Analysis and Implemented on Automated Text Summarization using Transformer Model



Shruti J. Sapra (Thakur), Avinash S. Kapse, Mohammad Atique

Abstract: Even though the work on automatic text summarization initially began 70 years prior, it has seen a remarkable development in the recent years due to new and advanced technologies. With the increasing significance of time, the need for condensed and precise information is at its peak. No one has time to review all the articles to obtain the correct data. With the help of an automatic text summarizer, we can condense the source text while preserving its content and overall meaning, thereby saving the reader's time. Text summarisation can be broadly categorised into two types: Abstractive Summarisation and Extractive Summarisation. Extractive summarisation aims to distinguish the most vital information that is currently separated and assembled in a system into a condensed summary. The abstractive summary group includes rewriting the entire article, and the summary is created using natural language processing techniques. In this paper, we discuss various text summarisation models and present the results of our implementation of an automatic text summarizer, which was trained using the CNN Daily Mail dataset.

OPEN ACCESS

Keywords: Abstractive Summarization, Extractive Summarization, Natural Language Processing, Text Summarization, Deep Learning, Transformers

I. INTRODUCTION

In the modern world, a large amount of data is readily

available through the internet. No one has time to review all the articles to obtain the correct data. Thus, the need for precise information has become a necessity, given the popular minimum amount of time. Although the effort on Automatic Text Summarisation (ATS) began over 70 years ago, it has experienced exponential growth in recent years due to the introduction of new and advanced technologies. With the help of an automatic text summarizer, we can condense the source text while preserving its information and overall meaning, thereby saving the reader's time. The reader can then judge whether to read the complete document or not based on the summary.

Manuscript received on 24 March 2024 | Revised Manuscript received on 03 April 2023 | Manuscript Accepted on 15 April 2024 | Manuscript published on 30 April 2024. *Correspondence Author(s)

Dr. Shruti J. Sapra (Thakur)*, Research Scholar, Department of Computer Science and Engineering, Sant Gadge Baba Amravati University, Amravati (M.H), India. E-mail: sjssapra123@gmail.com, ORCID ID: 0000-0002-0619-8500

Dr. Avinash S. Kapse, Associate Professor and Head, Department of Information Technology, Anuradha College of Engineering, Chikhli, Sant Gadge Baba Amravati University, Amravati (M.H), India. E-mail: askapse@gmail.com, ORCID ID: 0000-003-4637-6006

Dr. Mohammad Atique, Professor and Head, Department of Computer Science and Engineering, Sant Gadge Baba Amravati University, Amravati (M.H), India. E-mail: <u>mohammadatique@sgbau.ac.in</u>, ORCID ID: <u>0000-0001-6976-9895</u>

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an <u>open access</u> article under the CC-BY-NC-ND license <u>http://creativecommons.org/licenses/by-nc-nd/4.0/</u>

In Natural Language Processing, text summaries can be broadly categorised into two types: Extractive Summarisation and Abstractive Summarisation. Extractive Summarisation methods depend on mining a few portions, such as sentences or phrases, from a portion of text and then combining them to form a summary. Consequently, selecting the exact sentences is of extreme importance in an extractive method. On the other hand,

Abstractive Summarisation strategies utilise advanced NLP procedures to create an entirely new list, wherein some parts of this list may not appear in the original text. It incorporates heuristic methods to train the framework, endeavouring to comprehend the entire setting and create a list based on that arrangement. This is a more regular method of producing rundowns, and these outlines are more beneficial when compared to the extractive methodologies. However, regardless of the technique, the correctness of the summarisation is problematic to promise. To increase the precision of summaries, various models have incorporated additional factors into their implementations. Some have utilised human aid; some have employed a hybrid of extractive and abstractive summarisation systems, while others have focused on headings to make their summaries more favourable. In this paper, we discuss various text summarisation models and the results of our implementation of an automatic text summarizer, which we trained using the CNN Daily-Mail dataset and is based on the Transformer model. This deep learning framework attempts to address sequential processes while handling long-range persistent data with minimal difficulty. And finally, we draw conclusions and future directions.

II. LITERATURE REVIEW

Ekaterina Zolotareva et al [1] In this work, the text rundown issue has been investigated utilizing Sequence-tosuccession repetitive neural organizations and Transfer Learning. Their experimental results showed that the Transfer Learning-created model achieved significant development for abstractive text summarisation.

Wen Kai et al [2]. This paper proposes a prepreparation technique that involves a dependent arrangement of Bidirectional Encoder Representations from Transformers (BERT) and then combines it with a LeakGAN model to generate summaries.

Elozino Egonmwan et al [3]. They propose a framework that further develops execution on single-record outline tasks utilising the CNN/DailyMail and News studio datasets.

Published By: Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP) © Copyright: All rights reserved.



It tracks the well-known encoder and decoder worldview, yet with an additional emphasis proceeding from the encoder.

P.Krishnaveni *et al* [4]. The planned approach delivers an automatic extractive heading and summary text summarizer to enhance the quality of the text, thereby improving the comprehensibility of the summary manuscript. It summarises the specified information report, utilising nearby scoring and neighbourhood positioning, which provides a synopsis.

Shuxia Ren et al [5]. In this paper, the authors have combined the advantages of extractive and abstractive outline frameworks to propose a text synopsis prototype for a global fenced unit and duplicate component (GGUC). The analysis of the consequences exhibits that the presentation of the model is superior to the following message outline framework on the LCSTS datasets.

Chandra Prakash et al [6]. In this paper, the humansupported manuscript summarizer "SAAR" is currently planned for single records. As per the report, a term-sentence grid is created. The passages that trend the grid are weighted according to Reinforcement Learning. Subsequently, the formed outline is displayed to the client. If the client supports it, the final list is produced; otherwise, a new synopsis is created based on the client's input in the form of keywords.

Ashish Vaswani et al [7]. They propose another basic organisation design, the Transformer, based on consideration instruments, which forgoes repeats and convolutions entirely. Probes into two machine interpretation undertakings demonstrate these models to be predominant in quality and require altogether less time to train.

Ayesha Ayub Syed et al [8]. This review will generally examine the logical writing to acquire data and information about the current research in involuntary text summarizers, specifically abstractive outline-dependent proceeding neural networks. An audit of different neural organisations has been conducted to introduce abstractive synopsis models.

Reeta Rani et al [9]. This survey paper outlines various past investigations and studies in the field of Automatic Text Summarisation.

Wang Guan et al [10]. This article provides a detailed review of various methods and evaluation patterns. The fundamental consideration is leveraging the use of the most recent patterns, including neural networks and previously trained transformer models.

III. METHODOLOGY

Rewriting the complete document. The essence of the Transformer model lies with the idea of "self-attention"[7]. the capability to address multiple locations in the input sequence to create a depiction of the output. Transformers create stacks of self-attention layers.



Fig. 1. Neural Attention Model

The encoder (left) and decoder (right) blocks are at least one of a similar encoder and decoder assembled. Together, the encoder stack and the decoder stack are to a similar extent. The quantity of these stacks is a hyperparameter [11].

- The word of the input sequence is passed to the first encoder.
- These are situated further, transmuted, and passed on to the followingencoder.
- Output given by the final encoder in the encoders stack is given to every decoder in the decoder pile to processfurther [12].

$$Attention(Q, K, V) = softmax\left(\frac{QK^{T}}{\sqrt{d_{k}}}\right)V$$

Fig. 2. Attention formula

- Q is a representation of the words in the input in matrix form.
- K is a representation of the vocabulary related to the words in the matrix form.
- V is again the representation of the words in the input, which is used for the dot product and getting practical inferences [13].

A. Theory

The transformer is a deep learning framework that attempts to address sequential processes while addressing long-range persistent data with minimal difficulty. It comes under abstractive summarization and gives a summary after

B. Steps / Algorithm

While implementing the text summarizer, we have done it in the following stages:

Published By: Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP) © Copyright: All rights reserved.



Retrieval Number: 100.1/ijeat.D440413040424 DOI: <u>10.35940/ijeat.D4404.13040424</u> Journal Website: <u>www.ijeat.org</u>



Step 1: Importing the Data and Pre-Processing the Input Data. In this step, we import the dataset from the source (CNN Daily) and perform the necessary pre-processing on the data. We also tokenise the input so that it can be processed as individual words instead of a single string of words. We implement the detokenize routine to detokenize the output.

Step 2: - Encoder-decoder block, Dot-product and Causal Attention, Transformer Decoder Block and Transformer Language Model.

In this step, we implement the Transformer Model with the encoder-decoder, Attention, and Feedforward subroutines.

Step 3: Training. In this step, we train the model using the summaries that have already been generated in the dataset.

Using the pre-determined hyperparameters, the model trains itself and improves accuracy.

Step 4: - Evaluation. In this step, the model's accuracy will be evaluated at various steps [14].

IV. RESULTS

We trained the model with six layers and eight heads as hyperparameters. We found that after 1000 steps, the accuracy. It increased to 0.1173 as compared to the accuracy of 0.04255 obtained at 10 steps. Further results are illustrated in Fig. 3. The progression of accuracy with the number of training steps is shown in Fig. 1.

Table 1 Results	Obtained W	hile Training	the Model
Table 1. Results	Obtained W	nne rranning	the Mouel

Heads	6	6	6	6	6	6
Layers	8	8	8	8	8	8
Steps	10	20	50	100	500	1000
Accuracy	0.04255	0.03876	0.03947	0.03446	0.09693	0.1173
Train Cross-Entropy Loss	10.29155	9.8105	7.93204	7.54951	7.09017	7.002833
Eval Cross Entropy Loss	10.09326	9.6462	8.32800	7.69350	7.61404	7.55096
Time Per Step	83.4	107.4	62.9	72.9	74.9	72.1

- Heads No. of Heads in the feed-forward neural network
- Layers No. Number of Layers in the feed-forward neural network
- Steps No. Number of steps in the loop while training the dataset
- Accuracy Accuracy obtained while testing through the CNN Daily Mail Dataset
- Train/Eval Cross Entropy Loss Cross-entropy is a commonly used loss function to optimize the model further
- Time pre-step Time required to run each step in the loop for the stipulated steps



Fig. 3. Accuracy on Y-axis and No. Number of Steps on X-axis

V. CONCLUSION

With the development of digital media and the steadily growing publishing industry, no one has time to sift through all the articles to obtain accurate data. Thus, there is a demand for an automated text summarizer that will condense the source text while maintaining its information and overall meaning. We implemented the transformer model for the task of summarisation. We concluded from our results that after 1000 steps, the accuracy increased to 0.1173, compared to the accuracy of 0.04255 obtained at 10 steps.

DECLARATION STATEMENT

Funding	No, I did not receive.
Conflicts of Interest	No conflicts of interest to the best of our knowledge.
Ethical Approval and Consent to Participate	No, the article does not require ethical approval or consent to participate, as it presents evidence.
Availability of Data and Materials/	Not relevant.
Authors Contributions	All authors have equal participation in this article.

REFERENCES

- Ekaterina Zolotareva, Tsegaye Misikir Tashu and Tomáš Horváth, "Abstractive Text Summarization using Transfer Learning", In:20th ITAT (Information technologies Applications and Theory), Conference At: Tyrapol, Volume: 2718, August 2020.
- Wen Kai, Zhou Lingyu, "Research on Text Summary Generation Based on Bidirectional Encoder Representation from Transformers", Published in: 2020 2nd International Conference on Information Technology and Computer Application (ITCA). https://doi.org/10.1109/ITCA52113.2020.00074
- Elozino Egonmwan and Yllias Chali, "Transformer-based Model for Single Documents Neural Summarisation", Proceedings of the 3rd Workshop on Neural Generation and Translation, Association for Computational Linguistics, November 2019. https://doi.org/10.18653/v1/D19-5607
- P.Krishnaveni, Dr.S. R. Balasundaram, "Automatic Text Summarization by Local Scoring and Ranking for Improving Coherence", IEEE 2017 International Conference on Computing Methodologies and Communication (ICCMC). <u>https://doi.org/10.1109/ICCMC.2017.8282539</u>
- 5. Shuxia Ren, Kaijie Guo, "Text Summarization Model of Combining Global Gated Unit and Copy Mechanism", 2019 IEEE 10th International Conference on Software Engineering and Service Science
- (ICSESS).
 6. Chandra Prakash, Dr. Anupam Shukla, Human Aided Text Summarizer "SAAR" using Reinforcement Learning, 2014 *Published By:*

Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP) © Copyright: All rights reserved.



Analysis and Implemented on Automated Text Summarization using Transformer Model

International Conference on Soft Computing & Machine Intelligence. https://doi.org/10.1109/ISCMI.2014.2

- 7. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin, "Attention Is All You Need", arXiv:1706.03762, 2017.
- 8. Ayesha Ayub Syed, Ford Lumban Gaol, Tokuro Matsuo, "A Survey of the State-of-the-Art Models in Neural Abstractive Text Summarization", Published in: IEEE Access (Volume: 9), 2021. https://doi.org/10.1109/ACCESS.2021.3052783
- 9. Reeta Rani and Sawal Tandon, "Literature Review On Automatic Text Summarization", International Journal of Current Advanced Research, Volume 7, Issue 2(C), February 2018.
- 10. Wang Guan, Ivan Smetannikov, Man Tianxing, "Surveyon Automatic Text Summarization and Transformer Models Applicability", CCRIS 2020: 2020 International Conference on Control, Robotics and Intelligent System, October 2020. https://doi.org/10.1145/3437802.3437832
- 11. Python Programming: Using Problem-Solving Approach, Reema Thareja, Oxford University Press, June 2017)
- 12. Natural Language Processing with Python: Analysing Text with the Natural Language Toolkit, Steven Bird, Ewan Klein, Edward Loper, January 2011.
- 13.Ms. Shruti Sapra (Thakur), Dr. Avinash S. Kapse, "Analysis of Effective Approaches for Legal Text Summarization Using Deep Learning", "International Journal of Scientific Research in Computer Science Engineering and Information Technology", http:// ijrcseit.com/paper/CSEIT21849, ISSN: 2456-3307, pp. 53-59.
- 14. Shruti Sapra (Thakur), Dr. Avinash S. Kapse, "Constructive Approach for Text Summarization Using Advanced Techniques of Deep Learning", 5 th International Conference on Intelligent data Communication and Technologies and Internet of Things (ICICI 2021), http://icoici.org/2021/978.981-16-7610-9 Series of Springer.

AUTHORS PROFILE



Dr. Shruti Sapra (Thakur) is a Research Scholar in the PG Department of Computer Science of Sant Gadge Baba Amravati University. She completed her Ph.D. on March 22, 2024, and is an Assistant Professor at G.H. Raisoni College of Engineering, Nagpur. She has over 12 years of experience in teaching and more than 3 years in research. Published

25+ research papers in national and international journals and conferences. I have published five patents, one chapter in a book in SCOPUS, and two copyrights.



Dr. Avinash S. Kapse is an Associate Professor & Head of Department Computer Science & Engineering, Anuradha Engineering College, Chikhli, MS, India and Research Supervisor of Sant Gadge Baba Amravati University, Amravati, in Computer Science & Engineering and Information Technology. He is a second-term Board of Studies Member (BoS

Member) in Information Technology at Sant Gadge Baba Amravati University, Amravati. He has over 18 years of experience in teaching and research. He has published over 53 research papers in national and international journals and conferences. He published six patents and seven copyrights.



Dr. Mohammad Atique is a Professor and Head in the PG Department of SGBAU, Amravati. He received BE Computer Science and Engineering from Government College of Engineering, Amravati, India, in 1990, ME Computer science and Engineering (with Distinction) from Sant Gadge Baba Amravati University, Amravati India, in 1998 and the PhD degree in Computer Science and Engineering

from the Government College of Engineering, Amravati, in 2009. Since October 2011, he has been working as a full professor and head in the P.G. Department of Computer Science and Engineering at Sant Gadge Baba Amravati University, Amravati.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of the Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)/ journal and/or the editor(s). The Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP) and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Retrieval Number: 100.1/ijeat.D440413040424 DOI: 10.35940/ijeat.D4404.13040424 Journal Website: www.ijeat.org

Published By: Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP) © Copyright: All rights reserved.