International Journal of Computing and Applications | **ISSN: 0973-5704**
Volume 13, Number 1, (January-June 2018), pp 29-37

A STUDY ON EFFECT OF POSITIONAL INFORMATION IN SINGLE DOCUMENT TEXT SUMMARIZATION

Santanu Dam[1], Kamal Sarkar[2] and Sohini Roy Chowdhury[3]

[1]Department of CSE, Future Institute of Engineering & Management, Kolkata-150, India. Email: sntndm@gmail.com
[2]Department of Computer Science and Engineering, Jadavpur University, Kolkata, India. Email: jukamal2001@yahoo.com
[3]Department of Computer Science and Engineering, Jadavpur University, Kolkata, India

Abstract: In this paper, we present a study on how positional feature may affect single document summarization performance. In our approach, after ranking sentences based on the combination of sentence centrality score, TF-ISF based score, a summary is generated in two step process- the first step selects sentences in the summary if the sentences satisfy stringent positional criteria. If the desired length of the summary is not reached at the first step, the second step starts. In the second step, the position based criteria is relaxed. Sentences are added to the summary one by one from the ranked list in order if the sentences are not previously selected in the summary. The experimental results on DUC 2002 summarization data set show that our proposed method of using positional information based criteria boosts up the text summarization performance. Our proposed approach has been compared with the state-of-the art summarization approaches. The experimental results show that the performance of proposed approach is comparable to the state-of-the art summarization approaches.

Keywords: Single Document Text Summarization; Positional Information; Centrality Score; Centroid.

1. INTRODUCTION

Information today in digital world plays a key role. Now the today's people are overwhelmed by the huge amount of information available on the internet. Due to availability of the large amount of information, the users take long time to read each documents in the collection and find out the relevant topic. So it is very essential to devise an improved mechanism that can effectively represent information. Primary idea of any type of representation is to present main ideas of the document in very less space. Text summarization is a mechanism where original text compressed into shorter form by extracting salient information from the source text and to present that information to the user in the form of summary to provide same meaning and same amount of information in respect to original text. So, any summary helps the reader to quickly and easily understand the content of the original documents without reading the entire document individually. Manual summarization is a process that produces summary of the given document with the help of human. Manual summary takes huge amount of time and cost. For this reason automatic text summarization methods has been introduced. Automatic text summarization is a process which reduces the amount of text from the original document by eliminating relatively less important content of the document. So, to create the summary with the help of the computer programs we have to retain most important points of the document.

Automatic text summarization may be used for managing huge amount of data that helps the reader to get quick overview of the entire document thus reducing the reading time. Automatic text summarizer widely used in variety of applications like document indexing, question-answering systems, sentiment analysis etc. aiding retrieval system and document classification.

On the other hand, summaries generated by the machine are also free from any biasness of human-made summary and they may be used for commercial abstracting systems.

Automatic text summarization can be classified on the type of the users the summary intended for, it may be user focused (query focused) summaries or it may be generic type that is tailored for the requirement of particular group or community type. A summary produced further categorized in abstractive and extractive summary. Extraction methods works by selecting a subset of existing word phrases or sentences from the original documents to generate the target summary. On the other hand, abstraction methods are based on understanding of the source text by using linguistic method to produce generalized summary and it requires advanced language generation and compression techniques [1].

Based on the number of documents the summarization process can accept, it can be classified as single document summarization and multi document summarization. When input set is a single document, it is called single document summarization and if the input set is a group of related documents it is called multi-document summarization.

The earliest works on text summarization have been done by sentence extraction. Here sentences are extracted after ranking them based on the combination of a variety of features like sentence position, word importance, cue phrases, title information and article's first sentence. The centroid based summarization has been proposed in [2], where centroid is a pseudo document that consists of a set of terms whose TF*IDF value (TF: Term frequency, IDF: Inverse document frequency) is greater than some predefined threshold value. In this novel research work [2] the centroid feature are also combined with some of features we discussed above, to increase the performance of the extraction based summarization system. Major drawback of sentence scoring based algorithms is that it may not be fluent due to the fact that high scoring sentences are dispersed in the summary. By considering only the structured and unstructured feature cohesion is not increased enough among the sentences in summary. Cohesion may be increased by introducing word or sentence similarity metric to get the most salient word or sentence from the document. Similarity of each sentences with other sentences are measured in [3]. Cohesion based approach sometimes uses WordNet [4] for detecting cohesion.

Though the various state-of–the-art summarization approaches use the content based analysis of the sentences for identifying summary sentences, many researchers has highlighted the effectiveness of positional information in text summarization [11] [12][13][14]. Positional information is useful for text summarization of news documents since the sentences occurring in the beginning of the document contain more important sentences due to the journalistic way of news reporting. Most existing approach that uses positional information assigns positional score to the sentences based on a function which is monotonically decreasing function of sentence position. In the paper, we use a two level discrete function of sentence position which assigns positional score of 1 to the sentences if the position is below a predefined threshold, otherwise assigns a score of 0. We observe that the sentence ranking based approach when combined with our proposed positional feature, gives better performance on the benchmark data set.

In section II we discuss the previous works related to our proposed method. Our proposed summarization method discussed in section III. Experimental results and the summary evaluation process are discussed in section IV and determinately section V concludes the paper.

2. RELATED WORK

In this section, we present the brief survey of earlier works on various single document and multi-documents text summarization methods.

A. Single Document Text Summarization

The areas of text summarization received attention of the researchers from early 50's and for the past six decade extensive work done in this area. One novel summarization approach is presented in 1958 by Luhn [5], where sentences of a document are assigned weights based on high frequent words. Disregarding the very common words (stop words) in a document, another system presented in [6] used standard keyword method (keywords are words whose frequency is

greater than a threshold) and the methods which are used to determine the weights of a sentence are: (1) Cue Method: it's basically gives impact on the most relevant sentences which is measured by the presence or absence of certain set of cue words in the cue dictionary, (2) Title Method: here sentence weight is computed based on overlap between the sentence and the title or subheading, (3) Location Method: this is based on the assumption that the highly relevant sentences occurs earlier in the document.

Many experimental research works rebuild the fact that the best correlation between human-made and automatic summaries are obtained when automatic summarizers use the combination of the above stated methods [7-10]. Researchers have used combination of statistical and other features like sentence position [11][12][13][14], topic signature[12], lexical chains[15], to compute the saliency of the sentences. Ko and Seo [16] proposes to combine two consecutive sentences into pseudo sentence (bigram) where the bigrams are considered as the context.

A machine learning based text summarization approach has been proposed in [17]. A training corpus of document-summary pairs are given as input, the summarizer uses a learning algorithm to classify the sentences as which is summary worthy and which is not. They applied the machine learning algorithm called, bagging for learning task and as a base learner C4.5 decision tree has been. An EM algorithm based summarizer introduced in [18], forms a groups of similar sentences and finally sentences are picked up from each group to form the system summary. The work in [19] considers the probability of inclusion of the sentence in a summary depending on whether the previous sentence has been included or not. A maximum entropy based model for text summarization system has been proposed in [20]. In this work, features are taken like word pairs, sentence length, sentence position and discourse features (e.g. whether sentence following any heading like "Introduction," etc.) to choose most salient sentences from the document in a summary.

In the work presented in [21], sentences are extracted in two steps which are combination of statistical method and data noise reduction. Various works in single-document summarization has been proposed which has been studied and presented by Gupta and Lehal [22]. An algorithm for language independent generic extractive summary proposed by Patel et. al., [23] uses structural and statistical parameters. They applied their approach to single-document summarization for English, Hindi, Gujarati and Urdu documents. Mann and Thompson (1987) introduced structured feature and their proposed method creates a rhetorical relations between sentence segments and documents [24].

Graph based summarization using ranking algorithm such as TextRank algorithms [25] represent the whole document as a graph of sentences or words and measure the syntactic similarities among the sentences. In [26] sentence level semantic similarity is computed to eliminate redundancy from summary.

B. Others Variation in Text Summarization Methods

A very short summary generation (usually less than 10 words) process is called headline generation. This is only an indicative summary about the content of the main document. The rule based approach that uses a set of hand crafted rules and named entity cues for headline generation has been presented in [27]. The approach used some statistical method to generate summary. The HMM (Hidden Markov Model) used for headline generation [28], Dorr et. al., [29] developed Hedge Trimmer, which is basically uses parse-and-trim based approach to generate headlines of a given document. In this work sentences are parsed through parser and then parsed sentences are compressed by eliminating unimportant low information by set of linguistic rule to generate headlines.

3. OUR PROPOSED SUMMARIZATION METHOD

The main focus of the work is to judge the effect of positional information on sentence ranking based summarization methods and improve text summarization performance with suitable use of positional information. To keep the summarization process as simple as possible we rank the sentences

based on the simple scoring system. Our proposed summarization approach has the following important steps:

Step 1: Pre-processing of the document.

Step 2: Calculate sentence score based on similarity with the centroid (T).

Step 3: Calculate Sentence score based on centrality of the sentence in the semantic similarity graph representing the document (D).

Step 4: Calculate the overall score of the sentence.

Step 5: Summary generation method that use positional information.

A. Pre-processing

Documents are preprocessed in following way.

- First stop words are removed from the documents
- Rather than perform the stemming operation on the document in next step, we use lemmatizer[1] [30] (a tool from Natural Language Processing (NLP), which does full morphological analysis to accurately identify the lemma for each word). In NLP, especially for English, has evolved into the stage where stemming would become an archaic technology if "perfect" lemmatizes exist. The stemmer reduces all words to the same stem with a common form which may not meaningful whereas lemmatization removes inflectional endings and returns the base or the root or the dictionary form of a word.

For an example the phrase "Variation of individual genes are not visible" will be reduced by using Porter stemmer to "Variat of individu gen ar not eas vis" which actually have no meaning at all semantically. But for the same input phrase, lemmatizer gives the output as "Variation of individual gene be not visible", which is more meaningful and it may be used for further processing in knowledge based retrieval process.

B. Calculate Sentence Score based on Similarity with the centroid

We consider the centroid [2] as collection of words for a given document whose weights are higher than some predefined threshold and assign scores to the sentences based on their similarity with the centroid. So, it is necessary to calculate term weight in the input document for centroid calculation.

- *Term Weighting:* We have used TF-ISF based term weighting scheme. The frequency of the j-th term is calculated as $TF_j * ISF_j$ where,

$$TF_j = \text{Frequency of the term } j \text{ in the document} \quad (1)$$

$$ISF_j = \log \frac{N}{n_j} \quad (2)$$

TF_j is the term frequency of jth term in the document, ISF_j is the inverse sentence frequency of the term and N is the total number of sentences in the document and n_j is the number of sentences containing the jth term.

- *Centroid calculation:* To select the terms which are the members of the centroid, it is necessary to set a threshold value. Here threshold is set to $\mu + \sigma$, where μ is mean of the weights of terms in the document and σ is the standard deviation of term weights. Now the terms whose TF-ISF values are higher than the defined threshold are selected as the members of the centroid.
- *Sentence Score calculation:* We calculate sentence score based on its similarity with the centroid. To calculate the similarity with the centroid, we used cosine similarity measure using equation discussed in the next sub-sections.

C. Calculate Sentence Score based on Centrality of the Sentence in the Similarity Graph

The next phase deals with the ranking of the sentences based on the centrality score [25] of the sentence in the graph representing the document. We represent a document as graph in which a sentence corresponds to a vertex of the document and the arc between any two vertices exists if similarity between the two sentences is greater than a threshold. We have a sample similarity graph in Figure 1.

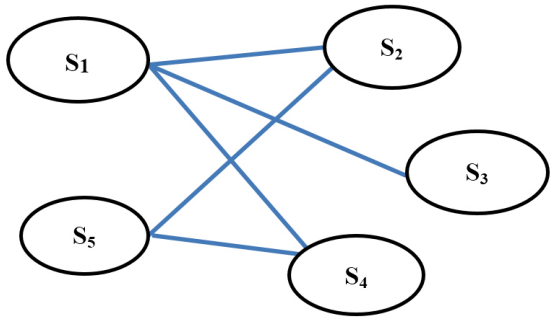

Figure 1: Sample Similarity Graph.

- *Calculation Similarity between Sentences:* Given two sentences T_1 and T_2, a joint word-set is formed for two sentences: for example, for the following two sentences, T is the joint word set constructed from the compared sentences and does not carry any redundant words.

  1. T_1: I did my homework.
  2. T_2: I completed my assignment.

  T = {I did my homework completed assignment}

  Now the vector derived from the joint word set is called lexical vector which has number of entries which is equal to number of words in the set T. The value of an first entry ($i = 1, 2, \ldots, n$) into the lexical vector corresponding to the sentence T_1 is determined by the TF × ISF weight of the corresponding term if the term is present in T_1. If the term is not present in T_1, we set to 0. Similarly the second word of the set T is checked in T_1 and so on.

  Similarity between two sentences is defined as the cosine similarity between two vectors S_1 and S_2 obtained for two different sentences:

$$\text{Cosine} - \text{sim} = \frac{s_1 \cdot s_2}{||s_1|| \cdot ||s_2||} \quad (3)$$

- *Calculate the Sentence Score:* Sentence score is basically the centrality of the sentence in the graph. Centrality score of the sentence is measured by the degree of the node that corresponds to the sentence in the graph. In other words, centrality score of sentence S is a count of other sentences to which S is similar. Two sentences are said to be similar if similarity between the sentences is greater than a predefined threshold (we set the threshold value to 0.6 for our experiments).

D. Calculate Overall Sentence Score

To calculate overall score (F) for each sentence we combine two sentence scores –(1) sentence score based on its similarity with the centroid (T) and (2) normalized centrality score (D):

$$F = (w_1 \times T + w_2 \times D) \quad (4)$$

Centrality score is normalized using traditional min-max procedure. For the best results, we set the value of w_1 to 0.6 and w_2 to 0.4.

E. Summary Generation

The summary generation step in our text summarization method is a crucial step, because it uses not only the overall sentence scores to rank sentences in descending order but also it takes into account the positional information of the sentences while selecting summary sentences. The algorithm that uses the ranked sentence list and positional information for summary generation is given below:

Step 1: Choose the top ranked sentence in the summary

Step 2: Select the next sentence from the ranked list in order and add it to the summary if the position of the sentence to be selected is <= predefined position threshold && similarity between the sentence and summary created so far is less than a predefined similarity threshold.

Step 3: Continue selecting sentences from the ranked list satisfying the positional criteria specified at step 2 until a given summary length is reached.

Step 4: If the summary of the given summary length is not reached due to the stringent criteria specified at step 2, we relax the positional criteria and select sentences in the summary in the following way:

Step 4.1: Scan the ranked list from the beginning again and set up the pointer to the beginning of the list. If the sentence pointed to by the pointer is not previously selected in the summary and similarity of

this sentence with summary created so far is <= a predefined similarity threshold, add the sentence to the summary.

Step 4.2: Advance the pointer to the next sentence and repeat the step 4.1 until the desired summary length is reached.

Our proposed summary generation process stated above has two important steps. In the first step, the criteria of sentence selection are more stringent and the sentences which have occurred earlier in the document are given preference. In the first step, the first *k* sentences of the document compete to be selected in the summary. If the summary of desired length is not achieved at the first step, the step 2 begins and positional restriction is relaxed and the remaining summary sentences are selected based on their importance in the document irrespective of their positions in the document. The reason for considering the positional information in the first step is that the sentences occurring in the beginning sections of the news documents are important due to journalistic way of news reporting.

Our summary generation algorithm considers two threshold values: (1) positional threshold (2) the similarity threshold. We experimentally decide the values of these two threshold values.

4. EVALUATION AND RESULTS

For experimental study, we have used summarization data sets released for summarization tasks carried out in several DUC (Document Understanding Conference)[1] conferences. Out of the DUC conferences, the single document summarization tasks were only considered in 2001 and 2002. We have used DUC 2001 and DUC 2002 datasets for training and testing our system. Since the target summary length in both DUC conferences was 100 words or less, we also generate 100-word summary for each document. The baseline (called lead baseline) in both years was the same: *taking the first n words of the input document*. We have used DUC 2001 task1 data set containing 309 English news articles for implementation of our system and tuning the system parameters. DUC 2002 task1 dataset was used for testing our system. DUC2002 task1 dataset contains 567 English news articles.

For summary evaluation, we have used the commonly used automatic evaluation tool called ROUGE package which is developed by Lin (2004) [31]. ROUGE measures a summary quality by counting overlapping units such as the word *n*-gram (when $n = 1$, it is called uni-gram and when $n = 2$, it is called bigram and so on), word sequences, and word pairs between the candidate summary and the reference summaries (Lin, & Hovy, 2003)[32]. We have used ROUGE version 1.5.5 for our system evaluation, evaluates summaries based on three metrics such as ROUGE-N precision, ROUGE-N recall and ROUGE-N F-score, where N can be 1, 2, 3, 4 etc. We have considered ROUGE-1 F-score for evaluating the system generated summaries, because among the various ROUGE scores, the unigram-based ROUGE score (ROUGE-1) has been shown to most agree with human judgment (Lin, & Hovy, 2003)[32].

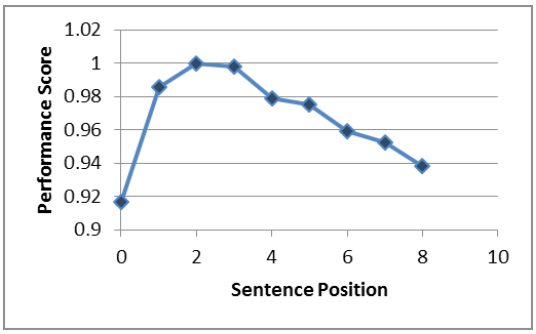

Figure2: The effect on Performance Score of our proposed system when position threshold is varied.

During experimentations, we observed that the system performance is affected when we vary threshold on sentence position during summary generation. In order to select an optimal position threshold value, an experiment is set up. In this experiment, a DUC 2001 dataset is used for tuning the threshold values. In the devised experiment on adjusting the threshold value for sentence position, the summarizer was run on the data set with a position threshold value, with this value ranging from 1 to 10. Each time the experiment was run, the position threshold value is

[1]http://www-nlpir.nist.gov/projects/duc/guidelines/2001.html
[2]http://www-nlpir.nist.gov/projects/duc/guidelines/2002.html

incremented by 1 and the summaries generated by the summarizer are evaluated using the ROUGE-1 average F-score [32]. After all summaries for the input documents are generated, the average ROUGE-1 score for each threshold value is calculated. Since the whole idea of having a threshold value is to maximize the summarization performance, a score was obtained by normalizing these values by dividing each average value by the maximum obtained average. So the value of one represents the highest obtainable summarization performance using the devised algorithm. We have shown in Figure 2 how the summarization performance is affected when the positional threshold value is varied. As we can see from Figure 2, the best performance is achieved when the positional threshold value is set to 2. During this experiment, we set the similarity threshold value to 0.7.

During system run, we set the summary length to 100 words. We consider the first 100 words of a document as a baseline summary to check whether the system generated summary is better than the baseline summary for a document. The ROUGE score obtained by our proposed system is shown in Table 1.

Table 1
System comparison results on DUC2002 data, the summaries are stemmed but stop-words not removed. 95% confidence intervals are shown in brackets

Systems	*ROUGE-1 F-score*
Proposed System	0.4880 [0.4706 – 0.5041]
Sys 28	0.4830 [0.4757-0.4898]
Sys 21	0.4757 [0.4688 - 0.4829]
DUC baseline	0.4751 [0.4679-0.4824]
Sys29	0.4685 [0.4616 - 0.4758]
Sys 27	0.4651 [0.4576- 0.4728]
Sys31	0.4599 [0.4528- 0.4664]

Our proposed system has also been compared with five top-performing systems participating in the single document summarization task of DUC2002. To compare our proposed method with the DUC systems, we have used the summaries released by DUC official on the web. Table 1 also shows the comparisons of our proposed summarization approach with the DUC baseline and the top five systems, sys 28, sys 21, sys 29, sys 27 and sys 31 participating in DUC 2002. As we can see from the Table 1, our proposed summarization approach performs significantly better than the DUC baseline and the performance of the approach is comparable to performances of other systems which it is compared to.

Table 2
Performance Comparisons of the proposed System with positional feature and the proposed system without positional feature (the summaries are stemmed but stop-words not removed. 95% confidence intervals are shown in brackets)

Systems	*ROUGE-1 F-score*
Proposed System with our defined positional feature	0.4880 [0.4706 – 0.5041]
Proposed System without positional feature	0.45150 [0.43156 – 0.47171]

To prove the effectiveness of our proposed method of using positional information, we develop a version of our system excluding the positional feature. We observe that exclusion of the positional feature from our proposed system drastically degrades the summarization performance. We have shown the results of our study in Table 2.

Our investigation in the DUC 2002 data set reveals that the reason of having better performance with our defined positional method is that DUC 2002 dataset is a collection of news articles. Due to the journalistic way of new reporting, the sentences occurring in the beginning section of the document carry important information.

We have also compared our proposed method with an existing single document summarization method [33] that also positional information. We have observed that the performance of the system presented in [33] is also comparable to our system. Despite this fact, our system differs from the system proposed in [33]. Like the system presented in [33], our proposed system does not use any sophisticated keyphrase extraction module. Moreover, the work in [33] does not report how much performance difference can be found if the positional information feature is removed. We have shown in Table 3 the comparison of our proposed system with the system presented in [33].

Table 3
Performance comparison of the proposed system and the system presented in [33] (the summaries are stemmed but stop-words not removed. 95% confidence intervals are shown in brackets)

Systems	*ROUGE-1 F-score*
Our Proposed System with our defined positional feature	0.4880 [0.4706 – 0.5041]
Sarkar(2013)[33]	0.4855 [0.4783 - 0.4925]

5. CONCLUSION AND FUTURE WORKS

In this work, we describe a summarization approach that use sentence centrality score and centroid based sentence score for calculating importance of sentence content. A novel use of positional information has been used during summary generation. The experimental study reveals that the performance of our system is comparable to the state-of-the art single document summarization systems.

We observe that, some sentences selected in the summary are verbose. We have planned to trim those sentences automatically to generate more concise summary. The deep linguistic analysis of the sentences may help to eliminate irrelevant elements from the summary. We will investigate this issue in future.

Acknowledgment

This research work has received support from the project entitled "Design and Development of a System for Querying, Clustering and Summarization for Bengali" funded by the Department of Science and Technology, Government of India under the SERB scheme.

References

[1] Kasture, N.R., et. al., "A survey on methods of abstractive text summarization." Int. J. Res. Merg. Sci. Technol 1.6 (2014): 53-57.

[2] D.R. Radev, H. Jing, M. Sty and D. Tam, "Centroid-based summarization of multiple documents," Journal of Information Processing and Management, Elsevier, Volume 40, No. 6, 2004, pp. 919-938.

[3] Kupiec, J., Pedersen, J.O. and Chen, F. (1995). A trainable document summarizer. Proceedings of18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 68-73.

[4] Miller, George A. "WordNet: a lexical database for English." Communications of the ACM 38.11 (1995): 39-41.

[5] H.P. Luhn, "The Automatic Creation of Literature Abstracts", Presented at IRE National Convention, New York, 159-165, 1958.

[6] H.P. Edmundson., "New methods in automatic extracting", Journal of the ACM, 16(2):264-285, April 1969.

[7] P. Baxendale, "Man-made index for technical literature-An experiment," IBM Journal of Research and Development, Vol. 2, No. 4, 1958, pp. 354-361.

[8] H.P. Edmundson, "New methods in automatic extracting," Journal of the Association for Computing Machinery, Vol. 16, No. 2, 1969, pp. 264-285.

[9] K. Sarkar, "An approach to summarizing Bengali news documents," Proceedings of the International Conference on Advances in Computing, Communications and Informatics, ACM, 2012, pp. 857-862.

[10] K. Sarkar, "Bengali text summarization by sentence extraction," Proceedings of International Conference on Business and Information Management, NIT Durgapur, 2012, pp. 233-245.

[11] C_Y.Lin and E. Hovy "Identifying Topics by Position," Proc. Fifth Conf. Applied Natural Language processing, pp.283-290, http://dx.doi.org/10.3115/974599, 1997.

[12] E. Hovy and C_Y. Lin, "Automated text Summarization and the Summarist System", Proc. Workshop Held at Baltimore, Maryland(TIPSTER '98'), pp.197-214, http://dx.doi.org/10.3115/1119089119121,1998.

[13] R. Katragadda, P. Pingali, and V. Varma, "Sentence Position Revisited: A Robust Light Weight Summarization 'Baseline' Algorithm," Proc. Third Int'l Workshop Cross Lingual Information Access: Addressing the Information Need of Multilingual Societies, pp. 46-52, http://portal.acm.org/citation.cfm?id=1572433.1572440,2009.

[3]Retrieved from http://wn-similarity.sourceforge.net/
[4]Retrieved from http://nltk.org

[14] Ouyang, Y., Li, W., Lu, Q. and Zhang, R., 2010, August. A study on position information in document summarization. In Proceedings of the 23rd International Conference on Computational Linguistics: Posters (pp. 919-927). Association for Computational Linguistics.

[15] Barzilay, R. and Elhadad, M., 1999. Using lexical chains for text summarization. Advances in automatic text summarization, pp. 111-121.

[16] Ko, Y. and Seo, J., (2008), An effective sentence-extraction technique using contextual information and statistical approaches for text summarization. Pattern Recognition Letters, 29(9), pp. 1366-1371.

[17] K. Sarkar, M. Nasipuri, S. Ghose, "Using Machine Learning for Medical Document Summarization," International Journal of Database Theory and Application, Vol. 4, No. 1, pp. 31-48, 2010.

[18] Y. Ledeneva, R.G. Hernández, R.M. Soto, R.C. Reyes and A. Gelbukh, "EM clustering algorithm for automatic text summarization," In Advances in Artificial Intelligence, Springer Berlin Heidelberg, 2011, pp. 305-315.

[19] J.M. Conroy and D.P. O'Leary, "Text summarization via hidden Markov models and pivoted QR matrix decomposition," Tech. Rep., University of Maryland, College Park, 2001.

[20] M. Osborne, "Using maximum entropy for sentence extraction," Proceedings of the ACL-02, In Proceedings of Workshop on Automatic Summarization, (Philadelphia, Pennsylvania), Annual Meeting of the ACL, Association for Computational Linguistics, Morristown, Vol. 4, 2002.

[21] W. Jung, Y. Ko, and J. Seo, "Automatic Text Summarization Using Two-Step Sentence Extraction", AIRS 2004, LNCS 3411, pp. 71 – 81, (2005).

[22] V. Gupta, G. Singh Lehal, "A Survey of Text Summarization Extractive Techniques", Journal of Emerging Technologies in Web Intelligence, Vol. 2, No. 3, pp. 258-268, August (2010).

[23] Patel. A., Siddiqui. T., Tiwary. U.S., "A language independent approach to multilingual text summarization", Conference RIAO2007, Pittsburgh PA, U.S.A., (2007).

[24] Mann, W.C., Thompson, S.A.: Rhetorical Structure Theory: A Theory of Text Organization. Technical Report ISI/RS-87-190 (1987).

[25] Mihalcea, R., Graph-Based Ranking Algorithms for Sentence Extraction, Applied to Text Summarization. In: Proceedings of the ACL 2004 (2004).

[26] Chatterjee, N. and Mohan, S., (2007), October. Extraction-based single-document summarization using random indexing. In 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007) (Vol. 2, pp. 448-455). IEEE.

[27] Sarkar, K. and Bandyopadhyay, S., (2005), February. Generating headline summary from a document set. In International Conference on Intelligent Text Processing and Computational Linguistics (pp. 649-652). Springer Berlin Heidelberg.

[28] M. Banko, V. Mittal and M. Witbrock, "Headline generation based on statistical Translation," Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics (ACL-2000), Hong Kong, 2000, pp. 318–325.

[29] D. Zajic, B. Dorr and R. Schwartz, "Automatic Headline Generation for Newspaper Stories," Workshop on Automatic Summarization, Philadelphia, PA, 2002, pp. 78-85.

[30] Manning, Christopher D., Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pp. 55-60.

[31] Lin, C. Y. (2004, July). Rouge: A package for automatic evaluation of summaries. In Text Summarization Branches Out: Proceedings of the ACL-04 Workshop (pp. 74-81).

[32] Lin, C. Y., & Hovy, E. (2003, May). Automatic evaluation of summaries using n-gram co-occurrence statistics. In Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1 (pp. 71-78). Association for Computational Linguistics.

[33] K. Sarkar. (2013). Automatic single document text summarization using key concepts in documents. Journal of information processing systems, 9(4), 602-620.