Improving text summarization using neuro-fuzzy approach

Muhammad Azhari and Yogan Jaya Kumar

Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

ABSTRACT

In today’s digital era, it becomes a challenge for netizens to find specific information on the internet. Many web-based documents are retrieved and it is not easy to digest all the retrieved information. Automatic text summarization is a process that identifies the important points from all the related documents to produce a concise summary. In this paper, we propose a text summarization model based on classification using neuro-fuzzy approach. The model can be trained to filter high-quality summary sentences. We then compare the performance of our proposed model with the existing approaches, which are based on fuzzy logic and neural network techniques. ANFIS showed improved results compared to the previous techniques in terms of average precision, recall and $F$-measure on the Document Understanding Conference (DUC) data corpus.

ARTICLE HISTORY

Received 3 May 2017
Accepted 2 August 2017

KEYWORDS

Text summarization; ANFIS; fuzzy logic; neural network; classification

1. Introduction

Text summarization has grown into a field of interest to explore, as information from online sources has become the current trend for information seekers. A brief summary of a text document is useful for readers to quickly extract the important information in the texts. For example, when a user looks for a news topic on the web, the search engine would retrieve many articles related to that news. It would be helpful to provide the summary of these articles to the readers instead of having them going through those lengthy texts. Thus, a summary can aid humans in obtaining and understanding the main idea discussed in the texts provided.

Generally, there are two types of summary which can be generated, that is, extractive and abstractive-based summaries (Kumar, Goh, Halizah, Ngo, & Puspalata, 2016). An extractive summary comprises of the original sentences which are selected from the input document. Such summaries can be obtained using methods of sentence extraction, statistical analysis and machine learning techniques. On the other hand, an abstractive summary contains sentences that have to be reconstructed using deep natural language analysis (Salim, 2015). Most studies in text summarization resolves around extractive-based approaches.

CONTACT Yogan Jaya Kumar yogan@utem.edu.my Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, 76100 Melaka, Malaysia

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Many research studies have applied soft computing approaches to improve text summarization and one of the technique is fuzzy logic-based approach (Patil, 2014; Suanmali, Salim, & Binwahlan, 2009). Fuzzy logic is widely used as it could handle uncertainties in data (such as text data) and can interpret the results from the generated rules. However, in all these studies, human or linguistic experts are required to determine the rules for their fuzzy system. These can be a very tedious and time-consuming process. Moreover, the performance of the fuzzy system can be affected by the choice of rules and parameters of membership function.

The motivation of this study is to model an optimized fuzzy-based summarization system and investigate its performance. In our previous published work (Kumar, Kang, Goh, & Khan, 2017), we modelled a neuro-fuzzy system for single document summarization. However, the proposed model was not evaluated using the same set of fuzzy rules created for fuzzy-based system (baseline model). The present work is an extension of the abovementioned paper whereby we validated our findings by comparing the performance of optimized fuzzy-based system (using the same rule base as in the baseline model). This is important to be investigated as using a different set of fuzzy rules would probably give different results. Furthermore, to conclude our findings, we will test our proposed model for multi-document summarization.

The rest of this paper is organized as follows: Section 2 discusses on the related works concerning this study. Section 3 outlines the proposed model. The experimental results and discussion are given in Section 4. Finally, we end with conclusion in Section 5.

2. Related works

In the recent past, soft computing-based approaches have gained popularity in its ability to determine important information across documents (Dixit & Apte, 2012; Megala & Processing, 2014; Patil, 2014; Sarda & Kulkarni, 2015). For instance, a number of studies have modelled summarization systems based on fuzzy logic reasoning in order to select important sentences to be included in the summary (Patil, 2014; Suanmali et al., 2009). First, the features influencing the importance of a sentence are determined, such as title word, sentence position, and thematic word. Then selected sentence features are used as input to the fuzzy system. The scores for each sentence are then derived using fuzzy rules scoring. The sentences with high-fuzzy score will be finally selected to be included in the summary until the desired summary length is obtained.

Apart from sentence scoring, fuzzy logic has also been used for semantic analysis to produce text summary. For example, Kumar, Salim, Abuobieda, and Tawfik (2014) investigated the cross document relations that exist between sentences and used fuzzy logic to rank sentences based on the type of cross document relations. Babar and Patil (2015) extracted the semantic relations between concepts using fuzzy reasoning to select summary sentences. This method (which is based on latent semantic analysis) improves the quality of summary.

Although all the above works support the benefits of employing fuzzy-based reasoning for extracting important sentences from the document, there is a limitation concerning this method. Human or linguistic experts are required to determine the rules for the fuzzy system. Furthermore, the membership functions need to be manually tuned. These can be a very tedious and time-consuming process. Moreover, the performance
of the fuzzy system can be affected by the choice of rules and parameters of membership function (Albertos, 1998).

Besides fuzzy logic, neural network models have also been employed in text summarization studies whereby its learning capabilities are used to identify summary sentences from the input text document. Megala, Kavitha, and Marimuthu (2014) used a three-layered feed-forward network model to learn the patterns in summary sentences. The resulting trained network is then applied to new input documents to determine if a sentence should be included in the summary.

In another related work, Sarda and Kulkarni (2015) used a similar neural network model with the combination of Rhetorical Structure Theory (RST). The RST relations that exist in the sentences are selected by their neural network model and used to form high-quality summaries. Fattah and Ren (2008) proposed an improved content selection approach using probabilistic neural network. They used probability function to better estimate the weights of their neural network model.

Although neural network model has been useful in terms of its learning capabilities, the model provides little information about the relationship between the input and output (a black box approach). The user cannot explain how learning from input data was performed. Thus, it can be observed that both fuzzy logic and neural network have their limitations. Unlike neural network which has adaptation ability, in most fuzzy-based approach, they cannot adapt to changing situations and require an expert to reconstruct the rules, while the result depends on those rules.

In this paper, we propose an extractive multi-document text summarization model based on classification using a hybrid neuro-fuzzy approach known as Adaptive Neuro-Fuzzy Inference System (ANFIS) to increase the capability of fuzzy-based summarization system. The proposed model will be used to classify sentences as summary sentence and non-summary sentence. We also compare our proposed model with existing models which are based on fuzzy logic and neural network technique.

3. Proposed ANFIS-based text summarization

From the related works discussed above, it can be observed that among the two soft computing techniques that have been associated with text summarization, fuzzy logic implementation, which exploits the tolerance for imprecision, is based on knowledge-driven reasoning whereas neural network (which learns to do tasks by considering examples) is based on data-driven approximation. Taking these observations into consideration, a better summarization system can be modelled by considering the advantages of both approaches and avoiding their drawbacks. This has led to the development of an approach which is mostly known as neuro-fuzzy approach. It has the benefits of both neural networks (neuro) and fuzzy logic (fuzzy).

ANFIS is one example of such hybridization (Loganathan & Girija, 2014). It combines the explicit knowledge reasoning of fuzzy logic system, which can explain input output relationship and the implicit knowledge of neural networks, which can be learnt. Past studies have shown that there are limitations with regard to fuzzy-based approaches as human experts are required to determine the rules and tune the membership functions for the fuzzy system (Suanmali et al., 2009).
Our hypothesis is that a better fuzzy-logic-based summarization system can be produced by integrating the learning and adaptive capabilities of neural network to improve the identification of summary sentences.

The architecture of our proposed model is shown in Figure 1. The ANFIS model is trained to classify document sentences as summary and non-summary sentences. Based on the available Document Understanding Conference (DUC) 2002 dataset, we prepared our training set, which comprises the features representing sentences with its corresponding output type, that is, summary or non-summary sentence. After computing the feature values for every sentence from the training set, we input them for the training of ANFIS. Once the training is completed, the resulting classifier model will be able to predict the output score of a new document sentence to determine its class, that is, as summary or non-summary sentence.

### 3.1. Extraction of features

The input documents which have been preprocessed are represented as a vector of features. These features are the attributes that are used to determine the importance of each sentence and sentences with high-feature scores are likely to be selected to form the summary. We extract five features from each sentence. These features have been extensively used in text summarization studies (Kumar et al., 2016; Suanmali et al., 2009). Each feature is given a value between 0 and 1, where values close to 0 indicate a low presence of the feature in the sentence while values close to 1 indicate a strong presence of the feature in the sentence. The five features that were selected as the input for ANFIS include title feature, sentence length, proper noun, thematic word, and term weight.

#### 3.1.1. Title feature

The sentence that contains the word(s) in the document title will be given a high score. Occurrence of words from the document title in a sentence indicates that the sentence is highly relevant to the document. This can be computed by counting the number of

![Figure 1. Text summarization based on ANFIS model.](image-url)
matching characters between the words in sentence and the words in the document title.

\[ f_1 = \frac{\text{number of title words in sentence}}{\text{number of words in document title}}. \]  

### 3.1.2. Sentence length
Sentence that is long is considered to inherit important information. Hence, the sentence length score is computed using the equation below.

\[ f_2 = \frac{\text{number of words occuring in the sentence}}{\text{number of words occuring in the longest sentence}}. \]

### 3.1.3. Proper noun
Sentence containing proper noun is considered to be an important sentence. The scores of sentence that contains proper noun are computed with the equation below.

\[ f_3 = \frac{\text{number of proper nouns in the sentence}}{\text{number of words occuring in the sentence}}. \]

### 3.1.4. Thematic word
This feature is used to determine the commonness of a term. A term that is used frequently is probably related to the topic of the document. We consider the top 10 words as the maximum number of frequent semantic terms.

\[ f_4 = \frac{\text{number of frequent terms in the sentence}}{\max(\text{number of frequent terms})}. \]

### 3.1.5. Term weight
The importance or weight of each word in the document can be computed. The weight \( W_i \) of word \( i \) can be calculated by the traditional \( tf.idf \) method (Suanmali et al., 2009). We adopted this method as \( tf.isf \) (term frequency, Inverse sentence frequency): \[ W_i = tf_i \times isf_i = tf_i \times \log_{n_i}^N, \] where \( tf_i \) is the term frequency of word \( i \) in the document, \( N \) is the total number of sentences and \( n_i \) is number of sentences in which word \( i \) occurs. Using Equation (5), the term weight score for a sentence can be computed as follows:

\[ f_5 = \frac{\sum_{i=1}^{k} W_i(S)}{\max\left( \sum_{i=1}^{k} W_i(S) \right)}, \]

where \( W_i (S) \) is the term weight of word \( i \) in sentence \( S \) and \( k \) is the total number of words in sentence \( S \).

### 3.2. ANFIS model
The five sentence features which have been described in the previous section will become the input to our ANFIS model. Each crisp input will be transformed into fuzzy value using a
membership function. Parameters in this layer are generally referred to as premise parameters and are used to adjust the shape of the membership function. The fuzzy values will be used as incoming signals to compute the firing strength of the corresponding rule. The output of each rule is combined with the linear combination of input variables. Parameters in this layer are referred to as the consequent parameters. The final output is then computed by measuring the aggregation of all incoming signals. Figure 2 depicts our ANFIS model structure. The detailed description of the basic ANFIS architecture is not presented in paper; however, it can be found in our past paper (Aik, 2008).

3.2.1. ANFIS learning method

In conventional fuzzy reasoning-based text summarization, the rules were to be decided by an expert, which is the limitation concerning fuzzy inference system (FIS) in text summarization. However, in the ANFIS model, no expert is required to manipulate the rules as the rules can be generated automatically by using the subtractive clustering method.

Subtractive clustering algorithm estimates the cluster number and cluster centres automatically by mapping the input–output training data. Each instance is seen as a potential cluster centre and the instances that have a value that is in the range of the first cluster will be included as the first cluster. Else, the instance will form a new cluster. The process will repeat until all instances are included in the clusters. An important advantage of using a clustering method to find rules is that the resultant rules are more tailored to the input
data than they are in an FIS generated without clustering (Moh’d Arikat, 2012). This reduces the problem of combinatorial explosion of rules when the input data has a high dimension (the curse of dimensionality). Figure 3 shows the fuzzy rules obtained based on the created data clusters.

To train the ANFIS model, a hybrid method that is a combination of least-square estimation and backpropagation gradient descent method is used. Least-squares Estimate (LSE) is used to minimize the squared error of the actual output and the target output. The backpropagation method is combined with LSE to update the parameters of the membership functions. Backpropagation method originates from multilayer feed-forward neural networks where the network is computed by using the gradient descent method to minimize the sum of squared errors. Backpropagation works by each input weights having their own learning rate, where the learning rate will change over time for each iteration.

A forward pass and a backward pass are included in the hybrid optimization method where forward pass is used to calculate the error measure. The error rates are propagated from the output end towards the input end in the backward pass, where all parameters are updated. The combination of fuzzy inference to represent knowledge in the form of fuzzy rules and membership functions with the learning ability of neural network enables the membership functions parameters to be adjusted directly from the output data. Figure 4 shows the membership function of our input data after training.

Figure 3. Rules generated using subtractive clustering method.
3.2.2. Sentence classification

The trained ANFIS model is then used to classify new input sentences to one of its class, that is, summary or non-summary sentence. The ANFIS model output, which is the predicted sentence score is used to set the classification rule for ANFIS to classify the sentence into binary value (1 or 0). Sentences which are classified to class ‘1’ represents summary sentence, while sentences which are classified to class ‘0’ represents non-summary sentence. The threshold value used to classify the predicted output to one of these two classes were selected based on experimental observation which gave us the least root mean square error (RMSE).

4. Experimental results and discussion

For this study, we used the DUC 2002 dataset for multi-document summarization task. In 2008, DUC became a summarization track in the Text Analysis Conference (TAC). TAC is a series of evaluation workshops organized to encourage research in Natural Language Processing and related applications, by providing a large test collection and common evaluation procedures. TAC is organized by the Retrieval Group of the Information Access Division (IAD) in the Information Technology Laboratory at the National Institute of Standards and Technology (NIST).
The DUC 2002 dataset contains multiple news articles with sample of summaries that have been produced by humans. Based on this dataset, we prepared our training and testing data by labelling each sentence with its class (i.e. 1 or 0). The dataset was preprocessed first before extracting the sentence features. In this phase, the steps that are involved include word tokenization, stop-words removal and stemming. From the dataset, 10 multi-document clusters which comprises of 402 sample sentences were split into 70% training data and 30% testing data using 5 hold out cross validation with balanced class distribution.

The hold out function in MATLAB is used to divide the dataset into balanced amount of class samples in the dataset. This function enables the permutation of different training and testing data using different fold of data in each iteration. Hence, multiple results and performance measure can be obtained from different testing data created using the hold out function. The parameter setting for the baseline methods (i.e. neural network and fuzzy logic) were tuned to give optimal results. The neural network model consists of four hidden nodes and was trained using Levenberg–Marquardt (LM) back propagation technique, which is proved to be one of the fastest and efficient algorithms for training small- and medium-sized feed-forward neural network patterns. The fuzzy logic model which was compared in this study uses Sugeno-type system with three membership functions for each input and five membership functions for the output.

We also ran statistical significance tests (T-Tests) to show the difference in performance between ANFIS and fuzzy-logic-based classification. The T-test is used to determine if there is a significant difference between the results of two groups. A low significance value for the T-test (typically less than 0.05) indicates that there is a significant difference between the two results. It means that there is less than a 5% chance that the two results came from the same group; therefore showing that the results between the two groups are significantly different.

Table 1 and Figure 5 show the precision and recall for summary sentence (class ‘1’) and non-summary sentence (class ‘0’) and using ANFIS, neural network, fuzzy logic and optimized fuzzy logic (the same fuzzy logic model where its membership functions had been tuned by ANFIS). In addition, Table 2 and Figure 6 show the average precision, average recall and average F-measure of the overall classification results. Table 3 present the Paired-Samples T-Test between ANFIS and fuzzy logic (Figure 7).

From this experiment, the results obtained for summary sentence classification using ANFIS, neural network and fuzzy logic gives us an insight into the performance of these soft computing-based approaches towards text summarization. It is clear from Table 1 that ANFIS produces better precision for class ‘1’ when compared to neural network and fuzzy logic and for recall class ‘1’ neural network a slightly better than ANFIS. For overall performance, ANFIS obtained the highest scores compared to other techniques. Based on literature, the performance of the fuzzy logic-based approach is often affected

| Model        | Precision ‘0’ | Precision ‘1’ | Recall ‘0’ | Recall ‘1’ |
|--------------|---------------|---------------|------------|------------|
| ANFIS        | 0.71409       | 0.86704       | 0.90000    | 0.64000    |
| Neural Network | 0.73846       | 0.82863       | 0.85000    | 0.70000    |
| Fuzzy logic  | 0.45464       | 0.14117       | 0.80667    | 0.03333    |
| Optimize FL  | 0.71332       | 0.85132       | 0.88667    | 0.64333    |
by the selection of fuzzy rules and membership functions; we also tested using the same fuzzy model (using the same rule base) and implement it using ANFIS (Optimized FL). It can be observed that much better results were obtained after the membership functions of the fuzzy logic model have been tuned.

### Table 2. Classification results using average precision, recall and F-measure.

| Model            | Average F-measure | Average precision | Average recall |
|------------------|-------------------|-------------------|----------------|
| ANFIS            | 0.7801            | 0.7906            | 0.7700         |
| Neural network   | 0.7792            | 0.7835            | 0.7750         |
| Fuzzy logic      | 0.3479            | 0.2979            | 0.4200         |
| Optimized FL     | 0.7736            | 0.7823            | 0.7650         |

### Table 3. Paired samples test between method ANFIS and fuzzy logic.

| Paired differences | Mean    | Std. deviation | Std. error mean | 95% Confidence interval of the difference | Sig. (two-tailed) |
|--------------------|---------|----------------|-----------------|------------------------------------------|------------------|
| Accuracy           | .348416667 | .044549985    | .004454999      | .339576983 - .357256350                  | .000             |
| F-Measure          | .430768650 | .051535892    | .005153589      | .420542811 - .440994489                  | .000             |
Next, from the significance test values shown in Table 3, we can observe that the results between the ANFIS and fuzzy logic were statistically significant. The comparison models achieved \( p \)-value < .005 and can therefore conclude that there is a significant difference between the obtained results.

It can be noted that, ANFIS, the hybrid approach which takes the advantages from both neural network and fuzzy logic, has improved the accuracy of the summarization model in determining the sentences which should be included in the summary. With the aid of a training algorithm, it enables the process of tuning of the parameters of membership functions for each sentence feature. These results support our hypothesis that a better fuzzy system can be produced by integrating the learning and adaptive capabilities of neural network to improve the identification of summary sentences using optimized fuzzy membership functions and rules. However, in order to affirm the effectiveness of the classification results, performance evaluation is needed to evaluate the final summary generated for a document. This can be achieved using standard evaluation metric for summarization. To further improve the results, the effect of normalization can be studied to increase its recall performance. It should be noted that achieving higher recall rates at better (but not necessarily top) ranks would uncover important sentences more efficiently (Kontostathis & Kulp, 2007).

**5. Conclusion**

In this paper, a study on hybrid soft computing-based approach to improve summary sentence selection is investigated. The key motivation of this paper is the growing number of research studies in text summarization based on soft computing approaches. In order to compensate the disadvantages of one approach with the advantages of another approach, a neuro-fuzzy method called ANFIS is proposed to increase the capability of fuzzy-based summarization system to better identify summary sentences. The proposed approach was able to alleviate some of the limitations in the current text summarization models. The experimental results show that ANFIS achieved better classification results in terms of precision, recall and \( F \)-measure. It should be noted that the quality of summary sentences was not evaluated in this study. In our ongoing work, we attempt to train the ANFIS classifier on larger data samples to generate summaries and evaluate the summaries using Recall Oriented Understudy for Gisting Evaluation (ROUGE) – a standard evaluation metric for summarization.
metric for summarization and implement normalization method to improve the performance of recall.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This research work supported by Universiti Teknikal Malaysia Melaka and Ministry of Higher Education, Malaysia [Research Acculturation Grant Scheme (RAGS) No. RAGS/1/2015/ICT02/FTMK/02/B00124].

**Notes on contributors**

**Yogan Jaya Kumar** received a PhD degree in Computer Science from Universiti Teknologi Malaysia in 2014. His research interest are Text Mining, Information Retrieval, and Soft Computing. He is currently a Senior Lecturer in the Department of Intelligent Computing and Analytics, Universiti Teknikal Malaysia Melaka, Malaysia.

**Muhammad Azhari** is a postgraduate researcher in the Department of Intelligent Computing and Analytics, Universiti Teknikal Malaysia Melaka, Malaysia. He received his Bachelor degree in Artificial Intelligence from Universiti Teknikal Malaysia Melaka, Malaysia, in 2016. His research interests are in the areas of Data Mining and Artificial Intelligence.

**References**

Aik, L. E. (2008). A study of neuro-fuzzy system in approximation-based problems. *Neuro-Fuzzy System ANFIS: Adaptive Neuro-Fuzzy Inference System, 24*(2), 113–130.

Albertos, P. (1998). Fuzzy logic controllers. Methodology. Advantages and drawbacks. In *X Congreso Espanol Sobre Technologias Y Logica Fuzzy (ESTYLF)*, Sevilla, España, pp. 1–11. doi:10.13140/RG.2.1.2512.6164.

Babar, S. A., & Patil, P. D. (2015). Improving performance of text summarization. In *Procedia – Procedia Computer Science* (Vol. 46, pp. 354–363). Elsevier Masson SAS. doi:10.1016/j.procs.2015.02.031

Dixit, R. S., & Apte, P. S. S. (2012). Improvement of text summarization using fuzzy logic based method. *IOSR Journal of Computer Engineering (IOSRJCE)*, 5(6), 5–10.

Fattah, M. A., & Ren, F. (2008). Automatic text summarization. *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, 2(1), 90–93.

Kontostathisis, A., & Kulp, S. (2008). The effect of normalization when recall really matters. International conference on Information and Knowledge Engineering (IKE), Las Vegas, NV, pp. 96–101.

Kumar, Y. J., Goh, O. S., Halizah, B., Ngo, H. C., & Puspalata, C. S. (2016). A review on automatic text summarization approaches. *Journal of Computer Science*, 12(4), 178–190. ISSN 1549-3636.

Kumar, Y. J., Kang, F. J., Goh, O. S., & Khan, A. (2017). Text summarization based on classification using ANFIS. In D. Król, N. Nguyen, & K. Shirai (Eds.), *Advanced topics in intelligent information and data base systems (ACILDS)* (Vol. 710, pp. 405–417). Studies in Computational Intelligence. Cham: Springer.

Kumar, Y. J., Salim, N., Abuobieda, A., & Tawfik, A. (2014). Multi document summarization based on news components using fuzzy cross-document relations. *Applied Soft Computing Journal*, 21, 265–279. doi:10.1016/j.asoc.2014.03.041

Loganathan, C., & Girija, K. V. (2014). Investigations on hybrid learning in ANFIS. *International Journal of Engineering Research and Applications*, 2(10), 31–37.
Megala, S. S., Kavitha, A., & Marimuthu, A. (2014). Enriching text summarization using fuzzy logic. International Journal of Computer Science and Information Technologies, 5(1), 863–867.

Megala, S. S., & Processing, A. P. (2014). Enriching text summarization using fuzzy logic. (IJCSIT) International Journal of Computer Science and Information Technologies, 5(1), 863–867.

Moh’d Arikat, Y. (2012). Subtractive neuro-fuzzy modeling techniques applied to short essay auto-grading problem. In 2012 6th International conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT) (pp. 889–895). IEEE.

Patil, M. P. D. (2014). Text summarization using fuzzy logic. International Journal of Innovative Research in Advanced Engineering, 1(3), 42–45.

Salim, N. (2015). Genetic semantic graph approach for multi-document abstractive summarization genetic semantic graph approach for multi-document abstractive summarization (December). doi:10.1109/ICDIPC.2015.7323025

Sarda, A. T., & Kulkarni, A. R. (2015). Text summarization using neural networks and Rhetorical Structure Theory. International Journal of Advanced Research in Computer and Communication Engineering, 4(6), 49–52. doi:10.17148/IJARCEE.2015.4612

Suanmali, L., Salim, N., & Binwahlan, M. S. (2009). Fuzzy logic based method for improving text summarization. International Journal of Computer Science and Information Security, 2(1), 4–10.