the use of n-grams to get a better string-matching. Crawling methods can also be considered to enrich the database and implementing the ranking methods in presenting search results so that the most relevant results are presented in the first place.

References

- [1] J. Pustejovsky and A. Stubbs, *Natural Language Annotation for Machine Learning*. Sebastopol: O'Reilly Media, Inc., 2013.
- [2] A. Vichare, A. Gyani, Y. Shrikhande, and N. Rathod, "A chatbot system demonstrating Intelligent Behaviour using NLP," *Int. J. Adv. Res. Comput. Eng. Technol.*, vol. 4, no. 10, pp. 4–6, 2015.
- [3] I. Jaya, I. Aulia, S. Hardi Melvani, J. Timanta Tarigan, M. Silvi Lidya, and Caroline, "Algorithm Scientific Documents Classification Using Support Vector Machine Algorithm," *3rd Int. Conf. Comput. Appl. Informatics*, 2019.
- [4] H. T. Ciptaningtyas, "Enhanced Confix Stripping Stemmer and ANTs Algorithm for Classifying News Document in Representation of Textual," *5th Int. Conf. Inf. Commun. Technol. Syst.*, pp. 149–158, 2007.
- [5] F. W. Onifade, M. Iaeng, O. Thiéry, O. Osofisan, and G. Duffing, "Dynamic Fuzzy String-Matching Model for Information Retrieval Based on Incongruous User Queries," *Proc. World Congr. Eng.*, vol. I, 2010.

- [6] Priandini, Novita, Badrus Zaman, and Endah Purwanti. "Categorizing document by fuzzy C-Means and K-nearest neighbors approach." In *AIP Conference Proceedings*, vol. 1867, no. 1, p. 020012. AIP Publishing LLC, 2017.