International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-13 Issue-4, April 2024

Analysis and Implemented on Automated Text Summarization using Transformer Model

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

***Abstract:** Despite the fact that 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 of condensed and precise information is on peak. No one has time to go through all the articles to get the right data. With the help of automatic text summarizer, we can shorten the source text while maintaining its data and overall meaning, thus saving time of the reader. Text summarization can extensively be alienated into two classifications, Abstractive Summarization also Extractive Summarization. Extractive summarization goals at distinguishing the foremost vital info that is at that moment separated and assembled system to a condensed summary. Abstractive summary group includes rewriting the complete article and the summary is created using natural language processing techniques. In this paper, we have discussed various text summarization models and presented theresults of our own implementation of automatic text summarizer which was trained using the CNN Daily mail dataset.*

***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 available at hands through the internet. No one has time to go through all the articles to get the right data. Thus, the need of precise information popular minimum amount of time has become a necessity. Even though the effort on Automatic Text Summarization (ATS) first on track 70 years ago, it has seen an exponential growth in the recent years due to new and advanced technologies. With the help of automatic text summarizer, we can shorten the source text while maintaining its information and overall meaning, thus saving time of the reader who can 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: 0000-0003-4637-6006

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

In Natural Language Processing, text summaries can be broadly alienated into two types, Extractive Summarization and Abstractive Summarization. Extractive Summarization methods depend on mining a few portions, such as sentences then phrases, from a portion of text and then stacking to form a summary [19]. Consequently, picking out the exact sentences is of extreme position in an extractive method. On the other hand,

Abstractive Summarization strategies utilize progressed NLP procedures to create an altogether new rundown wherein a few pieces of this outline may not show up in the first text. It incorporates heuristic ways to train the framework endeavoring to comprehend the entire setting and create a rundown dependent on that arrangement. This is a more regular method of producing rundowns and these outlines are more beneficial when contrasted with the extractive methodologies. But whatever the technique be, the correctness of the summarization is problematic to promise. To increase the precision of summaries, different models have included different factors in their implementations. Some have taken human aid; some have used a hybrid of extractive and abstractive summarization systems while some have focused on the headings to make their summaries more favorable. In this paper, we discuss various text summarization models and the results of our own implementation of an automatic text summarizer, which we trained utilizing the CNN Daily-Mail dataset, and is based on the transformer model, a deep learning framework that attempts to address sequential processes while addressing long-range persistent data with minimal difficulty. And finally, we draw conclusion and future directions.

II. LITERATURE REVIEW

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

Wen Kai et al [2][15]. *In* this paper suggests a pre-preparing technique dependent arranged Bidirectional Encoder Illustration after Transformers (BERT) then joined with LeakGAN model toward produce summaries.

Elozino Egonmwan et al [3][17][18]. They propose a framework that further develops execution on single record outline task utilizing the CNN/DailyMail and News studio datasets.

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

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

*P.*Krishnaveni *et al* [4]. The planned approach delivers automatic extractive heading sage text summarizer to recover the quality of text thereby enlightening the comprehensibility of the summary manuscript. It sums up the specified information report utilizing nearby scoring and neighborhood positioning that is it gives heading wise synopsis.

Shuxia Ren et al [5]. In this paper they have crossover the upsides of the extractive also abstractive outline frameworks to recommend text synopsis prototypical of joining worldwide fenced unit and duplicate component (GGUC). The analysis consequences exhibit that the presentation of the model is superior to the next message outline framework on LCSTS datasets.

Chandra Prakash et al [6]. In this paper, Human supported manuscript summarizer "SAAR" is actuality planned for single record. As of the report, a term-sentence grid is created. The passages trendy the grid is weight as of Reinforcement Learning. Subsequently, formed outline is displayed to the client then on the off chance that the client supports it then it is the last rundown, in any case new synopsis is produced according to the client input in type of catchphrases.

Ashish Vaswani et al [7]. They propose another basic organization design, the Transformer, in view of on consideration instruments, forgoing repeat and convolutions totally. Probes two machine interpretation undertakings demonstrate these models to be predominant in quality and requiring altogether less an ideal less amount of time to train.

Ayesha Ayub Syed et al [8]. This review will in general study the logical writing to acquire data and information about the current research in involuntary text summarizers explicitly abstractive outline dependent proceeding neural networks. An audit of different neural organizations founded abstractive synopsis models take part to be introduced.

Reeta Rani et al [9]. This survey paper gives the outline of different past investigates and study in the arena of Automatic Text Summarization.

Wang Guan et al [10]. This article reviews in detail various methods and evaluation patterns. The fundamental consideration is happening the uses of the most recent patterns, neural network created, then 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 creates stacks of self-attention layers.

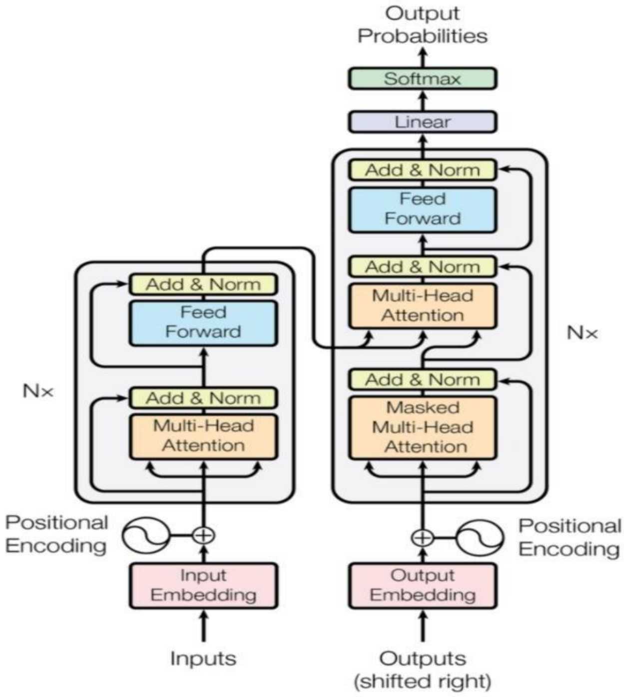

Fig. 1. Neural Attention Model

The encoder(left) and decoder(right) blocks are at least one of similar encoders and decoders assembled together. Together the encoder stack and the decoder stack are in similar extents The quantity of these stacks is a hyperparameter [11].

- The word of the input sequence is approved to the first encoder.
- These be situated further transmuted and passed on to the following encoder.
- 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 the 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 dot product and getting useful 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 summary after

B. Steps / Algorithm

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

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 necessary pre-processing on the data. We also tokenize the input so it can be processed as each word instead of a string of words. We implement the de-tokenize routine to detokenize it for 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, Feedforward subroutines.
Step 3: - Training In this step, we train the model using the summaries already generated in dataset. Using the hyperparameters pre- determined the model trains itself and improves the accuracy.
Step 4: - Evaluation In this step, the model's accuracy will be evaluated at various steps [14][16].

IV. RESULTS

We trained the model with 6 layers, 8 heads as hyper parameters. We found that after 1000 steps the accuracy. It increased to 0.1173 as compared to the accuracy of 0.04255obtained at 10 steps. Further results are illustrated in Fig 3. The way accuracy progresses with no. of steps in thetraining loop is shown in fig.1.

Table 1. Results Obtained While Training the Model

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. of Layers in the feed-forward neural network
- Steps – No. 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

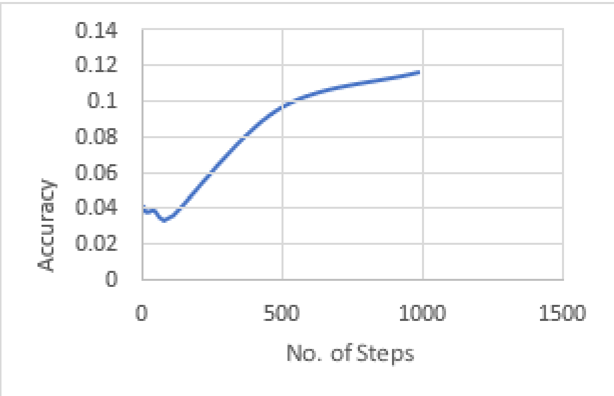

Fig. 3. Accuracy on Y-axis and No. of Steps on X-axis

V. CONCLUSION

With developing digital media and steadily growing publishing, no one has time to go through all the articles to get the right data. Thus, there is a demand for an Automated Text Summarizer which will shorten the source text, while maintaining its information and overall meaning. So, we implemented the transformer model for the task of summarization. We concluded from our results that after 1000 steps the accuracy increased to 0.1173 as 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 and consent to participate with evidence.
Availability of Data and Material	Not relevant.
Authors Contributions	All authors have equal participation in this article.

REFERENCES

1. 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.
2. 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
3. Elozino Egonmwan and Yllias Chali, "Transformer-based Modelfor Single Documents Neural Summarization", Proceedings of the 3rd Workshop on Neural Generation and Translation, Association for Computational Linguistics, November 2019. https://doi.org/10.18653/v1/D19-5607
4. 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). https://doi.org/10.1109/ICCMC.2017.8282539
5. Shuxia Ren, Kaijie Guo, "Text Summarization Model of Combining Global Gated Unit and Copy Mechanism", 2019 IEEE 10th International Conference on Software Engineering andServiceScience (ICSESS).
6. Chandra Prakash, Dr. Anupam Shukla, Human Aided Text Summarizer "SAAR" using Reinforcement Learning, 2014 International Conference on Soft Computing & Machine Intelligence. https://doi.org/10.1109/ISCMI.2014.22

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

7. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin,"AttentionIs 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.
15. Chellatamilan, T., Valarmathi, B., & Santhi, K. (2020). Research Trends on Deep Transformation Neural Models for Text Analysis in NLP Applications. In International Journal of Recent Technology and Engineering (IJRTE) (Vol. 9, Issue 2, pp. 750–758). https://doi.org/10.35940/ijrte.b3838.079220
16. Dhawale, A. D., Kulkarni, S. B., & Kumbhakarna, V. M. (2020). Automatic Pre Processing of Marathi Text for Summarization. In International Journal of Engineering and Advanced Technology (Vol. 10, Issue 1, pp. 230–234). https://doi.org/10.35940/ijeat.a1803.1010120
17. Saklecha, A., Uplavdiya, P., & Chawla, Prof. M. P. S. (2023). An Extensive Survey on Investigation Methodologies for Text Summarization. In Indian Journal of Signal Processing (Vol. 3, Issue 4, pp. 1–6). https://doi.org/10.54105/ijsp.d1016.113423
18. Pal*, A., Saha, S., & Anita. (2020). Musenet : Music Generation using Abstractive and Generative Methods. In International Journal of Innovative Technology and Exploring Engineering (Vol. 9, Issue 6, pp. 784–788). https://doi.org/10.35940/ijitee.f3580.049620
19. Arya, V., Khan, R., & Aggarwal, Prof. M. (2022). A Chatbot Application by using Natural Language Processing and Artificial Intelligence Markup Language. In International Journal of Soft Computing and Engineering (Vol. 12, Issue 3, pp. 1–7). https://doi.org/10.35940/ijsce.c3566.0712322

AUTHORS PROFILE

Dr. Shruti Sapra (Thakur) is a Research Scholar in PG Department of Computer Science of Sant Gadge Baba Amravati University Completed Ph. D on 22 March 2024 and Assistant Professor in G H Raisoni College of Engineering Nagpur. She has having has 12+years' experience in teaching and 3+Years in research. Published 25+ nos. research papers in national and international journals and conferences. I published 5 nos. patents, 1 Book Chapter in SCOPUS & 2 Copyright.

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 of Sant Gadge Baba Amravati University Amravati. He has 18+years' experience in teaching and research. He published 53+ nos. research papers in national and international journals and conferences. He published 6 nos. patents & 7 Copyrights.

Dr. Mohammad Atique is a Professor and Head in 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 & Head in P.G. Department of Computer Science and Engineering, Sant Gadge Baba Amravati University, Amravati.