International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-8 Issue-5S March, 2019

Building a Question Classification Model for a Malay Question Answering System

NurnasranPuteh, Mohd. ZabidinHusin, Hatim Mohamad Tahir, Azham Hussain

Abstract: Question answering system (QAS) is an example of an application of natural language processing where it is able to automatically return a specific answer to a question given in a natural language by a human. One of the important tasks in QAS is Question Classification which is the task to identify the semantic type of the required answer for the question posed to the QAS. Identifying the correct answer type is an important process before the required correct answer can be retrieved by the system. In this paper we presents a model of Answer Type Classification using machine learning approach targeted for a Malay QAS for the Quran, which is a restricted-domain QAS. The performance of the classification model using three different machine learning classification algorithms, namely Naïve Bayes, Random Forest and Support Vector Machine (SVM), are then evaluated. The results show that the classifier based on SVM has the best overall results in terms of accuracy, precision, recall and F1-score.

Keywords: Malay Question Answering, Question classification, Machine learning, Quran, QAS

I. INTRODUCTION

Question answering system (QAS) is an example of an application of natural language processing. Using this system, a user can pose a query in a natural language and the correct answer to the query will be automatically returned by the system (Hirschman &Gaizauskas, 2001). As such, QAS can be considered as a specific type of information retrieval. This can be contrasted with a search engine where the query is given in the form of keywords and a list of links or documents related to the query, which possibly contain the answer, will be returned. Popular examples of QAS are IBM's Watson and Apple's Siri. These are examples of general domain QAS, where the QAS can answer domain independent questions. QAS can also be categorized as restricted domain if it can answer only domain specific questions (Mishra&Jain, 2016). Examples of restricted domain QAS are geospatial domain QAS, medical domain

Revised Manuscript Received on March 10, 2019.

NurnasranPuteh, School of Computing, College of Arts and Science, University Utara Malaysia, Sintok, Malaysia,

Mohd. ZabidinHusin, School of Computing, College of Arts and Science, University Utara Malaysia, Sintok, Malaysia

Hatim Mohamad Tahir, School of Computing, College of Arts and Science, University Utara Malaysia, Sintok, Malaysia

Azham Hussain, School of Computing, College of Arts and Science, University Utara Malaysia, Sintok, Malaysia

The architecture of a QAS consists of three major stages (Jurafsky& Martin, 2009) which are Question Processing, Relevant Passage Retrieval, and Answer Processing. QAS, community-based QAS and patent QAS (Molla&Vicedo, 2007).

Question Processing involves the task of analyzing the question so that the meaning of the question can formally be represented. An important task in Question Processing is the Question Classification which is the task to classify the questions that are posed to the QAS according to the semantic type of their expected answers. The question classification process can significantly help a QAS in predicting the type of entities needed to be present in the relevant passage candidates (second stage in QAS) by classifying the questions into various correct answer types taxonomy (Santosh et al., 2010) such as person, location, time etc. Therefore, predicting the correct answer types is an important factor in the success of a QAS in producing the correct answer for a question. For example, the search for the correct answer to the question like "Who is the Prime Minister of Malaysia?" can be improved if the system can predict beforehand that the expected answer type to that question is "Person". Consequently, the search space for the correct answer for this question can be narrowed down to only those candidate passages that contain the "Person" entities. Since in question classification involves the task of predicting the answer type, it is also referred as answer type prediction.

Although currently there are a lot of QAS that have been built using the English and various other languages, QAS in the Malay language is still lacking. In this work, we focus on a domain-specific Malay QAS which supports restricted domain questions only. The domain that we choose is the Question Answering for the Holy Quran. The motivation for selecting this domain is that the current methods of getting information from the translated Malay Quran are limited only to keyword-based information retrieval (Ahmad et al., 2017). The main objective of this paper is to present our work in building the learning model for the Question Classification component of a Malay QAS and to report on the performance evaluation of the classifiers that are built with different machine learning classification algorithms. The rest of the paper is organized as follows: Section 2 provides the related work. Section 3 explains the approach used to build the model. Section 4 presents the evaluation results and finally conclusion and future work in Section 5.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

II. RELATED WORK

In general, there are two main approaches for question classification which are rule-based and learning-based approaches (Li & Roth, 2004). There are also some studies that uses the hybrid approach, which combines both rule-based and learning based approaches together (Loni, 2011). In the rule-based approach, rules are manually handcrafted which are used for matching the questions with the answer types (Prager et al., 1999) (Radev et at., 2002). Although this approach is effective, it is time-consuming and requires a lot of works since there are many rules that need to be considered and manually created. Furthermore, it is difficult to adapt if the answer type taxonomy changes such as when new categories are added. There are several web-based QA systems that have relied on such rules with limited success (Radev et at., 2002). There are also studies that employed rule-based approach for question classification in other languages such as in Indonesian (Gusmita et al., 2014) and Arabic (Al-Shawakfa, 2016).

In the learning based approach, classifier models built from machine learning algorithms are used to classify or predict the answer types from the questions. Most of the recent works on question classification are based on this approach as well as the hybrid approach. Supervised learning methods are used to learn classifiers from training data consisting of labeled questions according to the answer type taxonomy. Various works that use this approach mainly differ in the machine learning algorithms used and the features that are extracted from questions. For example, Li and Roth (2002)used the Sparse Network of Winnows (SNoW) learning architecture (Roth, 1998) to learn a question classifier for an open domain English question sets in Text Retrieval Conference 10 (TREC 10) competition. They used several types of features such primitive (words, pos tags, non-overlapping phrases), named entities, head chunks and semantically related words (words that often occur with a specific question class). The experimental results show that their approach can produce good results as compared to the rule-based approach.

Most similar to our work is the study by Abdelnasser et al. (2014) which learned an Arabic question classifier for the holy Quran as part of their work to build an Arabic QAS for the Quran. They make use of a corpus of Arabic questions for training and testing their classifier. The Support Vector Machine (SVM) (Cortes & Vapnik, 1995) is employed as the learning algorithm. However, no information is given regarding the features that are used for the question classification learning process. The study by Zhang and Lee (2003) compared the question classification accuracy of SVM with several machine learning algorithms such as SVM, Decision Tree, K-nearest Neighbor, Naïve Bayes and SNoW. Even though they used shallow features such as bag-of-words and bag-of-ngrams, the results appeared to be comparable to the well-established use of rules and rules-like features. There are also learning based question classification works that are done on languages other than English and Arabic. Examples are the study by Zhang and Zhao (2010) which employed SVM for Chinese datasets, and the study by Devi and Dua(2016) which compared the classifier performance of Nearest Neighbor and K-nearest Neighbor implemented on Hindi datasets.

In the hybrid approaches, which is the combination of both rule-based and learning approaches, the learned classifiers are augmented with the hand-built rules. For example, in the work by Silva et al. (Silva et al., 2011), they first used some pre-defined rules to match the questions and then used the matched rules as features in their learning classifier. SVM with linear kernel is utilized to train their classifier. They also used the TREC datasets and obtained the highest accuracy in their results as compared to their competitors. In terms of the extracted features, they employed headwords in combination with hypernyms. Sherkat and Fahoodi(2014) focused on hybrid question classification approach in a closed-domain QAS for the Persian language. They designed about up to three regular expression rules for each class of questions and used SVM to learn the classifier.

III. APPROACH

For our work in classifying Malay questions for the Quran, the learning approach is adopted. Specifically, we define the question classification task as follows:

Given a question in a Malay, q, it will map q to c, where c is an element in a set of predefined categories, C.

C is also known as the answer type taxonomy. In order to map or classify the questions to the specific taxonomy entry, a classifier is learned with a machine learning algorithm. To train and test the classifier, a corpus of Malay questions for the Quran is prepared where each question is manually labelled with the correct answer type classification.

Answer Type Taxonomy

The answers to the questions that are posed to the Quran can be classified into a set of categories which is known as the answer type taxonomy or answer type ontology. There are many question taxonomies that have been used in various question classification works. The most commonly used question taxonomy was proposed by Li and Roth (2002) which is a two-layered answer type taxonomy that contains six coarse-grained categories (ABBREVIATION, ENTITY, DESCRIPTION, HUMAN, LOCATION and NUMERIC VALUE) and 50 fine-grained categories, as shown in Table-1. Each coarse-grained category contains a non-overlapping set of fine-grained categories.

Table. 1 Two-layered Answer Type Taxonomy from (Li &Roth, 2002)

Coarse	Fine
ABBR	abbreviation, expansion
DESC	definition, description, manner, reason
ENTY	animal, body, color, creation, currency, disease/medical, event, food, instrument, language, letter, other, plant, product, religion, sport, substance, symbol, technique, term, vehicle, word
HUM	description, group, individual, title
LOC	city, country, mountain, other, state
NUM	code, count, date, distance, money, order, other, percent, period, speed, temperature, size, weight

For this study, which is the initial part of our work, we have adopted this taxonomy but we focus only on the coarse-grained category and thus, our taxonomy is a one-layered (flat) taxonomy. Furthermore, we omit the ABBREVIATION (ABBR) category due to its irrelevancy for the Quran domain. Instead, this category is replaced with the EVENT category, which we believe is the suitable category for questions such as: *Bilakah di bulan Ramadan*Revealed?*PadamalamapakahAllah menurunkanAlQuran?* (In what night that Allah revealed the Quran?) *Apakahnamamalam yang ganjaransamadenganseribubulan?* (What is the name of the night whose reward is equal to a thousand month?). The HUMAN (HUM) category is extended to cover not just human but every creations of Allah which include angels, Satan, animals and other *AlQuranditurunkan?* (When during Ramadan the Quran is living creatures. Consequently, the name HUM has been changed to CREATION. Table-2 lists the six category answer types in our taxonomy and gives some example of questions of each category.

Table 2. The Adopted Taxonomy and Sample Questions

Category	Example Questions
EVENT	*Bilakah di bulan Ramadan AlQuranditurunkan?* *Padamalamapakah Allah menurunkanAlQuran?*
DESCRIPTION	*Kenapakahmanusiainidiciptakanoleh Allah?* *ApakahkelebihanmalamLailatulQadar?*
ENTITY	*Apakah yang diarahkanoleh Allah supayadibuatolehNabiNuh?* *Daripadaapakah Allah menciptaNabi Adam?*
CREATION	*Siapakah yang membawawahyukepadaRasulullah?* *SiapakahRasul yang dipanggildengannamaibunya?*
LOCATION	*Dimanakah Allah berbicaradenganNabi Musa?* *Di negerimanakahtempatRatuBalqisdankaumnya?*
NUMERIC	*Berapalamakahtempohedahseorangisteri yang telahdiceraikan?* *Berapalamakah masa Allah menciptalangitdanbumi?*

IV. DATASET

The Malay Quranic question corpus is built from questions that are related to the holy Quran or the answers are to be found in the Quran. These questions are gathered from the Web from multiple source and are then merged to create the corpus. Since there are insufficient existing resources specifically developed for Malay Quran questions and answers, questions are also gathered from English resources and then translated to Malay. There are five web resources that were used for our question collections which are “Ask Islam” (“Ask Islam”, n.d.), “Back to Jannah” (“Back to Jannah”, 2018), “*Info Dan SoalJawabMengenai Al-Quran*” (“Info Dan SoalJawabMengenai Al-Quran”, 2011), “*Suara Islam, SoalJawabTentang Al Quran*” (“Suara Islam, SoalJawabTentang Al Quran”, 2016) and “Turn to Islam Community, Questions-on-Quran” (Basheerpkm, 2007). The collected questions were then selected to ensure the validity of the questions and their answers are related and mentioned in the Quran. Finally, the selected questions are manually labelled according to the six categories specified in the answer type taxonomy. The total number of questions labelled are 319. The distribution of the questions among the categories is shown in Figure-1. The highest number of questions (77) is in the Creation category, followed by Numeric (58), Description (52), Entity (52), Location (44), and Event (36).

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

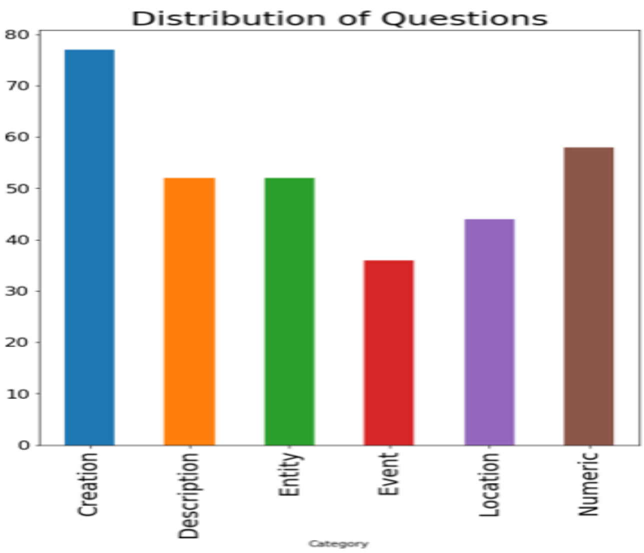

Fig.1 Questions Distribution according to Categories in the Malay Question Dataset

V. EXPERIMENTS AND RESULTS

In this particular study, for all questions in the dataset, the bag-of-words and n-grams (with n=1) models are chosen in selecting the features for the questions vector representations. In addition, to reduce the number of words in the vector space model, all stop words are removed from the questions. The Malay stop words (Abdullah et al., 2005) are employed for this purpose. Finally, the term frequency-inverse document frequency (TFIDF) weighting schemes is used to represent the selected text features in the vector space. In TFIDF, a higher weightage is assigned to a word based on two conditions. First, the word has a high occurrence in a question and second,the word has a low occurrence in other questions.

Three classifiers are implemented for our Malay questions classification task using three different machine learning classification algorithms, namely Naïve Bayes, Random Forest and Support Vector Machine (SVM). Basically, each algorithm represents a different machine learning approach. Naïve Bayes is chosen based on its simplicity and scalability, and can perform surprisingly well in practice. SVM is a more complex algorithm but can model non-linear decision boundaries. It is also quite robust against overfitting, especially in high-dimensional space like in the question and text classification problems. In this paper, the SVM kernel function that is adopted is the linear kernel with default parameter configuration. The linear kernel is used (instead of the non-linear kernel SVM) because it is often recommended for text classification. In contrast with Naïve Bayes and SVM, Random Forest is chosen because it is representing the ensemble learning algorithm in our experiment. Random Forest uses decision tree which also perform well in practice. In this work, all experiments are conducted using the Python Jupyter Notebook environment and the Scikit-learn machine learning library for the Python programming language. In order to evaluate the performance of the three classifiers, we opted for the stratified cross-validation instead of the data holdout (train/test split) strategy. According to Witten et al. (2017), cross-validation is the standard evaluation technique in situations where only limited data is available as in our case. Meanwhile, stratification is adopted to ensure that each of the data folds in the cross-validation contains similar proportions of all of the six different categories of questions. In other words, each fold should be a good representative of the whole questions dataset. In all of the experiments, the number of folds chosen for the cross-validation is ten.

The first performance metric that is evaluated is the overall classification accuracy of the classifiers. Accuracy is the proportion of the correct prediction achieved by a classifier, which is defined as follows:

$$\text{Accuracy} = \frac{\text{\# of correct predictions}}{\text{\# of predictions}}$$

Table-3 shows the average accuracies of the three classification algorithms where for each algorithm, the results from the 10-fold stratified cross-validation method are averaged. It can be seen from the table that with the average classification accuracy of 86.8%, the model based on SVM outperforms both Naïve Bayes and Random Forest methods, while Random Forest (84.6%) is better than Naïve Bayes (81.4%).

Table. 3 Mean Accuracies Comparison

Model Name	Mean Accuracy
Naïve Bayes	0.814313
SVM	0.867862
Random Forest	0.846472

Even though SVM achieves the highest accuracy, however, evaluating a classification model based on its accuracy alone is not sufficient in determining it as the best classifier model. Therefore, we need to further investigate on how well do the three classifier models perform on each of the categories of answer types. For this purpose, the performance metrics that are evaluated are the precision, recall and F1 score. Table-4, Table-5 and Table-6 show the evaluation results of Naïve Bayes, SVM and Random Forest respectively, for all of the six categories of answer types.

Table. 4 Classification Performance of Naïve Bayes

	Precision	Recall	F1-score
Creation	0.74	0.89	0.81
Description	0.75	0.69	0.72
Entity	0.60	0.46	0.52
Event	0.86	0.67	0.75
Location	0.89	0.73	0.80
Numeric	0.79	1.00	0.88

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Table. 5 Classification Performance of SVM

	Precision	Recall	F1-score
Creation	0.95	1.00	0.97
Description	0.67	0.92	0.77
Entity	0.89	0.62	0.73
Event	0.88	0.78	0.82
Location	0.90	0.82	0.86
Numeric	0.93	0.93	0.93

Table. 6 Classification Performance of Random Forest

	Precision	Recall	F1-score
Creation	0.84	0.84	0.84
Description	0.53	0.77	0.62
Entity	0.64	0.54	0.58
Event	1.00	0.78	0.88
Location	1.00	0.73	0.84
Numeric	0.94	1.00	0.97

VI. DISCUSSION

Comparing the results of precision, recall and F1 -score as shown in these tables, it can clearly be seen that SVM and Random Forest are better, as compared to Naïve Bayes, in classifying the questions into all of the categories of answer types. Between SVM and Random Forest, SVM appears to be the better classifier model for the Creation, Description, Entity and Location categories. However, Random Forest is the better classifier than SVM for the Event and Numeric category. It can also be noted that Naïve Bayes performs better than Random Forest only in the Description category but it is still below than that of SVM. Figure-2 depicts the results of averaging the scores of precision, recall and F1-score of the three models for all of the six categories. From these results, it can be seen that, on the average, SVM is the superior algorithm for the Malay questions classification task.

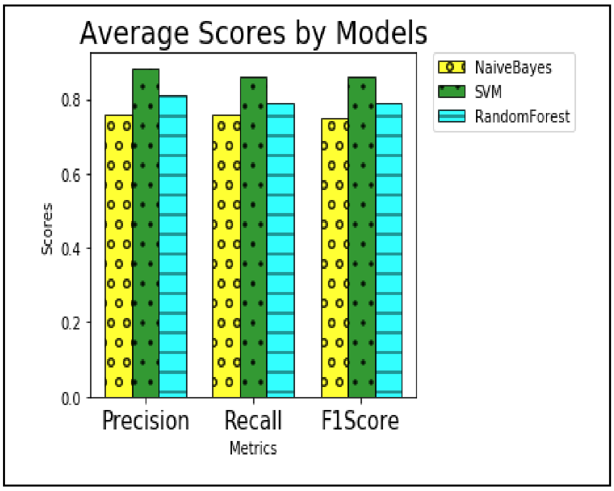

Fig. 2 Average Classifier Model Performance on all Categories of Answer Types

Although the size of our dataset is relatively small, the results are adequately satisfactory for the Malay question classification problem. Nevertheless, we can see that the performance of all of the classifier models are still low for the Description and Entity categories. All of the three classifier models achieve F1-score of below than 80% for these two categories. A further analysis of these two categories are done to identify the questions that are misclassified by all of the classifiers. These questions are shown in Table-7. All of the three algorithms failed to correctly classify these questions. As expected, these questions which have the "*apakah*" ("what is") question word are not easily predicted by the classifiers because they may be associated with many expected answer types. In contrast to the easier questions such as questions with "*siapakah*" ("who"), they are most likely to be related to the creation (person) type. Therefore, the "*apakah*" questions provide less information and impose a heavy burden on the classifiers using the bag-of-words strategy. Furthermore, these categories are relatively small categories in our dataset and perhaps they can be handled by re-training with a larger number of data in these categories.

VII. CONCLUSIONS AND FUTURE WORK

Question classification is an important task in QAS especially in the Malay language where the number of research works is still limited. In this paper, we presented a Malay question classification task using the learning based approach. For developing the classifier model, we have employed and evaluated three different machine learning classification algorithms which are Naïve Bayes, Random Forest and Support Vector Machine (SVM) and presented the results. All of the classifier models demonstrate good performance in term of accuracy, precision, recall and F1-score, with SVM has a superior overall performance.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Table. 7 Questions that are misclassified by all Models (related to Description and Entity categories)

Actual	Predicted (Error Type)	Questions
Description	Entity (False negative for Description)	*Apakah yang dimaksudkandengan Surah Madaniyyah?* (What does it mean by Madaniy chapters)
Event	Description (False positive for Description)	*Padamalamapakah Allah menurunkanAlQuran?* (In what night that Allah revealed the Quran?)
Entity	Description (False negative for Entity)	*MenurutnabiKhidirapakahbenda yang ada di bawahtembok yang hendakruntuhitu?* (According to prophet Khidir what was under the nearly collapsed wall?
Entity	Numeric (False negative for Entity)	*Apakah surah yang menjelaskanhukumperwarisan?* (What chapter that explain inheritance law?)

This study has offered us a good insight on how to proceed with the building of a model for Malay question classification. Firstly, we plan to make improvement to the current model by building a larger Quran question corpus since the current corpus contains only 319 questions. In addition, we plan to build a classifier that is trained using a hierarchical (fine-grained) answer type taxonomy instead of the flat (coarse-grained) answer type taxonomy that has been employed in this work. A fine-grained taxonomy is more useful for the question processing task in a realistic QAS. It is also interesting to experiment with different types of kernels (such as RBF and polynomial kernels) as compared to the linear kernel for the SVM algorithm that has been used in the current work. Furthermore, instead of using just the bag-of words strategy, we plan to further investigate and identify useful features of the Malay language that can be used to correctly classify the questions. This is especially true for the difficult type of questions such as those that we have identified in this work.

REFERENCES

1. Abdullah, M. T., Ahmad, F., Mahmod, R., &TengkuSembok, T. M.(2005). Improvement of Malay Information Retrieval Using Local Stop Word. *International Advanced Technology Congress*, Dec 6-8, Malaysia.
2. Ahmad, N. D., Bennett, B., & Atwell, E. (2017). Retrieval Performance for Malay Quran. *International Journal on Islamic Applications in Computer Science And Technology, Vol. 5*, Issue 2, 13-25
3. Al-Shawakfa, E. (2016). A Rule-Based Approach to Understand Questions in Arabic Question Answering. *Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2*, No. 3, pp. 210–231.
4. Ask Islam. (n.d.). Retrieved July 30 2018, fromhttp://www.askislam.org/literature_and_languages/quran/index.html.
5. Back to Jannah. (2018). Retrieved July 30 2018, from https://backtojannah.com/100-questions-on-quran/.
6. Basheerpkm. (2007, March 13). Retrieved July 30 2018, http://turntoislam.com/community/threads/100-questions-on-quran.10052/.
7. Cortes, C. & Vapnik, V. N. (1995). Support vector machines.*Machine Learning 20*: 273–297.
8. Devi, R., &Dua, M. (2016). Performance Evaluation of Different Similarity Functions and Classification Methods Using Web Based Hindi Language Question Answering System. In *Procedia Computer Science, Vol. 92*, 520–525.
9. Gusmita, R. H., Durachman, Y., Harun, S., Firmansyah, A. F., Sukmana, H. T., &Suhaimi, A. (2014). A rule-based question answering system on relevant documents of Indonesian Quran Translation. *In International Conference on Cyber and IT Service Management*, CITSM 2014, 104–107.
10. Hirschman, L. &Gaizauskas, R. (2001). Natural language question answering: the view from here. *Natural Language Engineering 7* (4): 275-300. Cambridge University Press.
11. Info Dan SoalJawabMengenal Al-Quran. (2011, August 25).Retrieved July 30 2018, https://shafiqolbu.wordpress.com/2011/08/25/info-dan-soal-jawab-mengenal-al-quran/.
12. Jurafsky, D.,& Martin, J. H. (2009). Speech and Language Processing. Prentice Hall.
13. Mishra, A., &Jain, S. K. (2016). A survey on question answering systems with classification. *Journal of King Saud University-Computer and Information Sciences 28*, 3, 345–361.
14. Li, X.,& Roth, D. (2002). Learning question classifiers. *In Proceedings of the 19th international conference on Computational linguistics, COLING '02 Association for Computational Linguistics,* 1–7.
15. Li, X.,& Roth, D. (2004). Learning question classifiers: The role of semantic information. In *Proceedings International Conference on Computational Linguistics*, 556–562.
16. Loni, B. (2011). A Survey of State-of-the-Art Methods on Question Classification. *Literature Survey*. Published on TU Delft Repository.
17. Molla, D., &Vicedo, J. L. (2007). Question answering in restricted domains: An overview. *Computational Linguistics 33*(1):41-61.
18. Prager, J., Radev, D., Brown, E., Coden, A., &Samn, V. (1999). The use of predictive annotation for question answering in TREC8, *NIST Special Publication 500-246: The Eighth Text REtrieval Conference* (TREC 8).
19. Radev, D., Fan, W., Qi, H., Wu, H.,& Grewal, A. (2002). Probabilistic question answering on the web, In*Proceedings of the eleventh international conference on World Wide Web (WWW2002),* Hawaii.

20. Roth, D. (1998). Learning to resolve natural language ambiguities: A unified approach. In *Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence, AAAI '98/IAAI '98. American Association for Artificial Intelligence*, 806–813.
21. Santosh, K. R., Singh, S., & Joshi B. P. (2010). A semantic approach for question classification using wordnet and wikipedia. *Pattern Recognition Letters Volume 31*, Issue 13:1935–1943.
22. Sherkat, E.,&Farhoodi, M. (2014). A Hybrid Approach for Question Classification in Persian Automatic Question Answering Systems. *In Proceedings of 4th International Conference on Computer and Knowledge Engineering (ICCKE)*, Mashhad, 279-284.
23. Silva, J.,Coheur, L., Mendes, A. C., &Wichert, A.(2011). From symbolic to sub-symbolic information in question classification. *Artificial Intelligence Review, 35*(2):137–154.
24. Suara Islam, SoalJawabTentang Al Quran. (2016, November 2).Retrieved July 30 2018, https://suaraislam.net/2016/11/02/soal-jawab-tentang-al-quran/.
25. Witten, I., &Frank, E. (2017). Data Mining Practical Machine Learning Tools and Techniques. Morgan Kaufmann, San Francisco, CA. Fourth edition.
26. Zhang, D.,&Lee, W. S. (2003). Question classification using support vector machines. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval, SIGIR '03*, 26–32.
27. Zhang, K.,&Zhao, J. (2010). A Chinese question-answering system with question classification and answer clustering. In *Proceedings of the 7th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2010, Vol. 6*, 2692–2696.