



**IJIRCCCE**

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

**Volume 10, Issue 4, April 2022**

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 8.165**



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

# Implementation Paper “Two-Layer LSTM Deep Learning Network for Abstractive Summarization”

Prof. S.A. Deshpande, Pranav Jadhav, Sanjeev Kadam, Deep Kasat, Anuja Mahajan

Department of Information Technology, Pune Vidyarthi Griha's College of Engineering and Technology, Pune, India

**ABSTRACT:** Automatic text summarization is an essential natural language processing application that goals to summarize a given textual content into a shorter model. The fast growth in media information transmission over the Internet demands text summarization using neural network from asynchronous combination of text. This paper represents a framework that utilizes the techniques of NLP technique to examine the elaborative information contained in multi-modal statistics and to enhance the aspects of text summarization. The basic concept is to bridge the semantic gaps among multi-modal content. After, the generated summary for important information through multi-modal topic modeling. Finally, all the multi-modal factors are considered to generate a textual summary by maximizing the importance, non-redundancy, credibility and scope through the allocated accumulation of sub modular features. The experimental result shows that Text Summarization framework outperforms other competitive techniques.

**KEYWORDS:** Recurrent Neural Network, Text Summarization, NLP, Latent semantic analysis (LSA), Text Rank,

## I. INTRODUCTION

TEXT summarization plays a vital role in our daily life and has been studied for several decades. From information retrieval to text mining, we are frequently exposed to text summarization. The goal of automatic text summarization is presenting the source text into a shorter version with semantics. The most important advantage of using a summary is, it reduces the reading time. Text Summarization methods can be classified into extractive and abstractive summarization. In simple words, data mining is defined as a process used to extract usable data from a larger set of any raw data. It implies analyzing data patterns in large batches of data using one or more software. Data mining is also known as Knowledge Discovery in Data. Summarization is a key data mining concept which involves techniques for finding a compact description of a dataset. Simple summarization methods such as tabulating the mean and standard deviations are often applied for exploratory data analysis, data visualization and automated report generation. Text summarizers automatically construct summaries of a natural language document. Text summarization is the condensed form of any type of document whether pdf, doc, or txt files but this condensed form should preserve complete information and meaningful text with the help of single input file and multiple input file. It is not an easy task for human being to maintain the summary of large number of documents. The abstractive text summarization algorithms create new phrases and sentences that relay the most useful information from the original text — just like humans do. Therefore, abstraction performs better than extraction. Classical methods operate by selecting and compressing content from the source document. Search Results Featured snippet from the web Formally, NLP is a specialized field of computer science and artificial intelligence with roots in computational linguistics. It is primarily concerned with designing and building applications and systems that enable interaction between machines and natural languages that have been evolved for use by humans. LexRank is an unsupervised approach to text summarization based on graph-based centrality scoring of sentences. TextRank is an algorithm based on PageRank, which often used in keyword extraction and text summarization. Latent semantic analysis (LSA) is a technique in natural language processing, in particular distributional semantics, of analyzing relationships between a set of documents and the terms they contain by producing a set of concepts related to the documents and terms. The main feature of fuzzy logic in text summarization is that it is able to deal with impreciseness and uncertainty of features in deciding the importance of sentences to be extracted into the summary.

## II. LITERATURE SURVEY

- 1. Cheng, Jianpeng, and Mirella Lapata. "Neural summarization by extracting sentences and words."2016.**

This paper, perform an extractive text summarization on single page document is demonstrated. The Method extraction of sentences, which gives the idea of the input text in a short form, is presented. Sentences are ranked by assigning weights and they are ranked based on their weights. Develop a general framework for single-document summarization composed of a hierarchical document encoder and an attention-based extractor. This architecture allows us to develop different classes of summarization models which can extract sentences or words. We train our models on large scale corpora containing hundreds of thousands of document-summary pairs. Data-driven neural summarization models require a large training corpus of documents with labels indicating which sentences (or words) should be in the summary.
- 2. Nallapati, Ramesh, et al. "Abstractive text summarization using sequenceto- sequence rnns and beyond."(2016)**

In this paper abstractive text summarization is the task of generating a headline or a short summary consisting of a few sentences that captures the salient ideas of an article or a passage. We use the adjective 'abstractive' to denote a summary that is not a mere selection of a few existing passages or sentences extracted from the source, but a compressed paraphrasing of the main contents of the document, potentially using vocabulary unseen in the source document. Use fuzzy text summarization. Content keyword feature Similarity of sentence and title of document Sentence location in the document. This paper implement fuzzy text summarization technique using content keyword feature and Similarity of sentence and title of document.
- 3. Freitas, N., and A. Kaestner. "Automatic text summarization using a machine learning approach."2005**

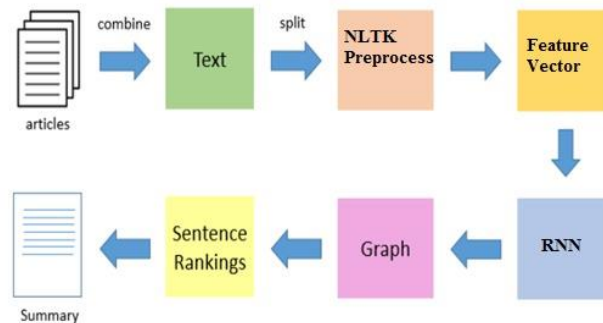
The proposed approach provides automatic feature based extractive heading wise text summarizer to improve the coherence thereby improving the Understandability of the summary text. It summarizes the given input document using local scoring and local ranking that is it provides heading wise summary. Generate summary using heading and sentence scoring. In this paper we address the automatic summarization task. Recent research works on extractive-summary generation employ some heuristics, but few works indicate how to select the relevant features. We will present a summarization procedure based on the application of trainable Machine Learning algorithms which employs a set of features extracted directly from the original text. These features are of two kinds: statistical – based on the frequency of some elements in the text; and linguistic – extracted from a simplified argumentative structure of the text. We also present some computational results obtained with the application of our summarizer to some well-known text databases, and we compare these results to some baseline summarization procedures.
- 4. Ferreira, Rafael, et al. "Assessing sentence scoring techniques for extractive text summarization." 2013**

Text summarization is the process of automatically creating a shorter version of one or more text documents. It is an important way of finding relevant information in large text libraries or in the Internet. Essentially, text summarization techniques are classified as Extractive and Abstractive. Extractive techniques perform text summarization by selecting sentences of documents according to some criteria. Abstractive summaries attempt to improve the coherence among sentences by eliminating redundancies and clarifying the contest of sentences. In terms of extractive summarization, sentence scoring is the technique most used for extractive text summarization. This paper describes and performs a quantitative and qualitative assessment of 15 algorithms for sentence scoring available in the literature. Three different datasets (News, Blogs and Article contexts) were evaluated. In addition, directions to improve the sentence extraction results obtained are suggested.
- 5. Chirantana Mallick, AjitKumar Das, Madhurima Dutta, Asit Kumar Dasand Apurba Sarkar, "Graph-Based Text Summarization Using Modified Text Rank", 2019**

In this paper the proposed method constructs a graph with sentences as the nodes and similarity between two sentences as the weight of the edge between them. Modified inverse sentence frequency-cosine similarity is used to give different weightage to different words in the sentence, whereas traditional cosine similarity treats the words equally. The graph is made sparse and partitioned into different clusters with the assumption that the sentences within a cluster are similar to each other and sentences of different cluster represent their dissimilarity. The performance evaluation of proposed summarization technique shows the effectiveness of the method.



### III. PROPOSED SYSTEM APPROACH



**Fig.2 Block Diagram of Proposed System**

We propose NLTK tools to preprocess the original Text text, Extract important features from Text text. Use algorithm to assign appropriate weights to the extracted Text text features and finally summarized it. System and users are two modules include in our system. System first gather the information about Text papers. After gathering of information like preprocessing on the data, feature extraction using genetic algorithm, training of that data, model generation according to the features of the data. User inserts the Text paper and finally gets the summarized of it. Using our proposed system, we summarized the articles at different type of stages. We are trying to develop system for summarization getting accurately and saves lots of retrieval of significant information from a long text within a short period.

- This system which will help in several areas such as Text articles, emails, research papers and online search engines to receive summary of results found.
- Our work is based on document with neural network techniques for summarize the Text.
- Module:
  - 1) Text Summarization.
- **Pre-processing**

In this part, preprocessing process contains four parts: (1) Segmentation by NLTK [40], (2) Removing stop words by NLTK, (3) part-of-speech tokenization, (4) Stem extraction based on NLTK and porter algorithm.

- **Feature Extraction and Summarization**

For extracting and analyzing information to retrieve useful knowledge within a reasonable time period, this information must be summarized. This paper reviews recent approaches for abstractive text summarization using deep learning models.

### IV. ADVANTAGES AND DISADVANTAGES

#### Advantages:

- Easy and rapid loading of the most important information, and resolution of the problems associated with the criteria needed for summary evaluation.
- To give unique solution for getting the summarization of long text, article.
- Even though the quality of summary generated might be excellent, manual summarization is a time consuming process



- It can solve the problem of information overload on the Internet and the other hand, it can simplify the information obtained by users.

**Disadvantages:**

- Large number of training dataset is required for more accuracy.

**V. RESULT**

As per the survey mentioned, we’ve devolved the system of abstractive summarisation of the text. Result of the same are below:

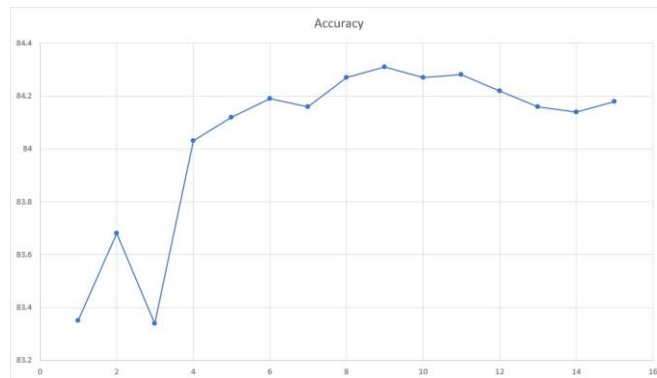


Figure: “Give Graph Name here”

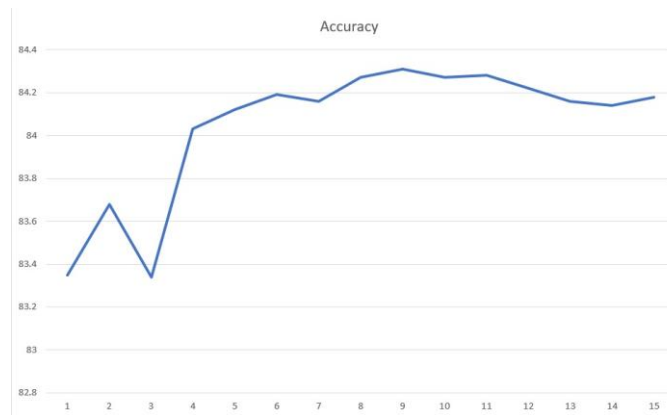


Figure: “Give Graph Name here”

	A	B	C	D	E	F	G	H	I	J	K	L
1	Paragraph	Review	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
2	The Best Hot Sauce in the World I don't know if it's the cactus or the	best hot sauce	7	8	7.5	7	8	8	8.5	7.5	8	8.5
3	If you like tangerine and you think you need carbonation, then I	refreshing	8	7	7.5	7	7.5	8	8	7	7.5	8
4	I'll admit that I don't care for drinks that are carbonated - the	refreshing	6	5	7	7.5	7	8	7.5	7	8	8.5
5	Ever since we discovered these several years ago, they have been a	best dark	9	8	8.5	8	9	7.5	8	8.5	8	9
6	Being an American in a small town on the West Coast, I had never	great taste	7	6	7.5	8	8.5	7	7.5	8	7.5	8
7	Sometimes with an active lifestyle comes bad joints. I have arthritis	best thing ever	8	8	8.5	7	8.5	7	6.5	7.5	8	7.5
8	Here is the skinny on using this as camping food: get some reflectix	best food ever	7	6	7	6.5	7	7.5	8	8.5	8	7.5
9	I give props to any mom, especially one that holds down a full time	Best baby food	8	7	7.5	8	8.5	9	8	9	8.5	8
10	Arrived in 6 days and were so stale i could not eat any of the 6 bags!!	stale	7	8	7	6.5	7	8	8.5	8	9	8
11	very bad product worst quality hate it	horrible	6	8	7	7.5	8	8.5	8	7.5	8	7.5

Figure: “Give Graph Name here”

## VI. CONCLUSION

This system addresses an asynchronous task, namely, how to use related text information to generate a textual summary. We formulate the task as an optimization problem with a budgeted maximization of sub modular functions. Automatic text summarization is a complex task which contains many sub-tasks in it. Every subtask has an ability to get good quality summaries. The important part in extractive text summarization is identifying necessary paragraphs from the given document. In this work we proposed extractive based text summarization by using statistical novel approach based on the sentences ranking the sentences are selected by the LexRank. The sentences which are extracted are produced as a summarized text. The proposed model improves the accuracy when compared traditional approach.

## VII. FUTURE WORK

In future work we are going to focus on more detailed study and developed android apps for Text summarization.

## ACKNOWLEDGMENT

This work is supported in Lung Cancer Detection System for medical research and development of any state in India. Authors are thankful to Faculty of Engineering and Technology (FET), Savitribai Phule Pune University, Pune for providing the facility to carry out the research work.

## REFERENCES

- [1] Cheng, Jianpeng, and Mirella Lapata. "Neural summarization by extracting sentences and words." arXiv preprint arXiv:1603.07252 (2016).
- [2] Nallapati, Ramesh, et al. "Abstractive text summarization using sequenceto- sequence rnns and beyond." arXiv preprint arXiv:1602.06023 (2016).
- [3] Freitas, N., and A. Kaestner. "Automatic text summarization using a machine learning approach." Brazilian Symposium on Artificial Intelligence (SBIA), Brazil. 2005.
- [4] Ferreira, Rafael, et al. "Assessing sentence scoring techniques for extractive text summarization." Expert systems with applications 40.14 (2013): 5755-5764.
- [5] Chirantana Mallick, AjitKumar Das, Madhurima Dutta, Asit Kumar Dasand Apurba Sarkar, "Graph-Based Text Summarization Using Modified Text Rank", Springer Nature Singapore Pte Ltd. 2019



INNO  SPACE  
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

 **doi**<sup>®</sup>  
**cross** **ref**

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  [ijircce@gmail.com](mailto:ijircce@gmail.com)



[www.ijircce.com](http://www.ijircce.com)

Scan to save the contact details