ISSN: 2321-7782 (Online)
Volume 2, Issue 11, November 2014
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

A Comprehensive Review on Fuzzy Logic & Latent Semantic Analysis Techniques for Improving the Performance of Text Summarization

Patil Pallavi D[1]
M.E Scholar Computer Engineering
Zeal Education's Dnyanganga College of Engineering & Research
Narhe (Pune), India

Mane P.M[2]
Asst. Professor Computer Engineering,
Zeal Education's Dnyanganga College of Engineering & Research
Narhe (Pune), India

Abstract: In this new generation, the quantity of information on the web is rising day by day, a significant quantity of time is exhausted in searching for related documents, it is tricky to drag out the information rapidly and most efficiently. There are so many text materials available on the web, in order to drag out the most related information from it, we need a good instrument .This problem is solved by the Automatic Text Summarization mechanism. "Text Summarization" is a process of creating a shorter version of original text that contains the important information.. The mission of this review paper is addressed Fuzzy logic & latent semantic Analysis techniques of various researchers are discussed from beginning of this research to this modern age.

Keywords: NLP, Text Summarization, Fuzzy logic, Latent Semantic Analysis.

I. Introduction

Text Summarization is a process of creating a shorter version of original text that contains the important information. The amount of information on the web is growing day by day. A considerable amount of time is wasted in searching for relevant documents. Hence text summarization technique came into existence which created a short summary for the text document by choosing important sentences of the document. An Automatic text summarization works very well on structured documents such as news articles, research publications and reports. Automatic Text summarization is tremendously helpful in tackling the information overload evils. It is the technique to recognize the most important pieces of information from the document, omitting irrelevant information and minimizing details to generate a dense consistent summary document. Automatic text summarization (ATS) deals with the process of sinking one or more texts to a summary text that very dense the most relevant information given a convinced information need, by concatenating literal fragments of the original text(s) or by identifying key concepts and generating a new text, all this by means of automatic techniques for natural language processing (NLP).

Natural Language Processing (NLP) is a field of computer science, artificial intelligence and linguistics as that entire particular field bring it into play. Generally it deals with the interactions between machines and human languages that complete task on inspecting, understanding and generating the language, which human use naturally in order to interact with computers in both oral and written contexts using natural human languages instead of computer languages. It is an interdisciplinary area based on versatile arena of study including computer engineering, which provides methods for model figure, algorithm devise and exploit; linguistics, which categorizes linguistic forms and practices; mathematics, which provides formal models and methods; psychology, which studies models and theories of human behaviour; statistics, which offers procedures for predicting measures based on sample records; and biology, which travels around the underlying architecture of linguistic processes in the human brain[25].

1.1 Categorization of Text Summarization:

1) Abstract vs. Extract summary - Abstraction is the process of paraphrasing sections of the source document whereas extraction is the process of picking subset of sentences from the source document and presents them to user in form of summary that provides an overall sense of the documents content.

2) Generic vs. Query-based summary - Generic summary do not target to any particular group. It addresses broad community of readers while Query or topic focused queries are tailored to the specific needs of an individual or a particular group and represent particular topic.

3)Single vs. Multi-document Summary - Single document summary provide the most relevant information contained in single document to the user that helps the user in deciding whether the document is related to the topic of interest or not whereas multi-document summary helps to identify redundancy across documents and compute the summary of a set of related documents of a corpus such that they cover the major details of the events in the documents, taking into account some of the major issues : the need to carefully eliminate redundant information from multiple documents and achieve high compression ratios; information about document and passage similarities, and weighting different passages accordingly; the importance of temporal information; co-reference among entities and facts occurring across documents.

4) Indicative vs. informative: An indicative summary provides merely an indication of the principal subject matter or domain of the input text(s) without including its contents. After reading an informative summary, one can explain what the input text was about, but not necessarily what was contained in it. An informative summary reflects (some of) the content, and allows one to describe (parts of) what was in the input text.

5) Background vs. just-the-news: A background summary assumes the reader's prior knowledge of the general setting of the input text(s) content is poor, and hence includes explanatory material, such as circumstances of place, time, and actors. A just-the-news summary contains just the new or principal themes, assuming that the reader knows enough background to interpret them in context.

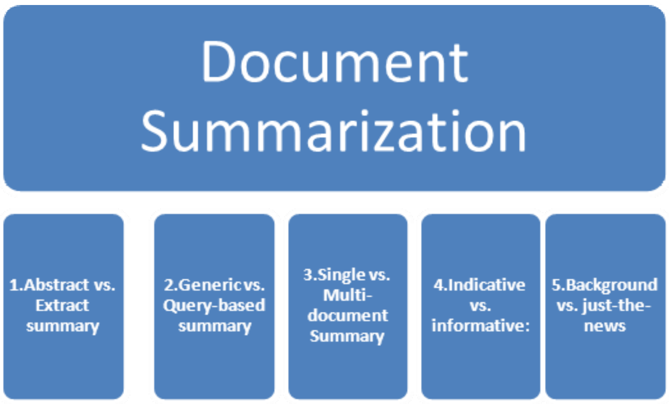


Figure 1.1: Categorization of Document Summarization

1.2 Text Summarization Approaches:

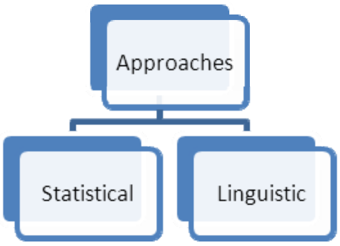
There are different types of summarization approaches depending on what the summarization method focuses on to make the summary of the text.

a) Statistical approaches:

In Statistical methods, sentence selection is done based on word frequency, indicator phrases and other features regardless of the meaning of the words. These methods are based on the idea that text surface cues are the most obvious indication of the text contents. There are several methods for determining the key sentences such as, The Title Method, The Location Method, The Aggregation Similarity Method, The Frequency Method etc.

b) Linguistic approaches:

Linguistic approaches are based on considering the connections between words and trying to find the main concept by analyzing the words. There are some methods such as, Lexical chain, WordNet,Graph theory, Clustering etc.



1.2 Approaches to Document Summarization

1.3 Main Steps in Text Summarization:

These steps are topic identification, interpretation and summary generation.

Topic Identification: In topic identification step, the most prominent information in the text is identified. Usually the system assign different precedence to different parts of the text like sentence, words, and phrases; then a fuser module mix the scores of each part in order to find the total score for a part. At last, the system presents the N highest score parts in accordance with predefined length. Several techniques for topic identification, including methods based on Position, Cue Phrases, word frequency, and Discourse Segmentation have been reported in the literature [1].

Interpretation: Abstract summaries need to go through interpretation step. In This step, different subjects are fused in order to form a general content.

Summary Generation: In this step, the system uses text generation method which itself is still an open research topic that has lots of similarities with text summarization. This step includes a range of various generation methods from very simple word or phrase printing to more sophisticated phrase merging and sentence generation.

2) REVIEW ON SUMMARIZATION TECHNIQUES

Extractive summarizer aims at picking out the most relevant sentences in the document while also maintaining a low redundancy in the summary. There are summarization techniques which works on the most relevant sentences

A. Term Frequency-Inverse Document Frequency (TFIDF) based

B. Cluster based

C. Graph theoretic approach

D. Machine Learning approach

E. Text summarization with neural networks

F. Query based extractive text summarization:

G. Automatic text summarization based on fuzzy logic

H. LSA (Latent Semantic Analysis) based

In this review we primarily aim to understand the how Fuzzy logic based & latent Semantic Analysis techniques have been used to build summarization systems.

2.1 General:

As accomplishment of automatic passage summarizer was often cited in the oldest publication in 1958 by H. P. Luhn [2]. This method is based on the word frequency and clearly emphasized that the most frequent words represent the most important concepts of the text. In the next step, frequencies are used to score and decide on sentences to be extracted. In this paper, application of machine is emphasized and expressed that because of the absence of the variations of human capabilities and orientation, auto abstracts have a high degree of reliability, consistency, and stability, as they are the product of statistical analysis of the author's own words. This paper is worthy of being appreciated as it is almost the earliest paper in this arena of automatic text abstraction [3]. Besides this, the proposed method mostly works on scientific article and journal paper.

P. B. Baxendale [4] in 1958 investigated machine techniques for reducing scientific credentials to their important sharp indices. Human scanning patterns were tried to be simulated here for selecting topic sentences special phrases. It was declared in this paper that sentence' position and containing certain cue-words (i.e., words like 'crucial' or 'pertinent' etc.) or the word exist in the heading are special indicator for being in the salient parts of the document.

G. J. Rath et al [5] in 1961 in their research scrutinized about four types of lexical indicators of content to determine which one is the best for detecting relevant document from repository and to answer a set of question. After their experiment, it was claimed that utilizing complete text is better than only abstract for answering question. But for distinguishing relevant document only abstract is enough. They also proclaimed that purely statistical method of producing extracts was suspected of being inadequate, and hence other methods were sought.

H. P. Edmundson [6] in 1969 accomplished a notable progress after ten years of the beginning of the research on text recapitulation. He described three additional methods with the standard keyword method, disregarding the very high frequency common words to determine the sentences' weight. Those are:

1. Cue Method: The hypothesis of this technique is that the presence or absence of certain cue words will compute the significance of a sentence.
2. Title Method: The weight of a sentence is calculated as a sum of all the content words materializing in the title, headings and sub-headings of a text.
3. Location Method: Here relevance is assumed on the basis of location, expressed that sentences taking place in initial position of paragraphs have a higher probability of being pertinent.

The result was very fruitful and assumed that by using a combination of these three schemes the best correlation between the automatic and human-made abstracts can be achieved. This paper emphasized on indicative abstracts rather than on the production of informative abstracts. So, if a user doesn't know about the document to be summarized earlier, can't get summary through this proposed methodology.

In [7], paper shows recent techniques and challenges on advances of automatic text summarization. Special attention is paid to the latest trends in text summarization. Author discusses the key challenges in automatic text summarization. These are inherent problem of overlapping of sets of similar text units or paragraphs; documents which contain long sentences are still a problem; another challenge is the word sense ambiguities which are inherent to natural language. The problem is that matching a system summary against the ideal summary is very difficult to establish. The problem of providing much accurate or efficient result for automatic text summarization. Author has discussed different types of summarization approaches

[8] Depending on what the summarization method focuses on to make the summary of the text. Automatic document summarization is extremely helpful in tackling the information overload problems. It is the technique to identify the most important pieces of information from the document, omitting irrelevant information and minimizing details to generate a compact coherent summary document. Author has given the types of summaries - Abstract vs. Extract summary, Generic vs. Query-based summary, Single vs. Multi-document summary, Indicative vs. Informative, Background vs. Just the- news. Author

has mentioned in detail the different approaches: Graph Theoretic Approach, Text summarization based on fuzzy logic, Text summarization using cluster based method.

In [1], paper author has concentrating on extractive summarization methods. An extractive summary is selection of important sentences from the original text. The importance of sentences is decided based on statistical and linguistic features of sentences. There are two broad methods of text summarization extractive and abstractive summarization. An extractive summarization method consists of selecting important sentences, paragraphs etc. from the original document and concatenating them into shorter form. An abstractive summarization method consists of understanding the original text and re-telling it in fewer words.

Paper [9], defines the most important criteria for a summary and different methods of text summarization as well as main steps for summarization process is discussed. There are different approaches for sentence selections are presented in order to generate a summary from a text. Author has explained detailed steps for summary of document, which are topic identification, interpretation and summary generation. Also author has given the different approaches for scoring and selecting sentences.

In paper [10], it is said that sentence scoring is the technique most used for extractive text summarization. Paper describes 15 sentence scoring algorithm and performs a quantitative and qualitative assessment of these 15 algorithms. Example Sentence length: It works as follows: (i) Calculate the largest sentence length; (ii) Penalize sentences larger than 80 percent of the largest sentence length; (iii) Calculate the Sentence Length Score for all other sentences. Author has used three different datasets (News, Blogs and Article contexts) . It is a single document summarization.

2.2 Fuzzy Logic Based Method

L. Suanmali et al.[11] has proposed text summarization based on fuzzy logic method to extract important sentences as a summary. This paper focuses on extraction approach. The sentence selection method had used to obtain the suitable sentences by assigning some numerical weighting to each sentences and selecting the best one. The 8 important features are used and calculate their score for each sentence. These are Title word, Sentence length, Sentence Position, Numerical data, Term weight, Sentence to Sentence Similarity, existence of thematic words and Proper Nouns. Author presents the some summarization approaches and describes pre-processing and the important features. The experimental results compared with the baseline summarizer and Microsoft Word 2007 summarizers.

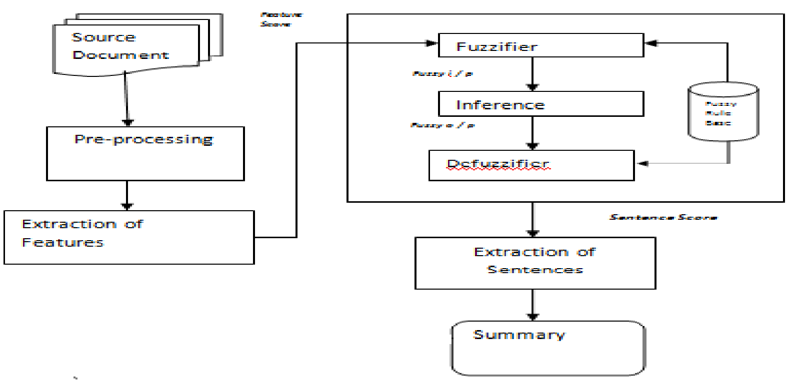


Figure 2.2: Text summarization based on fuzzy logic system architecture [6]

In paper [12], the focus is on Text Summarization by Extraction using Fuzzy Logic. Author has used the hybrid method-Statistical and Linguistic methods .The evaluation is carried out by using precision and recall method. Using all the 8 feature scores, the score for each sentence are derived using fuzzy logic method. The fuzzy logic method uses the fuzzy rules and triangular membership function. The fuzzy rules are in the form of IF-THEN. The triangular membership function fuzzifies each score into one of 3 values that is LOW, MEDIUM HIGH. Then we apply fuzzy rules to determine whether sentence is unimportant, average or important. This is also known as defuzzification. .

In paper [13] author has analyzed some state of methods for text summarization. He discussed the main disadvantages of these methods and proposed a new method of text summarization using fuzzy logic. The problems of machine learning method is features used in this method e.g. the occurrence of proper nouns have binary attributes such as zero and one which sometimes are not exact. It is clear that some of the attributes have more importance and some have less and so they should have balance weight in computations and hence the fuzzy logic is used to solve this problem. The implementation of text summarization is carried out using MATLAB.

In this [14] paper the focus is on the automatic text summarization by sentence extraction. Author used fuzzy logic to extract the sentences. The eight features and their calculated scores are used compare the results with the baseline summarizer and Microsoft

Word 2007 summarizers. Author has explained in detail the definitions and terminology of each eight sentence features with their formulas.

This paper[15], proposed an automatic text summarization approach based on sentence extraction using fuzzy logic, genetic algorithm, semantic role labelling and their combinations to generate high quality summaries. GA used in text summarization.

Author has mentioned the benefits of the genetic algorithm in the optimization problem in for feature selection. Fuzzy IF-THEN rules were used to balance the weights between important and unimportant features.

In [16], author has applied the fuzzy logic approach for important features to extract the sentences. In this paper the result is compared with baseline summarizer and Microsoft Word 2007 summarizers. Author has used 9 sentence features and 30 test documents in DUC2002 data set. Fuzzy logic techniques in the form of approximate reasoning provide decision-support.

In this [17] paper, author proposed a fuzzy-rough set aided method to extract key sentences as a summary of a document by estimating the relevance of sentences through the use of fuzzy-rough sets. This method uses senses rather than raw words to avoid the similar semantic meaning .Author has proposed the system in which the four phases of processes are organized. These are: Document Pre-processing, Word Sense Disambiguation, Semantic Patterns Retrieval, and Key Sentences Extraction.

2.3 Latent Semantic Analysis

In [18] Author defined Latent semantic analysis (LSA)is a technique for extracting the hidden dimensions of the semantic representation of terms, sentences, or documents, on the basis of their contextual use. There are some approaches that incorporate LSA in their MDS systems. Steinberger and Krista'ns (2007) developed an algorithm for single document summarization using LSA and apply it to the problem of MDS. Their system works as follows: first a term-by-sentence matrix (TSM) A is created where each row represents a term and each column represents a sentence. The cells of the matrix contain weighted term frequencies. Singular Value Decomposition (SVD) is applied to A which breaks down the original TSM into r base vectors which are linearly independent The result of SVD are three matrices T , S and DT . These sub matrices are used to calculate a ranking matrix S2DT, which is used to rank the sentences of the document collection. For each sentence the length of the corresponding vector in the ranking matrix is calculated. The sentences with the highest scores (length of vector divided by (number of terms) 0:4) are selected for the summary. To avoid redundancy only sentences that are not similar to sentences already in the extract, measured with the cosine similarity in the original term space, are added to it.

The evaluation was done on 50 DUC 2005 clusters containing 1300 documents in total. The resulting extracts were evaluated with ROUGE. The MDS system scored better than 27 other systems and worse than 5. In this approach SVD helps to capture the relationships between words by analysing co-occurrence patterns. Thus terms and sentences can be grouped on a more semantic basis than on word overlap only. The authors claim that each singular vector (rows in DT) represents a salient and recurring word usage pattern. They assume that each word usage pattern describes one topic in the document collection. Thus the topics of a document collection are represented by the singular vectors and the magnitude of the singular value represents the degree of importance of the corresponding pattern within the document set. The authors assume that the sentence that best represents this pattern will have the largest index value within the singular vector. The importance and the coverage of the word usage patterns within the sentences are taken into account by calculating the length of the sentence vectors in S2DT. The authors seek to choose sentences for their summary that have the greatest combined weight across all important topics. The problem is that in this approach the number of topics is linked to the number of singular values. In SVD the number of singular values and vectors is equal to the rank of the matrix, which is the number of linearly independent rows or columns of A. To reveal the latent semantic structure and the themes in a document collection the dimensions are reduced. Choosing the number of dimensions to keep also determines the number of topics. The unresolved problem is to determine the number of topics relevant for the document collection in advance, when the system claims to be fully automatic. The relation between word usage pattern and topic is debatable: is one word usage pattern equal to a topic of a document collection. In addition, even if there is a one-to-one relationship between patterns and topics, why are sentences selected that have the highest score over all topics? One could also choose one sentence for each topic. Also, only the resulting extracts were evaluated using ROUGE, where N-grams are compared to human written abstracts. A problem here is that human written abstracts are compared to extracts. ROUGE only takes matches of N-grams into account, but since humans created the abstracts, the content of the summaries compared can be similar but the words used can be very different. But evaluation of summaries is an open and very controversial problem. There was no evaluation of whether the sentences selected for the summary really do represent the most important topics of the document collection.

Another approach using sentence clustering and LSA is described in Bing et al. (2005). After a term-by-sentence matrix is built and factorized using SVD the sentences are compared pair wise. The sentences with the highest similarity are merged to a sentence cluster, called a fake sentence. This cluster and the rest of the sentences are used to create a new matrix and again (after applying SVD) all sentences are compared pair wise and so on. This process is repeated until the predefined number of clusters is reached. For each cluster the centroid sentence is determined. The centroid sentences are then sorted and included in the summary. This approach was tested on a dataset consisting of 20 documents clusters of 7-9 news articles from various websites. Judges were asked to grade the 20 extracts created by the system using a score between 1 (bad summary) and 5 (good summary). 75% of the summaries were marked with a score of 3 or 4. In this system only matrix D is used to cluster the sentences. I think it is unreasonable that the singular values are not taken into account since they correspond to the importance of each concept within the document collection. It is very time and memory consuming to build a new matrix and perform SVD every time two sentences are joined. The fake sentence is longer than the other sentences, therefore will score higher similarity with other sentences and attract more sentences. Again only the summary was evaluated as human judges were asked to score the summary. The judges were not provided with other summaries, to which they could compare the automatically created abstracts. It is not clear which instructions the judges received or if they had access to the original documents.

The Embra system for DUC 2005 (Hachey et al., 2005) uses LSA to build a very large semantic space to derive a more robust representation of sentences. In this space not only the documents that are summarized are included but also other documents from the DUC 2005 and AQUAINT corpus. First a term-by-document matrix (TDM) is created, SVD is applied and the resulting sub matrices are reduced to 100 dimensions. From the sub matrices a sentence representation is built, where each sentence is represented by a vector that is the average of the vectors of the words the sentence contains. This sentence representation is then passed to an MMR-style algorithm, which determines relevancy and redundancy (see above). The sentences with the highest MMR score are selected for the summary. In contrast to Goldstein et al. (2000) here the redundancy calculation is not based on single word overlap but on the word usage patterns revealed by LSA. SVD was performed on a term by document matrix, but it was not evaluated how the size of a semantic space influences redundancy identification. The approaches described here all lack the evaluation of the influence of LSA on detecting redundant information. Only indirect evaluation of the redundancy identification was carried out by evaluating the resulting summary. The influence of parameters like the number of dimensions (k), size of the semantic space or vocabulary included has not been properly analysed. However these parameters might have a great impact on the quality of redundancy identification. Once again in order to optimize these parameters a direct evaluation of the sentence clusters is required [19].

Asef, Kahani, Yazdi and Kamyar in 2011 proposed LSA for multi-document for text summarization for Persian language. In this LSA again it performs term selection. Its Limitation: used for Persian language differs from English language both morphologically and semantically [20].

Ozsoy, Cicekli in 2010 proposed LSA for multi-document for text summarization in Turkish language. In this LSA again explains its two approaches Cross and Topic which performs sentence selection. Its limitation is that it is used for Turkish language and sentence selection is done based on similarity of terms [21].

Makbule Gulcin Ozsoy , Ferda Nur Alpaslan and Ilyas Cicekli in 2011 proposed Text summarization using Latent Semantic Analysis In this paper, different LSA-based summarization algorithms are explained, two of which are proposed by the authors of this paper. The algorithms are evaluated on Turkish and English documents, and their performances are compared using their ROUGE scores. He mentioned that Latent Semantic Analysis (LSA) is an algebraic-statistical method which extracts hidden semantic structures of words and sentences. It is an unsupervised approach which does not need any training or external knowledge. LSA uses context of the input document and extracts information such as which words are used together and which common words are seen in different sentences. The summarization algorithms that are based on LSA method usually contain three main steps:

1. Input matrix creation: an input document needs to be represented in a way that enables a computer to understand and perform calculations on it. This representation is usually a matrix representation where columns are sentences and rows are words/phrases.
2. Singular Value Decomposition (SVD): SVD is an algebraic method that can model relationships among words/phrases and sentences. In this method, the given input matrix A is decomposed into three new matrices as follows:

   $A = U \sum V^T$

   A: Input matrix ($m \times n$)

   U: Words × Extracted Concepts ($m \times n$)

   $\sum$: Scaling values, diagonal descending matrix ($n \times n$)

   V: Sentences × Extracted Concepts ($n \times n$)
3. Sentence selection: using the results of SVD different algorithms are used to select important sentences.[22]

2.4 Evaluating Summary:

In this [23] paper, the investigation of the correlation between ROUGE and human evaluation of extractive meeting summaries is carried out. Both human and system generated summaries are used. The human evaluation of different summaries and calculated ROUGE scores are examined with their correlation .The better correlation can be achieved between the ROUGE scores and human evaluation. In text summarization, ROUGE has been correlate well with human evaluation when measuring match of content units.

In this [24] paper they explain the ideas of automatic text summarization approaches and the taxonomy of summary evaluation methods. The taxonomy of summary evaluation measures can be found in the below fig. 2.4. Text quality is often assessed by human annotators. They assign a value from a predefined scale to each summary. The main approach for summary quality determination is the intrinsic content evaluation which is often done by comparison with an ideal summary. For sentence extracts, it is often measured by co-selection.

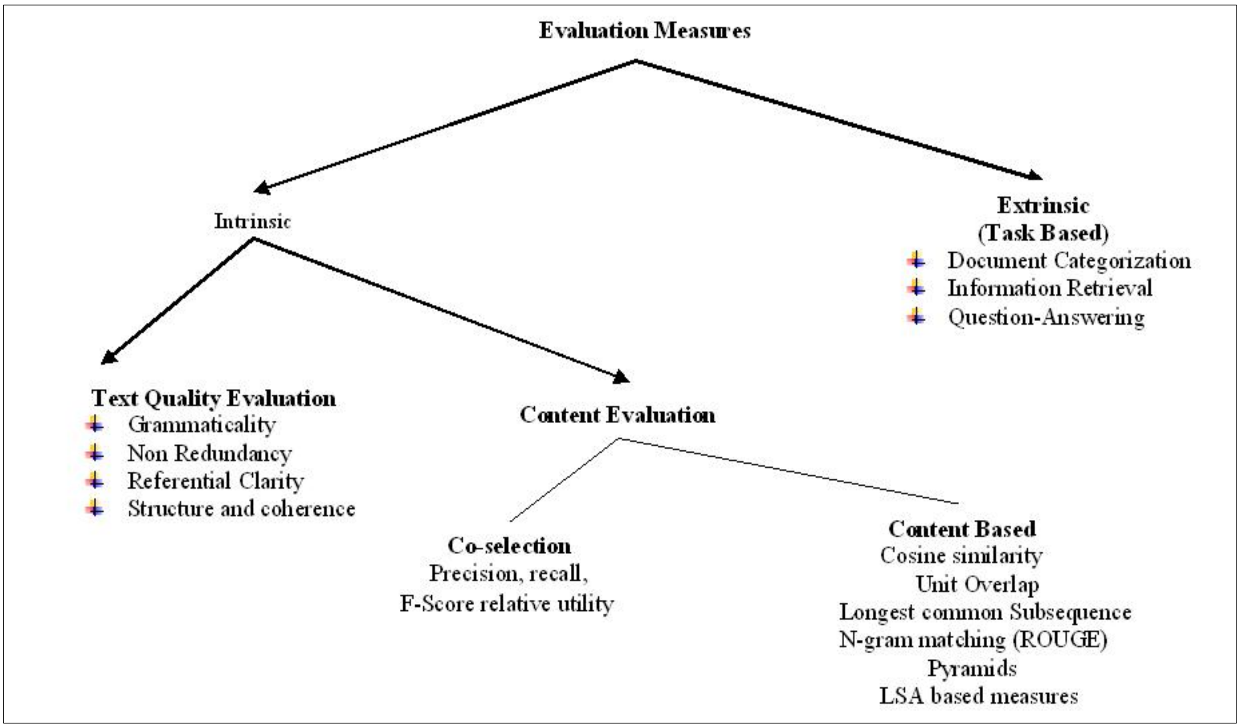


Fig 2.4: The taxonomy of summary Evaluation Measure

Evaluation methods are useful in evaluating the usefulness and trustfulness of the summary. In summary, evaluating the qualities like comprehensibility, coherence, and readability is really difficult. System evaluation might be performed manually by experts who compare different summaries and choose the best one. Two main criterion for evaluating the proficiency of a system is precision and recall which are used for specifying the similarity between the summary which is generated by a system versus the one generated by human. The weighted harmonic mean of precision and recall is called as F-measure. These terms are defined by following equations

$$Precision = Correct / (Correct + Wrong),$$

$$Recall = Correct/(Correct + Missed)\ ,$$

$$f - measure = 2 \times (precision \times recall) / (precision + recall)$$

References

1. Vishal Gupta and Gurpreet Singh Lehal "A survey of Text summarization techniques ",Journal of Emerging Technologies in Web Intelligence VOL 2 NO 3 August 2010.
2. Hans P. Luhn, “The Automatic Creation of Literature Abstracts,” IBM Journal of Research and Development, vol. 2, no. 2, pp. 159-165, 1958.
3. Dipanjan Das and André F. T. Martins, “A survey on Automatic Text Summarization,” Language Technologies Institute, Carnegie Mellon University, 2007. [Online] Available: http://www.cs.cmu.edu/~afm/Home_files/Das_Martins_survey_summarization.pdf (May 21, 2013)
4. P. B. Baxendale, “Machine-made Index for Technical Literature -An Experiment,” IBM Journal of Research and Development, vol. 2, no. 4, pp. 354-361, October 1958.
5. G. J. Rath, A. Resnick and T. R. Savage, “Comparisons of four types of lexical indicators of content,” Journal of the American Society for Information Science and Technology, vol. 12, no. 2, pp. 126-130, April 1961.
6. H. P. Edmundson, “New Methods in Automatic Extracting,” Journal of the Association for Computing Machinery, vol. 16, no. 2, pp. 264-285, April 1969.
7. Oi Mean Foong ,Alan Oxley and Suziah Sulaiman "Challenges and Trends of Automatic Text Summarization ",International Journal of Information and Telecommunication Technology Vol.1, Issue 1, 2010 .
8. Archana AB, Sunitha. C ,"An Overview on Document Summarization Techniques" ,International Journal on Advanced Computer Theory and Engineering (IJACTE) ,ISSN (Print) : 2319 ˝U 2526, Volume-1, Issue-2, 2013 .
9. Saeedeh Gholamrezazadeh ,Mohsen Amini Salehi, “A Comprehensive Survey on Text Summarization Systems ”, 978-1-4244-4946-0,2009 IEEE.
10. Rafael Ferreira ,Luciano de Souza Cabral ,Rafael Dueire Lins ,Gabriel Pereira e Silva ,Fred Freitas ,George D.C. Cavalcanti ,Luciano Favaro , "Assessing sentence scoring techniques for extractive text summarization ",Expert Systems with Applications 40 (2013) 5755-5764 ,2013 Elsevier .
11. L. Suanmali , N. Salim and M.S. Binwahlan,"Fuzzy Logic Based Method for Improving Text Summarization" , International Journal of Computer Science and Information Security, 2009, Vol. 2, No. 1,pp. 4-10.

12. Mrs.A.R.Kulkarni , Dr.Mrs.S.S.Apte "A DOMAIN-SPECIFIC AUTOMATIC TEXT SUMMARIZATION USING FUZZY LOGIC ",International Journal of Computer Engineering and Technology (IJCET), ISSN 0976- 6367(Print), ISSN 0976 - 6375(Online) Volume 4, Issue 4, July-August (2013).

13. Farshad Kyoomarsi ,Hamid Khosravi ,Esfandiar Eslami ,Pooya Khosravyan Dehkordy; "Optimizing Text Summarization Based on Fuzzy Logic ",Seventh IEEE/ACIS International Conference on Computer and Information Science ,978-0-7695-3131-1 ,2008

14. Ladda Suanmali ,Naomie Salim and Mohammed Salem Binwahla ,"Feature-Based Sentence Extraction Using Fuzzy Inference rules ",2009 International Conference on Signal Processing Systems ,978-0-7695-3654-5 ,2009 IEEE .

15. Ladda Suanmali, Naomie Salim and Mohammed Salem Binwahlan "Fuzzy Genetic Semantic Based Text Summarization ", 2011 Ninth Ninth International Conference on Dependable, Autonomic and Secure Computing ,978-0-7695-4612-4 ,2011 IEEE .

16. Ladda Suanmali, Mohammed Salem Binwahlan and Naomie Salim "Sentence Features Fusion for Text Summarization Using Fuzzy Logic ",2009 Ninth International Conference on Hybrid Intelligent Systems ,978-0-7695-3745-0 ,2009 IEEE.

17. Hsun-Hui Huang ,Yau-Hwang Kuo ,Horng-Chang Yang ,"Fuzzy-Rough Set Aided Sentence Extraction Summarization",Proceedings of the First International Conference on Innovative Computing, Information and Control (ICICIC’06),0-7695-2616-0/06 ,IEEE.

18. Landauer, T.K.—Dumais, S.T.: A Solution to Plato´ıs Problem: the Latent Semantic Analysis Theory of the Acquisition, Induction, and Representation of Knowledge. In Psychological Review, Vol. 104, 1997, pp. 211–240.

19. http://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-802.pdf

20. Asef, Kahani, Yazdi and Kamyar, “Context-Based Persian Multi-document Summarization”, In International Conference on Asian Language Processing (IALP), pp. 145-149, 2011.

21. Ozsoy, Cicekli, “Text Summarization of Turkish Texts using Latent Semantic Analysis", In Proceedings of the 23rd International Conference on Computational Linguistics, Beijing, pp. 869-876, 2010.

22. Makbule Gulcin Ozsoy and Ferda Nur Alpaslan, Ilyas Cicekli ,” Text summarization using Latent Semantic Analysis” Journal of Information Science 2616-0/06 ,IEEE.

23. Feifan Liu and Yang Liu, Member, IEEE "Exploring Correlation Between ROUGE and Human Evaluation on Meeting Summaries ",IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 18, NO. 1, JANUARY 2010

24. Josef Steinberger, Karel Jeˇzek ”EVALUATION MEASURES FOR TEXT SUMMARIZATION”, Computing and Informatics, Vol. 28, 2009, 1001–1026, V 2009-Mar-2

25. Md. Majharul Haque, Suraiya Pervin, and Zerina Begum” Literature Review of Automatic Single Document Text Summarization Using NLP”, ISSN 2028-9324 Vol. 3 No. 3 July 2013, pp. 857-865

AUTHOR(S) PROFILE

Ms. Patil Pallavi Dnyandeo, received the B.E degree in Information Technology and from Daulatrao Aher college of Engineering Karad (Banawdi) at Shivaji University .Now with Zeal Education’s Dnyanganga college of Engineering & Research for M.E.(Computer Engineering) purpose.

Mr. Mane Prashant M. (Assistant Professor) at Pune University , ZEAL education’s “Dnyanganaga College of Engineering And Research ”, Narhe, Pune. Qualification: M.E. (IT), Teaching Experience: 6 Years