BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **15**, No 4

Sofia • 2015 Print ISSN: 1311-9702; Online ISSN: 1314-4081
DOI: 10.1515/cait-2015-0056

Towards an Efficient Approach for Automatic Medical Document Summarization

Gayathri, P., N. Jaisankar

School of Computing Science and Engineering, VIT University, Tamil Nadu, India
Emails: pgayathri@vit.ac.in njaisankar@vit.ac.in

Abstract: Document summarization deals with providing condensed version of the original document. We present an extractive informative single medical document summarization approach. We compare the tokens in the sentence with cue words. A sentence ranking method is used to extract the important sentences. The existing summarizers are used for performance analysis.

Keywords: Document summarization, medical document, Domain-specific vocabulary, cue words, sentence ranking.

1. Introduction

There is a huge amount of information available in today’s fast growing information world. Document summarization has become an important tool of interpreting the usefulness of information [8]. Summarization is one of the information retrieval tasks. It helps to determine whether the retrieved document is relevant for information need by the user or not. The aim of the automatic summarization system is to shorten the length of the document without affecting the overall meaning. In general, the summarization system can take news article,a group of news articles, email, email threads or domain-specific information as input. The three basic primitive steps involved in the summarization task are as follows: representation of the input document, selection of the important content and generation of a novel content that corresponds to the gist of the document. Today, Evidence Based Medicine (EBM) is practiced in medical domain. Medical experts are not only based on years of their experience but on the recent discoveries also. In this scenario, medical professionals and researchers demand relevant medical information from a healthcare information system. However, available information systems fail to provide relevant information due to the overwhelmed data in the medical domain. The author written abstract of the document helps them to speedily understand what the article is about. Not all medical documents come with an author written abstract or summary. Because of this, medical professionals and researchers could not check the relevancy of retrieved document. So the summary produced by the summarization system may help them know about the document content and can decide whether the document is useful for in-depth study. Also it saves time of reading and understanding the large document.

The document summarization method can be broadly classified as extractive and abstractive. An extractive method extracts important sentences from the original document and thus creates a summary. Important sentences are defined based on the linguistic and statistical features. In the abstractive method, the sentences in the original document are re-phrased [1]. This method of summarization deals with understanding of the original content and retelling it in fewer words. Human beings generally write an abstractive summary by understanding the information content of the document. Thus, the abstractive summaries are summaries expressed in a clear natural language based on the main concepts in the original document [6]. The summarization system can also be sub-categorized based on the type of the detail provided, type of the content available, number of input documents and language [6]. Based on the type of the detail, the summary can be either indicative (provides the main idea of the original document in the summary) or informative (shortens the length without changing its meaning). Based on the type of the content, the summary can be either generic (not based on user's interest, it gives the same level of importance to all sentences when producing a summary) or query-based (based on the user's interest or query when producing a summary, it gives importance to certain sentences). Based on the number of input documents, the summary can be either a single document (produces a summary for only one document at a time) or multi-document (produces summaries for a group of related documents). Based on the language, the summary can be either mono-lingual (can produce a summary of documents written in one specific language) or multi-lingual (can produce a summary of documents written in different languages). The summary could be also genre-specific or domain-independent. In the first, the documents belonging to one specific domain are considered, whereas in the later, the documents belonging to any domain can be considered for summarization.

In this paper we describe a novel approach for single document extractive informative summarization of documents belonging to the medical domain. We treat the document as a set of sentences. We use the domain specific terms as cue words to score and rank the sentences. The cue words are the important words or terms specific to a domain [6]. For example, "heart disease", "risk factors", "blood pressure", "symptoms", "therapy", "treatment", etc., are domain-specific cue words which are good to preserve when creating a summary of a medical document on some of these topics. These cue words affect the summary worthiness because the sentences with cue words must get high preference to occur in the final summary than a sentence without cue words [10]. In order to identify all such terms, we maintain the domain-specific vocabulary which is built up by using Medical Subject Headings – MeSH (**http://www.nlm.nih.gov/mesh/**), a vocabulary Thesaurus controlled by National Library of Medicine (NLM). MeSH expressions are pre-processed to remove the stop words. A large syntactic lexicon of biomedical and general English is the SPECIALIST lexicon. Its coverage includes both commonly occurring English words and biomedical vocabulary. PhraseX program extracts noun phrase strings from the text by referring to the syntactic structure provided by the SPECIALIST minimal commitment parser, which relies on the SPECIALIST Lexicon and the Xerox stochastic tagger [14]. Thus, the noun phrases are identified from MeSH expressions and the vocabulary is maintained. If a sentence contains N cue words, then the score of this particular sentence is considered to be N. The score for each cue word is calculated based on the hits received from different sentences. All sentences are considered for ranking. Sentence ranking is done based on the cue word frequency of a sentence. The sum of cue word score of all cue words in a sentence gives the cue word frequency of the sentence. A sentence with high cue word frequency gets a high rank. Then based on the summary length, the sentences are extracted to create the summary. The similarity measure is used as an additional component to produce a more informative final summary. It helps the system to include highly dissimilar sentences in the final summary. Not all medical documents come with an author written abstract or summary. So, the medical documents with author written abstracts are used to test the proposed work. In our proposed work, all domain-specific cue words are considered to be equally important because prioritizing the cue words requires in-depth domain specific knowledge [10].

This paper is organized as follows: Section 2 contains related works in document summarization. Section 3 provides our proposed automatic extractive informative single document summarization model for medical documents. Section 4 describes the existing summarizers used for comparison. We present the evaluation, the experimental results and discussion in Section 5. At last, Section 6 concludes the proposed work.

2. Related works

The prior works mentioned within are the general approaches towards document summarization.

A language independent single document summarization using Wikipedia has been proposed in [2]. As the mobile devices are running out by the screen space and bandwidth, the authors have aimed at simplifying the information content of the document. They have presented only the most relevant information contained in the document. They mapped the sentences in the document to Wikipedia terms and based on the frequency of the mapped terms, they selected the sentences for summary. The mapping between a sentence and Wikipedia terms is represented as a bi-partite graph. The terms that receive hits above the maximum threshold and below the minimum threshold are excluded. The sentences that map to those terms are not included in the summary. In performance evaluation, the tool Recall-Oriented Understudy for Gisting Evaluation (ROUGE) (**http://www.berouge.com**) has been used to compute 1-gram score for 100-word summary. The work presented in [3] introduced two approaches to email thread summarization. Collective Message Summarization (CMS) applies multi-document summarization approach and Individual Message Summarization (IMS) applying the single document summarization approach. In this approach, the sentence compression technique is used. The approach was not purely extractive. The linguistic and statistical methods are used to generate the summary. The work provided efficient applications to access large email collections. The results are tested by using documents in Enron collection corpus. In [4], the authors have introduced a new approach for single document text summarization and simplification. The keywords are selected by using the weighted approach. The input document is divided into two parts: informative and non-informative. Summarization and simplification is done individually on both parts. Part of Speech (PoS) tags are obtained using the NLTK tagger. The nouns in the informative part are simplified using WordNet. The keyword selection approach was used in the non-informative part. The two output files obtained after summarization and simplification were merged to get the single output file.

The work presented in [5] exploits an extraction based single document summarization approach, using neural networks and fuzzy logic. In this approach, the feature based sentence scoring technique is used. Analysis is done by comparing the neural networks and fuzzy logic based on the performance measures precision, recall and f-measure. Precision is the ratio of the number of the system extracted sentences that match with the manual summary to the number of sentences in the system extracted summary. The recall is the ratio of the number of system extracted sentences that match with the manual summary to the number of sentences in the manual summary. The harmonic mean of precision and recall gives the f-measure. The experimental results show that fuzzy logic performs better when compared with neural networks.

A sentence ranking technique has been used for web document summarization in [7]. The sentences are ranked using two important measures, such as frequency in terms of the sentences and similarity to other sentences. Highly ranked sentences are used in the summary. Performance evaluation is done by using a recall measure. A language independent domain-independent automatic text summarization is proposed in [9]. The summarization is done by using the sentence extraction method. The unsupervised learning algorithm is used to create the clusters that contain similar sentences. Most representative sentence in each cluster is identified and it is used to generate the final summary. The experimental results obtained show that their approach provides 0.5%, 0.55%, 0.46%, 0.6% better accuracy when compared with the existing other approaches [15-18] respectively.

The number of approaches tackling the problem for document summarization in the medical domain has also been proposed. Machine learning approach for medical document summarization is used in [10]. Supervised learning method is used to classify sentences based on their worthiness. A bagging method with C4.5 decision tree algorithm is used as a base learner. The learning algorithm labels each sentence as summary worthy or moderately summary worthy, or summary unworthy. The authors used centroid overlap, sentence position, first sentence overlap, sentence length, domain-specific cue phrases, position of the cue phrases and acronyms as features to characterize and rank sentences. Based on this ranking, the sentences are selected to create a summary. The sentence similarity is measured by using idf-cosine similarity metric. The proposed model has been tested with medical news articles. The authors have projected their results for various ROUGE scores. Summaries, produced by human abstractors are used as reference summaries in ROUGE.

The problem of single document summarization is addressed as a binary optimization problem in [11]. The summarization is done based on the genetic operators and guided local search. The proposed method is compared with other methods like Unified Rank, DE, FEOM, Net Sum, CRF, SVM, QCS and Manifold Ranking using ROUGE measures. In [12] the authors have proposed a context-based word indexing model for document summarization. The authors have narrated that the existing models for document summarization use similarity between sentences to extract the most informative sentences in order to produce the final summary. They focused on single document summarization using a sentence extraction approach. Bernoulli model of randomness is used to calculate the similarity between sentences. The sentence similarity matrix is computed by finding lexical association between the terms in the sentence. They have shown that the system is capable of producing context-sensitive document summarization. The experimental evaluation is performed over the benchmark DUC data sets.

The approaches above presented are similar since they deal with single document summarization using the sentence ranking method. But these approaches differ in the features considered for summarization.

Based on understanding the above mentioned summarization approaches, we introduce an extractive informative single medical document summarization that exploits domain-specific knowledge. It is observed from literature that most of the summarization approaches create the final summary by using the sentence ranking method. The researchers have used a few or more number of features for sentence ranking. In our proposed summarization approach, we have opted for the best sentence feature for ranking.

3. Proposed system model

Our proposed summarization approach is an informative extraction based single document summarization that can be performed over the documents belonging to the medical domain. It extracts parts of the original document and provides them as a final summary. The sentence extraction method is used to condense the length of the document without affecting its information content and its meaning. The sentences are scored based on the occurrence of the cue words. Each cue word is scored based on the number of hits received from different sentences. Based on the cue word frequency, each sentence is ranked. The ranked sentences are presented to the user as a final summary based on the similarity measure and summary length. The similarity measure is used to avoid highly similar sentences in the final summary. The proposed approach contains three phases, namely: document pre-processing, sentence ranking and summary creation. Fig. 1 depicts our proposed summarization approach.

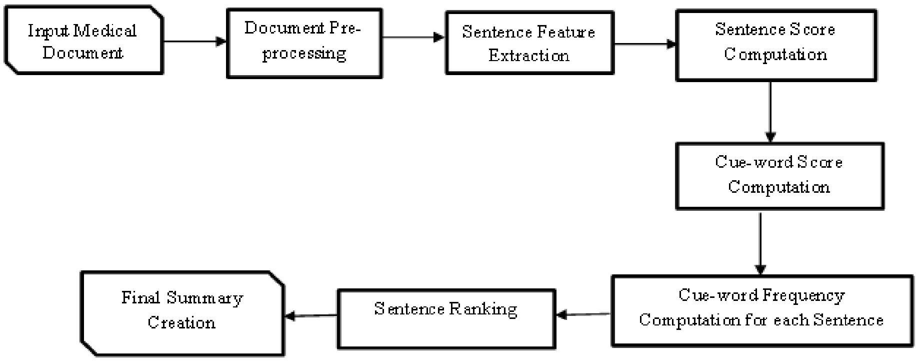

Fig. 1. Proposed summarization approach

3.1. Document pre-processing

Pre-processing gives structured representation to the original input document [8]. It performs removal of punctuation marks and stop words. Removal of the punctuation marks is the process of removing all punctuation marks in the document except for the dot (.) because the dot is used to identify the sentence boundary (end of each sentence). The stop words are the words that appear frequently in the document and which do not contribute any semantic information towards the final summary. Thus, the pre-processing removes the noisy text from the original input document.

3.2. Sentence ranking

To perform this, the medical domain-specific vocabulary is stored in a knowledge base. The dictionary entries are used as cue words. The individual words in each sentence are considered as tokens and compared with the cue words. The process is repeated for each sentence in the document. Bag-of-words representation with binary values is used to indicate the matching between a sentence and cue words [19]. The value 1 represents the match between the token in the sentence and the cue word, whereas the value 0 represents that there is no match between them.

Table 1. Bag-of-words representation

Sentence	Cue words (W_j)						
(S_i)	W_1	W_2	W_3	W_4	W_5	W_6	W_7
S_1	1	0	1	0	1	0	1
S_2	1	1	0	0	0	0	0
S_3	0	0	1	0	0	0	1
S_4	0	0	0	1	1	1	0
S_5	0	0	0	1	0	0	0

Consider a small document that contains 5 sentences and a knowledge base with 7 cue words which is represented as a bag-of-words in Table 1. The size of the table will be $r*c$ where r (rows) is the number of the sentences in the document and

c (columns) is the number of cue words in the knowledge base. This bag-of-words representation is used to compute the score for each sentence $S(S_i)$ and the score for each cue word $S(W_j)$. Then the cue word frequency of sentence $F(S_i)$ is calculated. The score for each sentence i is computed by

$$S(S_i) = \sum_{j=1}^{c} \alpha_{ij}, \quad i = 1, \ldots, r, \tag{1}$$

for example, $S(S_1) = 4$.

The score for each cue word j is computed by

$$S(W_j) = \sum_{i=1}^{r} \beta_{ij}, \quad j = 1, \ldots, c, \tag{2}$$

for example, $S(W_1) = 2$.

Then the sentences are ranked based on their cue word frequency. The cue word frequency for each sentence i is

$$F(S_i) = \sum\nolimits_{\text{for all words } Wj \text{ that occurs in sentence } i} S(W_j), \tag{3}$$

where W_j is the j-th word in the sentence i, and $S(W_j)$ is the score of the cue word W_j in sentence I; for example, $F(S_1) = S(W_1) + S(W_3) + S(W_5) + S(W_7) = 8$.

The sentence with high cue word frequency gets a high rank. Thus, the sentences are ranked and arranged in a decreasing order of their cue word frequency.

3.3. Summary creation

After ranking and ordering of the sentences, N sentences are used to create the final summary. The value of N depends on the compression ratio of the final summary. If N high ranked sentences are selected, there is a possibility to have a highly similar sentence in the final summary. In order to create a more informative final summary, it is necessary to keep the sentences in the summary sufficiently dissimilar from each other. Jaccard similarity measure is used to compute the similarity between the sentences. As the summary length is restricted, there is a need to restrict the similar sentences in order to create a more informative final summary. To include sufficiently dissimilar sentences in the final summary, the following algorithm is used.

Summary creation algorithm:

Input: Ordered list of sentences

Output: Final summary

Step 1. Include the first sentence in the final summary.

Step 2. Choose the next sentence from the ordered list.

Step 3. Calculate the dissimilarity between the chosen sentence and all other previously included sentences in the summary.

Step 4. Include the chosen sentence in the final summary if it is sufficiently dissimilar from all the other previously included sentences in the summary.

Step 5. Repeat Steps 2, 3 and 4 until a predefined length of the final summary is attained.

The Jaccard similarity coefficient (J) value between any two sentences [13] can be calculated by

(4) $$J(X, Y) = \frac{|X \cap Y|}{|X \cup Y|}, \qquad 0 \le J(X, Y) \ge 1,$$

where $X \cap Y$ and $X \cup Y$ can be defined as the size of the intersection of words in two sentences and the size of union of the words in two sentences respectively. We say two sentences are *dissimilar sentences* if its D_J value is greater than the minimum threshold. Otherwise, those sentences are *similar sentences*. A higher D_J value makes the sentences more dissimilar. The dissimilarity D_J between the two sentences X and Y can be calculated by

(5) $$D_J(X, Y) = 1 - J(X, Y) = \frac{|X \cup Y| - |X \cap Y|}{|X \cup Y|}.$$

In order to increase the readability of the final summary, the sentences included in the final summary are rearranged according to the order in which they appear in the original input document.

4. Existing summarizers used for comparison

We have compared the performance of the proposed summarization approach with the popular existing summarization system called MEAD (**http://www.summarization.com/mead/**) and with a free online text summarizer called Text compactor (**http://www.textcompactor.com**). MEAD is a single (individual document) and multi-document (clusters of related documents) summarizer which summarizes a document using features like sentence position, similarity of sentences to centroid, etc. The text compactor is a free online single document summarizer which summarizes document based on a compression ratio and sentence score. We have also compared our domain-specific vocabulary based summarization approach with the Wikipedia based summarization approach [2].

5. Performance evaluation, experimental results and discussion

For the purpose of performance analysis, we have used 100 medical documents with an authors written abstracts (summaries) from various sources like medical news today (**http://www.medicalnewstoday.com/articles**), Medline Plus (**http://nlm.nih.gov/medlineplus**), PubMed (**http://nlm.nih.gov/pubmedtutorial**), etc. The articles are related to the medical topics, such as malaria, dengue, heart disease, diabetes, communicable diseases, non-communicable diseases, deficiency diseases, drugs and their side effects, modern surgery, vaccinations, symptoms etc. From these test articles, we have removed the images, links and author written abstracts and then we have considered the documents for the summarization process. For each test document, we have created a system generated summary by using our proposed approach, Wikipedia based summarizer, MEAD and Text compactor. The reported results from different summarization algorithms are obtained on the same document corpus.

5.1. Performance evaluation

Performance evaluation measures are useful in evaluating the efficiency and usefulness of the system generated summaries. We have used primary measures precision, a recall and f-measure to evaluate the performance of the system. We have also used ROUGE for automated evaluation of the system generated summaries.

5.1.1. Precision

The precision is the measure that is used to evaluate the correctness of sentences in the final summary produced,

$$\text{Precision} = X/Y,$$

where X is the number of system extracted sentences that match with the manual summary and Y is the number of sentences in the system extracted summary.

5.1.2. Recall

Recall is the measure that is used to evaluate the relevant sentences which are included in the final summary,

$$\text{Recall} = X/Z,$$

where, X is the number of system extracted sentences that match with the manual summary and Z is the Number of sentences in the manual summary.

5.1.3. F-measure (F-score)

Weighted harmonic mean between the precision and recall measure is F-measure. F-measure attains its best score at 1 and the worst score at 0,

$$\text{F-measure} = (\text{Precision} * \text{Recall})/((\text{Precision} + \text{Recall})/2).$$

We have used author written abstracts in the test documents as a manual summary. In these summaries are abstractive, we create an extractive manual summary by selecting the sentences from the original document that best match the sentences in the abstract.

5.1.4. ROUGE

ROUGE requires the same length of the reference (model) summary and system generated summary. We have used the author written abstracts in the test documents as a reference summary. ROUGE reports the separate scores for 1-, 2-, 3- and 4-gram matching between the model summaries and generated summaries. In general, this can be represented as n-gram, where n indicates the sequence of the consecutive words in a summary sentence. For a single document summarization, the recall measure can be defined as percentage of n-grams in the model summary that also occurs in the generated summary. ROUGE-N measure is the n-gram recall between the candidate (system generated) summary and the reference summary. The ROUGE-N can be calculated by

$$\text{(6)}\qquad \text{ROUGE-N} = \frac{\sum_{S\in\{\text{Reference summaries}\}} \sum_{\text{gram}_n\in S} \text{Count}_{\text{match}}(\text{gram}_n)}{\sum_{S\in\{\text{Reference summaries}\}} \sum_{\text{gram}_n\in S} \text{Count}(\text{gram}_n)},$$

where n denotes the length of n-gram. The maximum number of n-grams co-occurring in a candidate summary and reference summary is represented by $gram_n$ and $Count_{match}(gram_n)$.

5.2. Experimental results and discussion

We now produce the final summary created by our proposed approach for the document which describes the civil society and medical profession in dengue prevention. Fig. 2 contains the original input medical document, in which the cue words are highlighted in bold. Fig. 3 contains the author written abstract for the original input medical document. Figs 4 and 5 show the final summary created for the compression ratio set to 15% and 25%, respectively. The compression ratio set to p% indicates that only p percent of the total sentences in the original input document are extracted to create the final summary.

> ***Dengue fever*** *caused by* ***dengue viruses*** *is an annual seasonal feature in many towns in Tamil Nadu.* ***Dengue viruses*** *are transmitted by* ***aedes*** *mosquitoes which are relatively small compared to other mosquitoes and have black and white striped body and legs. They breed in fresh water in small containers both inside and outside our own house. The mosquito needs only a very thin layer of water to lay eggs on. If sufficient water is present, the eggs hatch and adult mosquitoes emerge in 7 days after going through the larval and pupal stages. If the water dries up, the eggs remain in a dry stage until the next season when they get immersed in water and then they hatch. Since they breed prolifically in rain water collection, they are more prevalent during July to December.* ***Aedes*** *mosquitoes stay mostly indoor and bite anytime during the day. People get infected with* ***dengue virus*** *through the bite of already infected mosquitoes.* ***Virus*** *multiplies in the body to many millions and then more mosquitoes pick up the* ***virus*** *and transmit to others. Thus, the humans are involved in spreading* ***dengue viruses*** *in two ways: allowing mosquitoes to breed and also allowing them to bite. So, our awareness and behavior are keys to preventing* ***dengue fever****.*

Fig. 2. Original input medical document

> *Dengue fever caused by dengue viruses is an annual seasonal feature in many towns in Tamil Nadu. Dengue viruses are transmitted by aedes mosquitoes which are relatively small compared to other mosquitoes and have black and white striped body and legs. Aedes mosquitoes stay mostly indoor and bite anytime during the day. People get infected with dengue virus through the bite of already infected mosquitoes.*

Fig. 3. An author written abstract for the original input medical document

> *Dengue fever caused by dengue viruses is an annual seasonal feature in many towns in Tamil Nadu. Dengue viruses are transmitted by aedes mosquitoes which are relatively small compared to other mosquitoes and have black and white striped body and legs.*

Fig. 4. Final summary created for the compression ratio set to 15%

> *Dengue fever caused by dengue viruses is an annual seasonal feature in many towns in Tamil Nadu. Dengue viruses are transmitted by aedes mosquitoes which are relatively small compared to other mosquitoes and have black and white striped body and legs. People get infected with dengue virus through the bite of already infected mosquitoes.*

Fig. 5. Final summary created for the compression ratio set to 25%

Tables 2 and 3 show that our proposed method obtains better experimental results for the compression ratio set to 15% and 25%, respectively. Figs 6 and 7 give graphical representation of the obtained results.

Table 2. Results for the compression ratio set to 15%

Summarizer	Precision	Recall	F-measure
Proposed approach	0.68	0.41	0.51
Wikipedia based summary	0.65	0.39	0.49
MEAD	0.63	0.35	0.45
Text compactor	0.60	0.32	0.42

Table 3. Results for the compression ratio set to 25%

Summarizer	Precision	Recall	F-measure
Proposed approach	0.63	0.51	0.56
Wikipedia based summary	0.61	0.45	0.52
MEAD	0.58	0.43	0.49
Text compactor	0.55	0.39	0.46

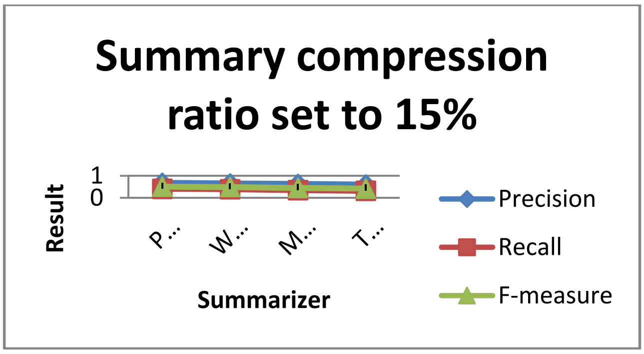

Fig. 6. Results with respect to precision, recall and f-measure for a compression ratio of 15%

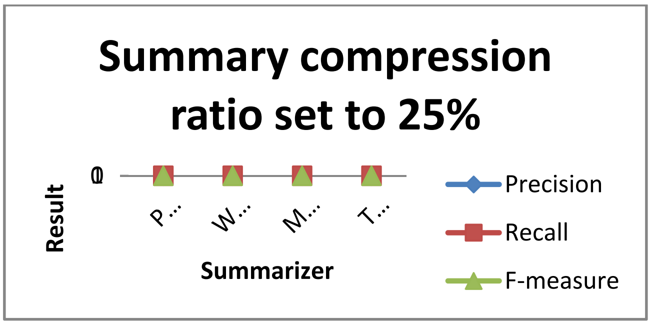

Fig. 7. Results with respect to precision, recall and f-measure for compression ratio of 25%

We consider first *n* words to perform the experimental evaluation of the proposed system using ROUGE. For 100-words and 150-words summary generation task, we set *n* value as 100 and 150 respectively. We report ROUGE 1-gram and 2-gram score at the confidence interval of 95%. Tables 4 and 5 show average ROUGE recall scores. Also, Figs 8 and 9 show the graphical representation of the obtained ROUGE results.

Table 4. Obtained ROUGE results for a 100-words summary

Summarizer	ROUGE-1 score	ROUGE-2 score
Proposed approach	0.50	0.35
Wikipedia based summary	0.49	0.32
MEAD	0.47	0.31
Text compactor	0.45	0.29

Table 5. Obtained ROUGE results for a150-words summary

Summarizer	ROUGE-1 score	ROUGE-2 score
Proposed approach	0.53	0.37
Wikipedia based summary	0.50	0.35
MEAD	0.49	0.34
Text compactor	0. 46	0.31

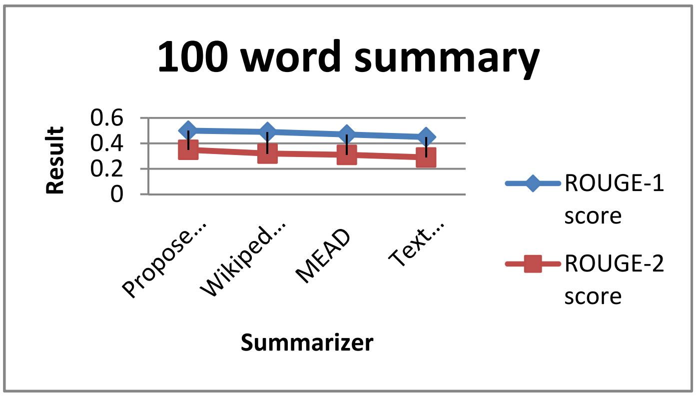

Fig. 8. ROUGE results for a 100-words summary

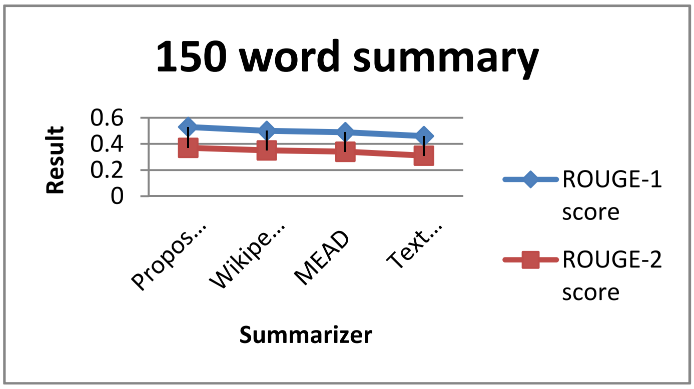

Fig. 9. ROUGE results for a 150-words summary

These results reveal that the quality of the summaries produced by the proposed system is better when compared with the existing summarizers. Also, our proposed domain specific vocabulary based summarization method is able to create better summary when compared with the Wikipedia based summarization method for medical documents.

6. Conclusion

This paper explains an efficient approach for automatic extractive informative single medical document summarization. Here the domain-specific vocabulary words are used as cue words. Pre-processing is done to remove the noisy data from the input document. The pre-processed document is used for ranking. In order to rank each sentence, the score is calculated for each sentence based on the number of occurrences of the cue words. The cue word score is calculated for each cue word in the knowledge base. Finally, the sentences are ranked based on their cue word frequency. The proposed system produces a final summary based on the compression ratio in order to restrict the length of the summary produced. Since the similarity measure is used in the final summary creation, the system includes highly dissimilar sentences and produces a more informative final summary. The system is tested with the input of 100 medical articles taken from various medical sources. The articles with author written abstracts are considered for testing. The result shows that the proposed approach performs better compared to the existing summarizers. The performance of the system is also measured with respect to the quality of the summary produced by using ROUGE.

References

1. Pimpalshende, A. N. Overview of Text Summarization Extractive Techniques. – International Journal of Engineering and Computer Science, Vol. **2**, 2013, No 4, pp. 1205-1214.
2. Ramanathan, K., Y. Sankarasubramaniam, N. Mathur, A. Gupta. Document Summarization Using Wikipedia. – In: Proc. of 1st International Conference on Intelligent Human Computer Interaction, Allahabad, India, 2009, pp. 254-260.
3. Zajic, D. M., B. J. Dorr, J. Lin. Single-Document and Multi-Document Summarization Techniques for Email Threads Using Sentence Compression. – Information Processing and Management, Vol. **44**, 2008, No 4, pp. 1600-1610.
4. Shubhangi, C. T. An Approach to Single Document Text Summarization and Simplification. – IOSR Journal of Computer Engineering, Vol. **16**, 2014, No 3, pp. 42-49.
5. Santhana, M. S., A. Kavitha, A. Marimuthu. Enriching Text Summarization Using Fuzzy Logic. – International Journal of Computer Science and Information Technologies, Vol. **5**, 2014, No 1, pp. 863-867.
6. Munot, N., S. S. Govilkar. Comparative Study of Text Summarization Methods. – International Journal of Computer Applications, Vol. **102**, 2014, No 12, pp. 33-37.
7. Reddy, Y. S., D. A. S. Kumar. An Efficient Approach for Web Document Summarization by Sentence Ranking. – International Journal of Advanced Research in Computer Science and Software Engineering, Vol. **2**, 2012, No 7, pp. 221-225.
8. Gupta, V., G. S. Lehal. A Survey of Text Summarization Extractive Techniques. – Journal of Emerging Technologies in Web Intelligence, Vol. **2**, 2010, No 3, pp. 258-268.
9. Garcia-Hernandez, R. A., R. Montiel, Y. Ledeneva, E. Rendon, A. Gelbukh, R. Cruz. Text Summarization by Sentence Extraction Using Unsupervised Learning. – In: Proc. of Mexican International Conference on Artificial Intelligence: Advances in Artificial Intelligence, 2008, pp. 133-143.

10. S a r k a r, K., M. N a s i p u r i, S. G h o s e. Using Machine Learning for Medical Document Summarization. – International Journal of Database Theory and Application, Vol. **4**, 2011, No 1, pp. 31-48.
11. M e n d o z a, M., S. B o n i l l a, C. N o g u e r a, C. C o b o s, E. L e o n. Extractive Single-Document Summarization Based on Genetic Operators and Guided Local Search. – Expert Systems with Applications, Vol. **41**, 2014, No 9, pp. 4158-4169.
12. G o y a l, P., L. B e h e r a, T. M. M c G i n n i t y. A Context-Based Word Indexing Model for Document Summarization. – IEEE Transactions on Knowledge and Data Engineering, Vol. **25**, 2013, No 8, pp. 1693-1705.
13. N i w a t t a n a k u l, S., J. S i n g t h o n g c h a i, E. N a e n u d o r n, S. W a n a p u. Using Jaccard Coefficient for Keywords Similarity. – In: Proc. of International MultiConference of Engineers and Computer Scientists, Vol. **1**, 2013, pp. 23-27.
14. H l i a o u t a k i s, A., K. Z e r v a n o u, E. G. P e t r a k i s. The AMTEx Approach in the Medical Document Indexing and Retrieval Application. – Data and Knowledge Engineering, Vol. **68**, 2009, No 3, pp. 380-392.
15. L e d e n e v a, Y., A. G e l b u k h, H. R. G a r c i a. Terms Derived from Frequent Sequences for Extractive Text Summarization. – Computational Linguistics and Intelligent Text Processing, Vol. **4919**, 2008, pp. 593-604.
16. M i h a l c e a, R. Random Walks on Text Structures. – Computational Linguistics and Intelligent Text Processing, Vol. **3878**, 2006, pp. 249-262.
17. H a s s a n, S., R. M i h a l c e a, C. B a n e a. Random-Walk Term Weighting for Improved Text Classification. – International Journal of Semantic Computing, Vol. **1**, 2007, No 4, pp. 421-439.
18. M i h a l c e a, R., P. T a r a u. TextRank: Bringing Order into Texts. – In: Proc. of Empirical Methods in Natural Language Processing, Barcelona, Spain, 2004.
19. F r u n z a, O., D. I n k p e n, T. T r a n. A Machine Learning Approach for Identifying Disease-Treatment Relations in Short Texts. – IEEE Transactions on Knowledge and Data Engineering, Vol. **23**, 2011, No 6, pp. 801-814.