IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Abstractive Summarizer – A meaningful gist

Swapnil Sagar,
Department of CSE,
MIT School of Engineering,
MIT ADT University,
Pune, India.

Monti Patel,
Department of CSE,
MIT School of Engineering,
MIT ADT University,
Pune, India.

Swapnil Jagtap,
Department of CSE,
MIT School of Engineering,
MIT ADT University,
Pune, India.

Umakant Kumar Yadav,
Department of CSE,
MIT School of Engineering,
MIT ADT University,
Pune, India.

Prof. Dipti.D.Pawar,
Assistance Professor
Department of CSE,
MIT School of Engineering
MITADT University
Pune, India

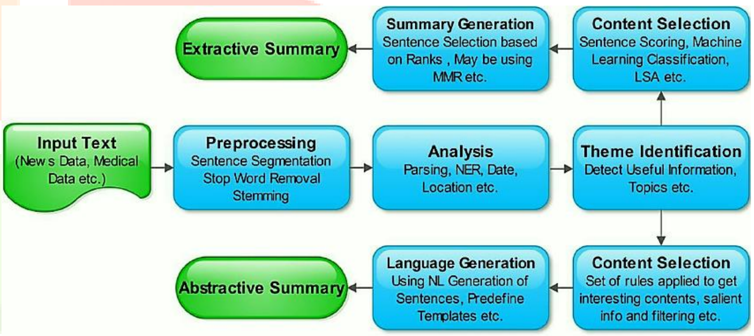
***Abstract*— Abstractive method of Text Summarization means generating a smaller and readable summary texts that captures the context of the source text in a lesser readable text form. The produced summaries contain new expressions and texts that might not be in the original paragraphs. Summaries play a vital role for readers who regularly reads and search for documents. We have created a model using Deep Learning and LSTM which generates a short abstractive summary. This process requires a lot of theoretical implementations and challenges which need to be overcome which a discussed in details in this paper.**

***Keywords*— Abstractive summarization, Natural Language Processing, Beautiful Soup, Deep learning, LSTM**

I. Introduction

TLDR is a common initialism used frequently by readers. Its full version means “Too Long; Didn't Read”. Most of the times, people state TLDR to say that the paragraph is too long to read which makes don’t want to read it. A short summary of the text is the most effective approach to understand an overall theme of a document. Summarization intends to diminish the size of the target phase while keeping its true meaning intact. It is most helpful in long formatted texts which required longer read time. Based on types of summarizations, they are made using different techniques. The first one takes most important sentence and makes one summary out of those sentences which is known as Extractive summarization the other one uses new words through **Natural Language Processing** technique to connect important phrases which is known as Abstractive summarization.

Extractive summarization technique is an extensively researched topic and has been developed for quite a mature level. Now the research has more focus towards the **abstractive summarization**. What makes abstractive summarization a time consuming and complicated task is the dynamic and elaborated algorithms in the natural language text which help to develop a newly created short sentences which has the same meaning as the original form.



This **Abstractive Summarizer** is considered as one of the most complex undertaking in the fields of Deep Learning and Natural Language Processing, which includes a proper knowledge of long sentences, information compression, and logical language generation [3]. The dominant paradigm for training machine learning models to do this is sequence-to-sequence (seq2seq) learning using (Long Short-Term Memory) **LSTM**, where a neural network learns to map input sequences to output sequences. While these seq2seq models were initially developed using recurrent neural networks, Transformer encoder-decoder models have recently become favored as they are more effective at modeling the dependencies present in the long sequences encountered in summarization [2].

II. Literature Survey

In 2016, N. Moratanch and Dr. S. Chitrakala have proposed a paper "A Survey on Abstractive Text Summarization". In this paper, the authors proposed an idea which uses semantic based approach to generate abstractive summary.

In 2017, Abigail See, Peter J. Liu and Christopher D. Manning have proposed a paper "Get to The Point: Summarization with Pointer-Generator Networks". In this document, authors have used hybrid pointer-generator network which inputs words via pointing and coverage to keep track of what has been summarized.

In 2018, Nithin Rapha, Hemanta Duwarah and Philemon Daniel have proposed a paper "Survey on Abstractive Text Summarization".In this paper, they have compared the different approaches(RNN variance, Word embedding,etc) to achieve the abstractive summary .

In 2019, Abu Kaisar Mohammad Masum, Sheikh Abujar, Md Ashraful Islam Talukder, Syed Akhter Hossain and AKM Shahariar Azab Rabby have proposed a paper "Abstractive method of text summarization with sequence-to-sequence RNNs". In this paper, they have used RNN Encoder and Decoder and LSTM to summarize the text.

In 2020, Pooja Batra, Kavya Bhatt, Sarika Chaudhary, Saloni Varshney and Srashti Verma have proposed a paper "A Review:Abstractive Text Summarization Techniques using NLP".In this paper,they have used LSTM and Encoder-Decoder networks with attention mechanism.

Comparison of various approach:

Year	Author	Methods	Description
2016 [9]	N. Moratanch and Dr. S. Chitrakala	Semantic based approach	In sematic based technique,linguistics illustration of documents is employed to feed into NLG system.
2017 [10]	Abigail See,Peter J. Liu and Christopher D. Manning	Pointer-Generation Networks	Pointing aids accurate reproduction of information while retaining the ability to produce novel words.
2018 [11]	Nithin Raphal, Hemanta Duwarah and Philemon Daniel	RNN Variance, Work Embedding	They have reviewed supervised learning and GRU(Gated Recurrent Unit) to compare the summary of text.
2019 [12]	Abu Kaisar Mohammad Masum, Sheikh Abujar, Md Ashraful Islam Talukder, Syed Akhter Hossain and AKM Shahariar Azab Rabby	Sequence to Sequence RNNs	They preprocess the data count, vocabulary size , word embedding file and build a sequence to sequence model and response summary.
2020 [13]	Pooja Batra, Kavya Bhatt ,Sarika Chaudhary, Saloni Varshney and Srashti Verma	LSTM and Encoder-Decoder architecture	They have used LSTM and RNN to generate the tetxt summary.

III. SYSTEM ARCHITECTURE AND IMPLEMENTATION

Abstractive Summarization

In this type of summarization, new sentences are being generated with the help of original text. The new sentences which are formed might not be present in the original text.
Introduction to Sequence-to-Sequence (Seq2Seq) Modelling
Our goal is to design a text summarizer where the input is a long sequence of words and the output is a short summary. Below is a typical Seq2Seq model architecture:
There are two major components of a Seq2Seq model:

- Encoder
- Decoder

Encoder-Decoder Architecture

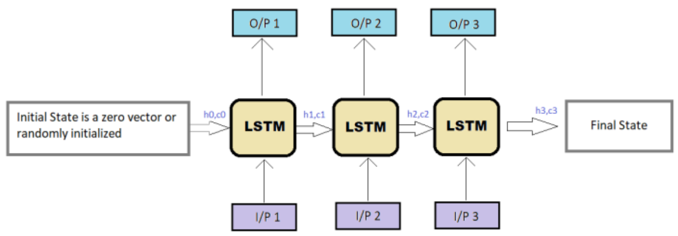
This architecture is used to solve the problems where the input and outputs are of different length.
The input is a long sequence of words and the output will be a short edition of the input sequence.
We have set up the Encoder-Decoder in 2 phases:

- Training phase
- Inference phase

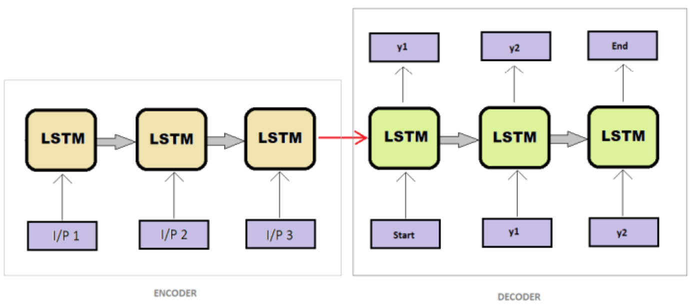
Let's understand these concepts through the lens of an LSTM model.

Training phase
In the training phase, we had set up the encoder and decoder. We have trained the model to predict the target sequence offset by one timestep.

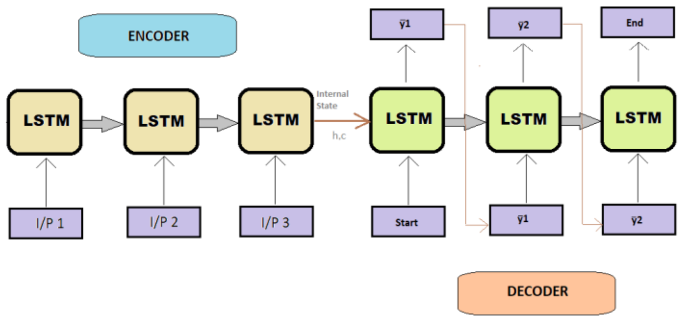
Encoder
An Encoder Long Short-Term Memory model (LSTM) reads the entire input sequence and at every timestep, one word is fed into the encoder [8], then it processes the information at every timestep and records the contextual information present in the input sequence.



Decoder
The decoder is an LSTM network which reads the entire target sequence word-by-word and forecast the same sequence offset by one timestep. [8] The decoder is prepared to forecast the next word in the sequence given the previous word.



Inference Phase
After training, the model is tested on new source sequences for which the target sequence is unknown [12].
The Intuition behind the Attention Mechanism



We have taken an example to understand how Attention Mechanism works:

- **Source sequence:** "Which sport do you like the most?
- **Target sequence:** "I love cricket"

The first word **'I'** in the target sequence is connected to the fourth word **'you'** in the source sequence, right? Similarly, the second-word **'love'** in the target sequence is associated with the fifth word **'like'** in the source sequence.
So, instead of seeing all the words in the source sequence, we can expand the significance of specific parts of the source sequence that result in the target sequence. This is the basic idea behind the attention mechanism.
There are 2 classes of attention mechanism depending on the way the attended context vector is derived:

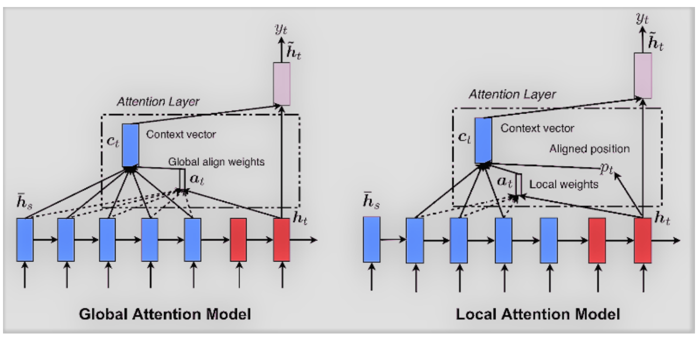
- Global Attention
- Local Attention

.

Global Attention
Here, the attention is placed on all the source positions. In other words [6], **all the hidden states of the encoder are considered for deriving the attended context vector.**

Source: Effective Approaches to Attention-based Neural Machine Translation – 2015

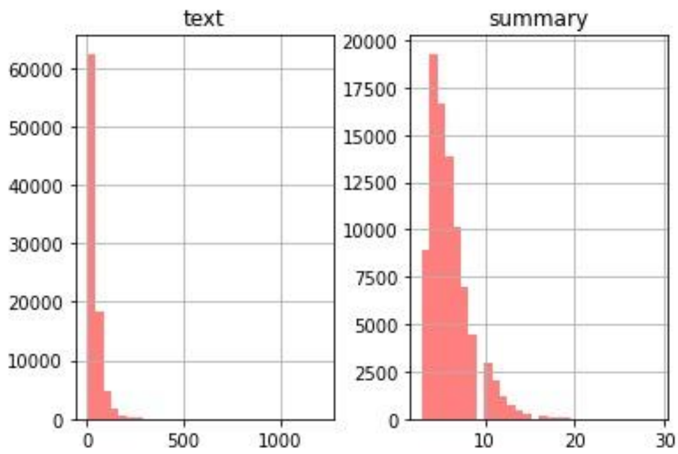
Local Attention
Here, the attention is placed on only a few source positions. **Only a few hidden states of the encoder are considered for deriving the attended context vector.**



IV. SYSTEM RESULT

In this model, we analyze the total length of words, reviews and the summary to get an overall stat about the distribution of total length of the provided text. This will help us to set a fixed maximum length of the sequence.
Here we use Python pyplot library to see the comparison graph of the summary and the text generated by the model.

Comparison Output:



Here is a summary generated by the model:

```
Review: used eating flaxseed brownie hodgson mill
brownies super easy wake taste great since like
dark chocolate usually add little cocoa
Original summary: delicious brownie
Predicted summary: best brownie mix
Review: favorite coffee keurig coffeemaker
convenient get amazon cheaper running around
stores trying find lowest price
Original summary: great coffee
Predicted summary: great coffee
```

The fact which makes it Abstractive is that the summary generated by the model does not have common words in the input passage but still both of them convey the same meaning.

V. CONCLUSION AND FUTURE WORK

In this paper we have discussed the implementation of an effective solution for Text Summarization using Deep learning and LSTM.
The future scope of this project is to increase the training dataset size and build the model. By using larger dataset, the capability of this deep learning model will enhance. Using beam search strategy for decoding the test rather than simple greedy approach (argmax) will result on more logical filtering of words and give better BLEU score.
Bi-Directional LSTM can also be implemented which is capable of capturing the context from both the directions and gives a better context vector as result.

VI. REFERENCES

[1] Gupta, Som & Gupta, S.K. (2018). Abstractive Summarization: An Overview of the State of the Art. Expert Systems with Applications. 121. 10.1016/j.eswa.2018.12.011.

[2] Sutskever, Ilya & Vinyals, Oriol & Le, Quoc. (2014). Sequence to Sequence Learning with Neural Networks. Advances in Neural Information Processing Systems. 4.

[3] Knight, K. and Marcu, D. 2000. Statistics-based summarization - step one: sentence compression. In Proc. Of AAAI 2000.

[4] Lloret, E., Palomar, M. 2011. Analyzing the Use of Word Graphs for Abstractive Text Summarization. In Proc. of IMMM 2011.

[5] S. Syed, Abstractive Summarization of Social Media Posts: A case Study using Deep Learning, Master's thesis, Bauhaus University, Weimar, Germany, 2017.

[6] A. Khan and N. Salim, "A review on abstractive summarization methods," Journal of @eoretical and Applied Information Technology, vol. 59, no. 1, pp. 64–72, 2014

[7] E. Baralis, L. Cagliero, A. Fiori, P. Garza, "MWI-Sum: A Multilingual Summarizer Based on Frequent Weighted Itemsets", ACM Transactions on Information Systems, vol. 34, no. 1, pp. 1-35, 2015

[8] Meena, Yogesh & Gopalani, Dinesh. (2015). Domain Independent Framework for Automatic Text Summarization. Procedia Computer Science. 48. 723-728. 10.1016/j.procs.2015.04.207.

[9] Moratanch, N. & Gopalan, Chitrakala. (2016). A survey on abstractive text summarization. 1-7. 10.1109/ICCPCT.2016.7530193.

[10] Abigail & Liu, Peter & Manning, Christopher. (2017). Get To the Point: Summarization with Pointer-Generator Networks. 1073-1083. 10.18653/v1/P17-1099.

[11] Raphal, Nithin & Duwarah, Hemanta & Daniel, Philemon. (2018). Survey on Abstractive Text Summarization. 0513-0517. 10.1109/ICCSP.2018.8524532.

[12] A. K. Mohammad Masum, S. Abujar, M. A. Islam Talukder, A. K. M. S. Azad Rabby and S. A. Hossain, "Abstractive method of text summarization with sequence to sequence RNNs," *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2019, pp. 1-5, doi: 10.1109/ICCCNT45670.2019.8944620.

[13] P. Batra, S. Chaudhary, K. Bhatt, S. Varshney and S. Verma, "A Review: Abstractive Text Summarization Techniques using NLP," 2020 International Conference on Advances in Computing, Communication & Materials (ICACCM), 2020, pp. 23-28, doi: 10.1109/ICACCM50413.2020.9213079.

[14] T. Shi, Y. Keneshloo, N. Ramakrishnan, and C. K. Reddy, Neural Abstractive Text Summarization with Sequence-To-Sequence Models: A Survey, http://arxiv.org/abs/1812.02303, 2020

[15] C. Sunitha, A. Jaya, and A. Ganesh, "A study on abstractive summarization techniques in Indian languages," Procedia Computer Science, vol. 87, pp. 25–31, 2016.