Journal of Engineering and Applied Sciences 13 (12): 4468-4475, 2018
ISSN: 1816-949X

Classification of Holy Quran Translation Using Neural Network Technique

Suhaib Kh. Hamed and Mohd Juzaiddin Ab Aziz
Center for Artificial Intelligence Technology (CAIT),
Faculty of Information Science and Technology, University Kebangsaan Malaysia,
43600 Bangi, Selangor, Malaysia

Abstract: The Holy Quran is the most significant religious text which is followed by the believers of the Islamic religion. The translations of the Quran are the interpretation of its meaning in different languages to assist those who are not familiar with the Arabic language. The Holy Quran consists of 114 Chapters (Surah) and 6236 Verses. These 114 Chapters have different length. The verse is the smallest segment of the texts of the Quran. The Holy Quran as a book is not classified on topics and its verses describe many topics and many verses even from different chapters converge within the same theme. The number of verses and chapters may share similar topics such as faith and morality. One of the solutions to tackle this issue is the Quran classification. Thus, the aim of this study is to classify the Quranic verses by using the Neural Network (NN) classifier based on the predefined topics in order to provide the readers the relevant Quranic verses depending on their need. This research used the most popular and the widespread of the Holy Quran translation in English language by Abdullah Yusuf Ali as the reference dataset. In this regard, the neural network classifier will address this issue through classifying the Al-Baqarah Surah that represents the Quran into two categories which are the Fasting and Pilgrimage topics. Finally, based on the F-measure, the evaluation of Al-Baqarah classification by using NN showed a level of approximately 90%. This proves that NN has succeeded in presenting a encouraging result in this critical area.

Key words: Holy Quran, information retrieval, machine learning, neural network classification, text, issue

INTRODUCTION

Is one of the most common and most significant problems in the fields of natural language processing, data mining and machine learning because of the large and growing information available on the internet or existing as the digital forms like documents. The aim of documents classification is to assign each document to one or more defined categories, depending on their contents. The Holy Quran consists of 114 Chapters (Surah) and 6236 Verses (Ayah) which is written in the Arabic language. These 114 Chapters have dissimilar length. The verse is the smallest segment of the texts of the Quran and the verse might consist of a one or more of the sentences. A group of verses will form the chapter. The availability of the Quran translations making the accessibility of written knowledge that related to Quran becomes faster and less complicated, especially for those who are not familiar with the Arabic language. The Holy Quran as a book is not classified based on topics and its verses describe many topics and many verses even from different chapters converge within the same subject. The number of verses and chapters may share similar topics such as Faith and Morality (Hamed and Ab Aziz, 2016). One of the solutions to tackle this issue is the Quran classification. Thus, there is an essential need for the Quran classification to classify the Quran based on its content. Several Quran classification algorithms have been developed among them Naive Bayes, support vector machines, k-nearest neighbors to classify any verses to the predefined subjects. Since, this research dealing with a sacred script, therefore, the classification must be context sensitive. Moreover, there is a need to comprehend the current and possible classifications for the Quranic verses (Al-Kabi *et al*., 2013). The Holy Quran as the text is a exquisite and free of weaknesses and difficult to formulate by the human being and is considered from the linguistic perspective is complex in its context and structure and as well is varied in its formulations of expression (Nassimi, 2008). The Artificial Neural Network (ANN) technique efficiently uses in different domains and has showed a promising performance in the field of texts classification because it has the ability for recognizing the complex patterns existing in the text (Ramlall, 2010; Mohammed and Omar, 2012; Patra and Singh, 2013). Therefore, the need of classification is important for better information management and to determine the passage of verses that

Corresponding Author: Suhaib Kh. Hamed, Center for Artificial Intelligence Technology (CAIT), Faculty of Information Science and Technology, University Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

contain relevant verses to the user's need and then reducing the computation of the search space in order to reduce the irrelevant verses. Consequently, this research proposes a Quran classification using ANN technique in order to address all these problems that mentioned previously with respect to the Holy Quran.

Literature review: Number of research have been produced to develop and facilitate the classification task on the Quran. Although, the most of these studies have been dedicated to the Arabic reader it has been presented several pieces of research discuss the classification process of the translations of the Quran. Al-Kabi *et al.* (2013) have examined the effectiveness of four common classification methods. Support vector machine, decision tree, Naive Bayes and k-nearest neighbors algorithm in terms of classifying various verses of the Holy Quran based on their topics. They suggested a pre-processing phase for the classification task by implementing the steps of removing the Arabic diacritics and also removing the Quranic symbols. The results of their evaluation demonstrated that the Naive Bayes classifier showed a high level of accuracy and the lowest rate of error. On the contrary, the decision tree classifier presented the lowest level of accuracy whereas the error rate was the highest among these classifiers. In that respect Shahzadi and Sawar (2012) have presented a research capable of classifying the Quranic verses of English translation to one or more predefined categories based on linear classifier depending on the concepts of the Quran through using a semantic network. They mentioned that most of the previous studies in that respect conducted only for some Quranic chapters or the classification based on term frequency. Their researches focuse on two main phases which are building the semantic network of Quranic concepts and classify the Holy Quran based on this semantic network. In the same vein, Nassourou (2011) presented a methodology of reconstructing the chronology of Quran (revealed in Mecca or in Medina) based on machine learning techniques. A hybrid statistical classifier which is a combination of Bayesian and distance-based classifiers have been employed in order to reach at plausible dates of revelation in accordance with the traditional Islamic scholars. In this regard, Al-Kabi *et al.* (2005) have proposed a system that classifies the verses of the Fatiha and Yaseen Chapters in the Quran according to the classifications of Islamic scholars. A system (classifier) has been designed and implemented to categorize the different verses in each Chapter (Surah). This system fully normalizes the verses in the first stage and then the verses are classified into classes for which they have the highest score. The accuracy of the system can be improved as they mentioned if a more powerful stemmer is used.

MATERIALS AND METHODS

Dataset: This research used the English translation of the Holy Quran by Abdullah Yusuf Ali (YA) as the reference dataset which is considered as the most popular translation of the Quran among others translations that covers a large number of readers of the Holy Quran in the English language. Al-Baqarah Chapter was used as a sample of the Holy Quran because the Al-Baqarah Surah is the longest chapter of the Holy Quran, it contains a large number of commands, prohibitions, judgments and moralizing. Since, the Fasting and the Pilgrimage are the two pillars of the five pillars of Islam. Therefore, this study focuses on the verses of Fasting and Pilgrimage that were mentioned in Al-Baqarah Chapter. Table 1 shows the verse's numbers of Fasting and Pilgrimage and none of these categories that belong to the Al-Baqarah Surah.

Document classification based on neural network technique: Since, the research scope focuses on the verses of Fasting and Pilgrimage, therefore, Al-Baqarah Surah will be classified into two classes, Fasting and Pilgrimage. The main goal of document classification is to reduce the searching space by identifying the passages of information that are relevant to the particular topic (Baharudin *et al*., 2010). This research used the Neural Network (NN) based on supervised learning technique to classify the verses of Al-Baqarah Surah. NN has been effectively used in many areas of artificial intelligence for example, NLP, pattern recognition and classification tasks (Mohammed and Omar, 2012). Moreover, NN technique produces better results in the complex domain than others classification algorithms (Patra and Singh, 2013). The most important type of Neural Network technique is Back-Propagation Neural Network (BPNN) which was employed in this research. BPNN could attain higher accuracy because it has the ability for learning and improving itself (Vora and Shruti, 2009). The classification of the Al-Baqarah Chapter based on neural network technique was carried out in the WEKA. Although WEKA includes different machine learning algorithms, it contains many tools for data pre-processing (Witten and Frank, 2005). The process of NN classification consists of the following phases, training set, filtered classifier that combines a filter and the neural network classifier and finally, the evaluation of the generated module based on the testing set. Figure 1 illustrates the process of verses classification in WEKA based on NN classifier.

Training set: Firstly, this data set is converted to the ARFF format to be understandable for the WEKA and to perform the verses classification. The training set consists of the list of 150 instances of verses sharing a set of attributes where this training set represents 80% of the

Table 1: The verses Fasting and Pilgrimage of Al-Baqarah Chapter

Variables	Fasting	Pilgrimage	The rest of the verses
Verses numbers	183, 184, 185, 186, 187	125, 158, 196, 197, 198, 199, 200, 201, 202, 203	-
Number of verses	5	10	271
Total number verses of Al-Baqarah Surah	286		

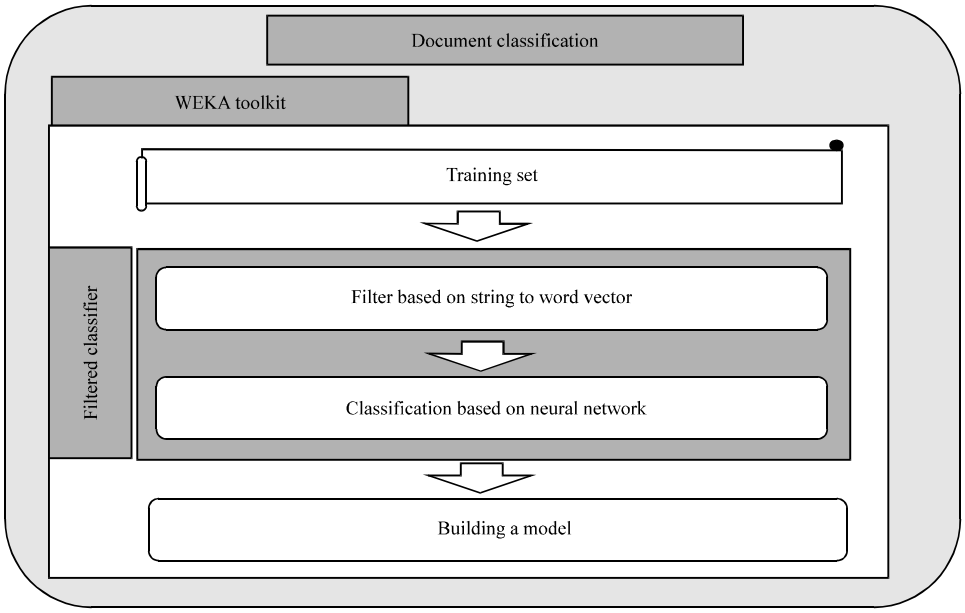

Fig. 1: Verses classification

total number of the dataset which means that the 20% represents the remaining data that will be used for the testing set. The examples that are used in this training set consisting of the English translation of Quranic verses and the interpretation of the Quran (English translation of Tafseer). The training set has three classes, 50 examples for the Fasting class, 50 examples for the Pilgrimage class and 50 examples for the none class that are not related to any of these mentioned classes. Each example is labeled with a value represents its class. This phase is considered the most important phase because based on these examples, the Neural Network (NN) classifier could learn and then able to predict the classes of the verses easily and hence, building the best classifier that mainly depends on the quality of these examples. The training of the NN classifier is achieved by means of using the back propagation learning rule based on supervised learning.

Filtered classifier: The classifier and the "String to word vector" filter could be combined to build the filtered-classifier in WEKA toolkit. This combination includes neural network classifier based on back propagation network and the filter which this filter could be able to deal with string attributes directly without the need for the filter in an isolated stage to process and transform the verses. In WEKA, the task of the filter is similar to the task of data pre-processing phase which includes several tasks that use to pre-process the data. The pre-processing technique based on the filter is used to reduce the complexity of the document in order to generate a clear text and then extracted the feature set based on feature selection method by transforming the verses into a vector representation of these verses. This vector is represented a bag-of-words (Ghiassi *et al*., 2012; Aggarwal and Zhai, 2012). The researchers prefer the filter methods in text classification to enhance the accuracy level of the classifier and then reducing the computation time (Uysal and Gunal, 2014).

String to word vector filter: The raw verses are data and firstly should transform them into a form appropriate for learning by generating a dictionary of terms from all these verses in the training set and assign a numeric attribute for each term which refers to the importance of terms in the text that will be fed later to the NN classifier, this process is implemented by using the filter "String to word vector". This filter is consisting of the pre-processing techniques such as the normalization, stop words removal and the stemming where these techniques are responsible for removing the redundancy and irrelevancy and noisy features from the text and hence to generate an initial set of features. Since, the dimensionality of the initial features space is still very high because the initial feature set consists of thousands of features (terms). Therefore, the

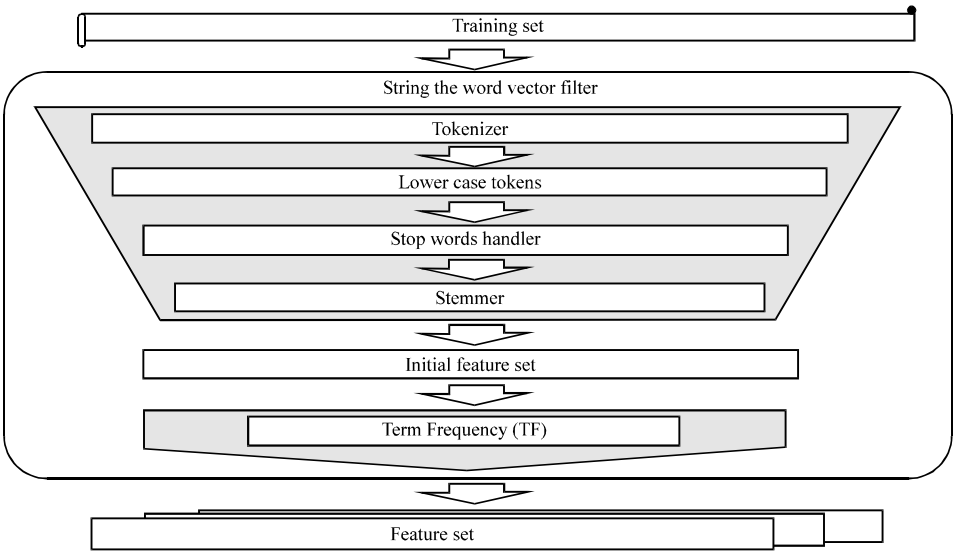

Fig. 2: String to word vector filter

methods of feature reduction should be applied to remove the irrelevant features from the initial feature set in order to enhance the effectiveness of the NN classifier by increasing the susceptibility of this classifier, through feeding these attributes to the classifier in order to lead to more understanding of the learning process of classifying verses. One of the methods of the feature reduction is the feature selection method which is performed by applying the feature weighting scheme.

The feature weighting scheme that applied by this research is Term Frequency (TF) technique (Babu *et al*., 2014) in order to rank the features in the initial feature vector and then choose a number of the high scoring features as a new feature subset (vector representation of verses) which is considered the distinctive attributes in classifying the verses. These feature vectors are then used to train the neural network. Based on the outcomes of the conducted experiments in this study and according to the best result from these outcomes that attained in classifying of the verses, the following tasks in the string to word vector filter have been adjusted such as illustrated in Fig. 2.

Tokenizer: Tokenization is the process of dividing a sequence of text in the verses into words, phrases, based on n-gram technique.

Lower case tokens: All the word tokens are converted to lowercase before being added to the feature set. Based on the experiments the normalization technique improved the classification accuracy (Patki and Kelkar, 2013; Uysal and Gunal, 2014).

Stop words handler: There are many words in the verses are repeated frequently and essentially do not carry any information (Aggarwal and Zhai, 2012). Thus, their presence in verses classification will present a lack of understanding properly to the content of the verses.

Stemmer: The aim of stemming is to minimize the words to their roots which could be easily used to differentiate the words.

Term Frequency (TF) transform: This technique represents the feature selection method which is responsible for generating the features set of the initial features set depending on the word frequencies by calculating their weight based on the following Eq. 1:

$$TF = \log (1+fij) \quad (1)$$

where, fij is the frequency of word i in verse j (instance). According to the significance of the terms in the verses, this process includes assigning a weight to each term which indicates the relative importance of the term in verse and this depends on the word frequency in the verses. Therefore, the most repeated words in the document (its term frequency is high) are regarded a more significant in this document (Wang *et al*., 2012). If a word (except stop words) appears within a particular category, the word should be considered as a feature or discriminator of this category. For example, “Fasting” or “Ramadan” frequently appears in the Fasting category.

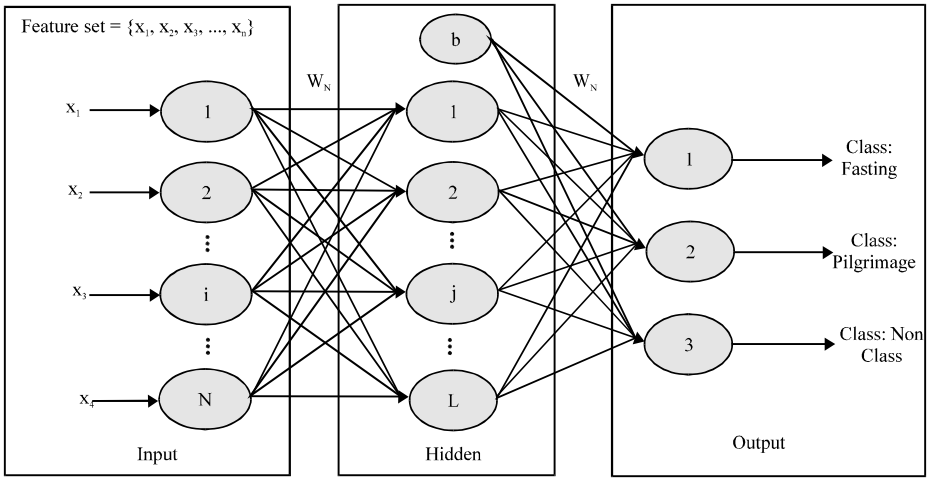

Fig. 3: The structure of neural network classifier

Neural network classifier: The major task of the Neural Network (NN) technique is to automatically learn to recognize component patterns and make intelligent decisions by many examples of verses that are presented in the training set, through an interactive process of adjustments applied to its synaptic weights and bias levels. NN classifier is a network of units where the input units usually represent the features set (the terms of the verses that extracted by the string to word vector filter) verses, the output units represent the classes (Fasting and Pilgrimage). In order to empirically estimate the weights, firstly, the string to word vector filter will provide the networks with the features set of distinctive terms that extracted from the training set. The function of NN classifier is to compare the weights of terms of the Quranic verses that should be classified with the weights of predefined classes based on the training set to determine each verse to its class. The learning process through training the neural network classifier is not merely a matter of saving the mapping of the relations between the inputs and the outputs of the presented instances but in point of fact, it is to deduce and extract the distinctive features and the internal rules from these instances which are unclear to the normal user. De Houwer *et al.* (2013). This research used the neural network classifier based on the Back-Propagation Network (BPN). The BPN is a multilayer neural network consisting of the input layer, the hidden layer and the output layer as it is shown in Fig. 3. All the neurons of the hidden and the output layers are non-linear units with the sigmoid function as the activation function. These neurons in the hidden and output layers are connected with biases, the aim of using the bias is similar to the function of the weight. There are two major stages of the backpropagation learning task, the first stage is a forward phase and the second stage is a backward phase. With regard to the forward stage, the input patterns transfer forward through the networks from one layer to the other and ultimately the actual output of the network will be generated. If there is any difference between the generated actual output and the desired output, this obviously indicates to an error. This requires some procedures to be taken in order to minimize this variance in order to reduce the value of the error that represents the difference between the actual output and the desired output, the generated error value will be propagated in a backward direction. With respect to the backward stage, it should be performed small readjustments in the weights of the network in order to decrease the sum squared errors. Back-propagation learning process has been successfully applied regarding solving many difficult problems (Mohammed and Omar, 2012; Aljawfi *et al.*, 2014). The process of the Back-Propagation Network (BPN) can be outlined in the following steps:

- The back-propagation network has randomly set initial weights in the range of -0.5, 0.5
- Updating the weights to obtain the output that is identical to the training set
- Computing the value of error based on the following Eq. 2:

$$\text{Error e} = \text{Desired output-Actual output} \quad (2)$$

- The weights of networks should be readjusted to reduce the value of error

Build the model: The building of classifier depends on the testing set that used to evaluate this classifier. It is an important matter to have a testing set that incorporates

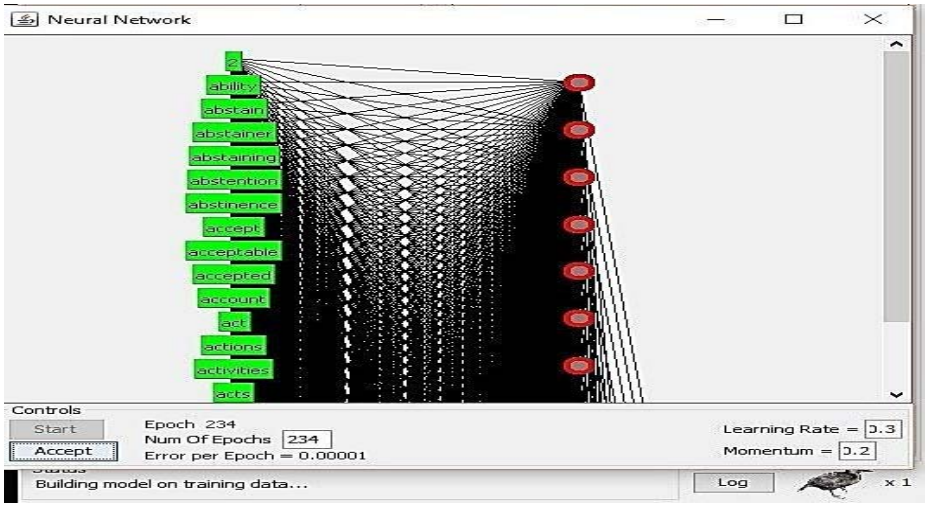

Fig. 4: Verses classification based on training set

the same concepts as the training set but differs with a training set in terms of the instances. Except that the classification will end up with poor performance and the result is not satisfactory. The testing set consists of 30 instances of verses that made up 20% of the dataset. These instances include the Quranic verses and interpretation of the Quran (Tafseer) which are similar to the training set with the respect that are discussing same classes of Fasting, Pilgrimage and none class. These instances are divided as follows, 10 instances about fasting and ten instances about the pilgrimage and ten instances do not belong to either of them. According to the experiments that performed on the testing set based on the process of the Back-Propagation Network (BPN) in order to build the classifier regarding achieving high performance and obtaining the best result as illustrated in Fig. 4, the configuration of BPN classifier as following.

Hidden layers: This specifies the number of hidden layers of the neural network based on this Eq. 3:

$$\text{Hidden layer} = \frac{(\text{Attributes}+\text{Classes})}{2} \quad (3)$$

where attributes indicate to the number of terms (features set) of the input verses and the value of classes will be three based on the classes of this research.

Learning rate = 0.3. According to Witten and Frank (2005). If the value of learning rate is large then the search might overstep and miss the global minima. Therefore, the value of learning rate should be small in order to find the global minima, although, this value lead the progress toward the global minima might be slow.

Momentum = 0.2. As stated by Witten and Frank (2005) to improve the performance of the classifier should adding a small value of momentum for updating weights to smooth the search process (Table 2).

Training time: This represents the number of epochs to train the classifier based on the training set. Depending on the GUI, the classification process was terminated at epoch number 234 because the error rate was the minimum error obtained by the research at this epoch which is 0.00001. The error rate in the training set will give a good indication of future performance, therefore, good classifiers have low error rates (Witten and Frank, 2005).

Evaluation: The results of the experiments that conducted based on different NN classifiers have been evaluated based the evaluation metrics. The evaluation of the effectiveness of document classifiers is commonly conducted experimentally based on many metrics such as Precision (P), Recall (R) and F-score (Korde and Mahender, 2012; Ghiassi *et al*., 2012). The equations of precision and recall have to be as follows:

$$P = \frac{\text{Number of correctly classified verses}}{\text{Total number of classified}} \quad (4)$$

$$R = \frac{\text{Number of correctly classified verses}}{\text{Total number of correct verses from the data}} \quad (5)$$

Since, the precision and recall is calculated earlier, hence, it is possible to calculate the F-score function based on the following Eq. 6:

Table 2: Different experiments of NN

Classifier name	Training set	Filter parameters	Minimum error rate at epoch
NN1	45 instances (each class 15)	String to word vector without using feature selection technique	Minimum error rate (0.00002) at Epoch 521
NN2	120 instances (each class 40)	String to word vector based on TF-IDF	Minimum error rate (0.00001) at Epoch 239
NN3	150 instances (each class 50)	String to word vector based on TF-IDF	Minimum error rate (0.00002) at Epoch 113
NN4	150 instances (each class 50)	String to word vector based on TF	Minimum error rate (0.00001) at Epoch 234

$$\text{F-score} = 2\times\frac{\text{Precision}\times\text{Recall}}{\text{Precision}+\text{Recall}} \quad (6)$$

RESULTS AND DISCUSSION

Based on the testing sets, several experiments were executed to choose the best NN classifier. However, in this research it was mentioned to four important experiments in terms of varying results and the stages of the evolution of the performance of these classifiers which differ with regard to using the training set and as well as the settings of the filter as Table 2. Table 3 demonstrate the stages of some experiments of NN classification that conducted on several testing sets based on the configurations that were mentioned in Table 2 where it produced different results based on the evaluation metrics depending on the changes made to the setting of the classifier or based on the used training set. According to Table 3, NN4 showed a better accuracy of classification than others classifiers based on the F-score evaluation which is approximately 0.86 level. This is due to the NN4 classifier, takes into consideration the only using of TF technique as a feature selection without IDF technique because there are many of frequent terms that are repeated in the verses that refer to the classes of Fasting and Pilgrimage such as the "fast" and "pilgrimage". However, if the IDF technique is used as a feature selection and then the classifier will focus on all the terms that are not repeated in verses based on the IDF such as "oath" and "bounty" which might not refer to the required classes, it will lead to classify many irrelevant verses to these classes. This study has adopted the configuration of the Neural Network classifier (NN4) because the result of this classifier showed a high level of accuracy compared to all experiments conducted in this critical domain that is because whole the verses that are not related to these categories have been significantly reduced. Table 4 illustrates the classification of Al-Baqarah Surah to their classes where all the verses of Fasting and Pilgrimage were assigned to the correct class and the irrelevant verses have been reduced in each class and excluded 246 verses that are not related to the topics of Fasting and Pilgrimage. Whereas Table 5 shows the truth table of the classification of 5 neural network classification. The evaluation of the NN classification based on the F-score showed a high result which is approximately 90% level such is shown in Table 6 which this result is considered a very high in such a critical domain.

Table 3: Classifications evaluations based on testing set

Classifier name	Testing set	Recall	Precision	F-score
NN1	16 instances	0.38	0.38	0.37
NN2	25 instances	0.50	0.50	0.50
NN3	30 instances	0.64	0.64	0.63
NN4	30 instances	0.87	0.87	0.86

Table 4: Al-Baqarah Surah classification

Variables	Fasting	Pilgrimage	The rest of the verses
Verses numbers	3, 56, 159, 172, 173, 183, 184, 185, 186, 187, 222, 225, 226, 229, 230, 242, 266	5, 37, 46, 70, 77, 95, 105, 114, 125, 127, 158, 196 197, 198, 199, 200, 201, 202,203, 227, 239, 245, 282	-
Number of verses	17	23	246
Total number of verses of Al-Baqarah Surah			286

Table 5: Neural network classification truth table

Class	Truth Fasting	Truth Pilgrimage	Truth none
Predicated as Fasting	5	0	0
Predicated as Pilgrimage	0	10	0
Predicated as none	12	13	246

Table 6: The evaluation of Baqarah Surah classification

Recall	Precision	F-score
0.91	0.91	0.90

CONCLUSION

The Holy Quran discusses a large number of topics. This research focused on the topics of Fasting and Pilgrimage. Therefore, it was classified the Quran that is represented by Al-Baqarah Surah to the classes of Fasting and Pilgrimage. Where Al-Baqarah Surah was classified by using neural network technique and this technique proved highly efficient in classifying this critical domain through a number of experiments were carried out to improve the performance of this proposed method. Consequently, the verses that are irrelevant to the scope of the user's need it will be excluded and provide them only with Quranic verses that satisfy their needs.

ACKNOWLEDGEMENT

This research has been supported by Center for Artificial Intelligence Technology (CAIT), Faculty of Information Science and Technology, University Kebangsaan Malaysia.

REFERENCES

Aggarwal, C.C. and C.X. Zhai, 2012. A Survey of Text Classification Algorithms. In: Mining Text Data, Aggarwal, C.C. and C.X. Zhai (Eds.). Springer, Boston, Massachusetts, USA., ISBN :978-1-4614-3222-7, pp: 163-222.

Al-Kabi, M.N., G. Kanaan, R. Al-Shalabi, K.M.O. Nahar and B.M. Bani-Ismail, 2005. Statistical classifier of the holy quran verses (Fatiha and Yaseen Chapters). J. Applied Sci., 5: 580-583.

Al-Kabi, M.N., I.M. Alsmadi, H.A. Wahsheh and B.M.Abu Ata, 2013. A topical classification of Quranic Arabic text. Proceedings of the NOORIC2013 International Conference on Advances in Information Technology for the Holy Quran and its Sciences, December 22-25, 2013, Taibah University, Medina, Saudi Arabia, pp: 272-277.

Aljawfi, O.M., N.M. Nawi and N.A. Hamid, 2014. Enhancing back propagation neural network with second order con-jugate gradient method for fast convergence. Proceedings of the 1st International Conference on Recent Trends in Information and Communication Technologies (ICT'14), September 12-14, 2014, Universiti Teknologi Malaysia, Johor, Malaysia, pp: 575-586.

Babu, B.V., A. Nagar, K. Deep, M. Pant and J.C. Bansal *et al*., 2014. Proceedings of the Second International Conference on Soft Computing for Problem Solving (SocProS 2012). Vol. 236, Springer, Berlin,Germany,ISBN:978-81-322-1601-8,Pages:1603.

Baharudin, B., L.H. Lee and K. Khan, 2010. A review of machine learning algorithms for text-documents classification. J. Adv. Inform. Technol., 1: 4-20.

De Houwer, J., D. Barnes-Holmes and A. Moors, 2013. What is learning? on the nature and merits of a functional definition of learning. Psychonomic Bull. Rev., 20: 631-642.

Ghiassi, M., M. Olschimke, B. Moon and P. Arnaudo, 2012. Automated text classification using a dynamic artificial neural network model. Expert Syst. Appl., 39: 10967-10976.

Hamed, S.K. and M.J. Ab Aziz, 2016. A question answering system on Holy Quran translation based on question expansion technique and neural network classification. J. Comput. Sci., 12: 169-177.

Korde, V. and C.N. Mahender, 2012. Text classification and classifiers: A survey. Intl. J. Artif. Intell. Appl., 3: 85-99.

Mohammad, N.F. and N.Omar, 2012. Arabic named entity recognition using artificial neural network. J. Comput. Sci., 8: 1285-1293.

Nassimi, D.M., 2008. A thematic comparative review of some English translations of the Qur'an. Ph.D Thesis, University of Birmingham, Birmingham, UK.

Nassourou, M., 2011. A knowledge-based hybrid statistical classifier for reconstructing the chronology of the Quran. Ph.D Thesis, University of Würzburg, Würzburg, Germany.

Patki, P.S. and V.V. Kelkar, 2013. Classification using different normalization techniques in Support Vector Machine. Intl. J. Comput. Appl., 1: 17-19.

Patra, A. and D. Singh, 2013. A survey report on text classification with different term weighing methods and comparison between classification algorithms. Intl. J. Comput. Appl., 75: 14-18.

Ramlall, I., 2010. Artificial intelligence: Neural networks simplified. Intl. Res. J. Finance Econ., 39: 105-120.

Shahzadi, N. and M.J. Sawar, 2012. Semantic network based classifier of Holy Quran. Intl. J. Comput. Appl., 39: 43-47.

Uysal, A.K. and S. Gunal, 2014. The impact of preprocessing on text classification. Inf. Process. Manage., 50: 104-112.

Vora, K. and Y. Shruti, 2009. A survey on backpropagation algorithms for feedforward neural networks. Intl. J. Eng. Dev. Res., 1: 193-197.

Wang, H., Y. Guo, X. Shi and F. Yang, 2012. Conceptual representing of documents and query expansion based on ontology. Proceedings of the International Conference on Web Information Systems and Mining, October 26-28, 2012, Springer, Berlin, Germany, ISBN:978-3-642-33468-9, pp: 489.

Witten, I.H. and E. Frank, 2005. Data Mining: Practical Machine Learning Tools and Techniques. 2nd Edn., Morgan Kaufman, San Francisco, CA., USA., ISBN-13: 9780080477022, Pages: 560.