iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** III **Month of publication:** March 2023

DOI: https://doi.org/10.22214/ijraset.2023.49329

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue III Mar 2023- Available at www.ijraset.com

Abstractive Text Summarization using Deep Learning

Rishank Tambe[1], Disha Thaokar[2], Eshika Pachghare[3], Prachi Sawane[4], Pranay Mehendole[5], Asst. Prof. Priti Kakde[6]
[1, 2, 3, 4, 5, 6]*Department of Information Technology G.H. Raisoni college of Engineering Nagpur, India*

Abstract: ***The number of text records has increased dramatically in recent years, and social media structures, including websites and mobile apps, will generate a huge amount of statistics about non-text content. structure, including blogs, discussion forum posts, technical guides, and more. Statistics, which constitute human behavior and intuitive thinking, consist of many records that are relatively difficult to manage due to their large number and various factors. However, the demand for statistics summarizing textual content is increasing. Text summarization is a way of analyzing unstructured text and converting it into meaningful statistics for evaluation that will produce the necessary number of useful records. This study describes a deep learning method for effectively summarizing textual content. As a result, the reader receives a condensed and focused model of the unique textual content.***

Keywords: Abstractive, Deep Learning, Text Summarization, encoder-decoder, Attention mechanism.

I. INTRODUCTION

Abstract text summarization is a technique of interpreting text while focusing on parts that convey useful information without changing the original meaning of the source result. An abstract summary algorithm generates new sentences from given sentences by changing the words but not changing the exact meaning.

Abstract text summarization is a strategy for creating a concise, accurate summary of long texts while focusing on paragraphs that provide relevant information and preserve the overall meaning. The current multiplication of unstructured textual data circulating in the digital realm requires the development of automated text synthesis technologies that allow users to draw simple conclusions. We now have immediate access to massive amounts of information. The huge volume of textual data on the web has multiplied as the rapid expansion of the Internet makes it significantly difficult for further activities, such as document management, text classification, and informational research. believe.

Advanced deep learning algorithms are used in abstraction-based synthesis to interpret and condense the original text, just as humans do. Extraction is still commonly used because the text synthesis algorithms required for abstraction are more difficult to generate. To triumph over the grammatical mistakes of extraction techniques, summary device studying algorithms can generate new terms and sentences that constitute the maximum critical statistics from the supply text. However, the field of DL-based abstract text summarization currently lacks a thorough literature review, however, this article provides a comprehensive overview of DL-based abstract text summarization.

II. METHODOLOGY

A. Sequence-to-Sequence Modeling

Machine learning includes the neural network method referred to as "sequence-to-sequence learning," which is primarily utilized in models for language processing.

It can be made using encoder-decoder-based machine interpretation, which transforms an input sequence into a series of output sequences with a tag and consideration esteem, and recurrent neural networks (RNNs).

The objective is to attempt to anticipate the following state arrangement based on the previous succession using two RNNs that will cooperate using a unique token.

Sequence-to-sequence models can also be implemented using attention-based techniques. Sequence-to-sequence models are neural networks that take a sequence from one domain (for instance, text vocabulary) as input and produce a new sequence in a different domain (i.e. summary vocabulary).

There are two major components of a sequence-to-sequence model are

1) Encoder
2) Decoder

©IJRASET: All Rights are Reserved

Encoder-Decoder: The encoder will record the input sequence's information and represent it in a hidden state. The output order will be predicted by the decoder using the encoder's most recent hidden input state. The two major methods used to increase the encoder's efficiency are reversing the input text (reverse encoder) and bidirectional encoding. The reverse encoder receives the input sentence in its reversed form. Alternative: bidirectional RNN. The past and future are only vaguely known.

B. Attention Mechanism

A complex cognitive ability required for humans is attention. An important feature of perception is that people often do not validate most of the information at once. Instead, people tend to selectively focus on a certain piece of information when and where it is needed while at the same time ignoring other noteworthy information. The efficiency and accuracy of cognitive information processing is significantly increased by the attention mechanism.

The encoder-decoder model for machine translation has been improved with the addition of an attention mechanism. The idea behind the attention mechanism is to give the decoder flexible access to the most important components of the input sequence through a weighted permutation of all the encoded input vectors, the best-fit vectors get the highest weights.

In this project we use global attention, which is a kind of attention mechanism.

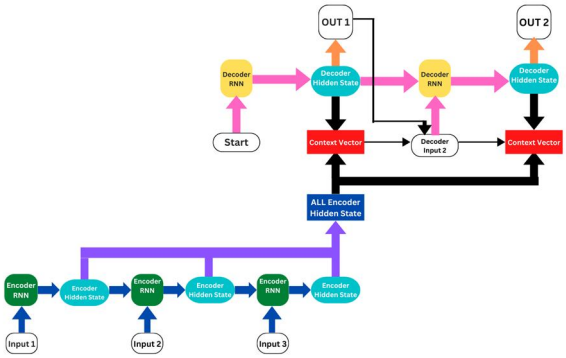


Fig. encoder-decoder architecture using an attention mechanism

III. IMPLEMENTATION

First of all, the text summarizer will take input from the user in the form of text which has to be summarized. Then the entered data is preprocessed to understand the distribution of the sequences. This data is summarized and presented as an output that is ready to share.

A. Datasets

Data set is a collection of data. Data sets can also contain a collection of documents or folders..

We used fine cuisine ratings from Amazon as the dataset for our model. The collection spans more than ten years and contains every one of the 500,000 reviews up until October 2012. Reviews include ratings, simple text reviews, information about the product and the user, and other information. Reviews from every other Amazon category are also included.

B. Libraries

1) *Numpy:* Working with arrays is done using Numpy.
2) *Pandas:* Python's primary library for scientific computing is called Numpy.
3) *Attention layer:* Official Keras support for the attention layer is not available. We'll pick the third-party model for our example.
4) *Keras:* A high-level deep learning API called Keras is developed in Python and is used to make it simple to create neural networks.
5) *Tensorflow:* It is used for data automation, model tracking, performance monitoring, and model retraining.

C. Flowchart

After extracting the dataset and importing the libraries, we preprocess the dataset. The dataset is then cleaned, i.e. removing:

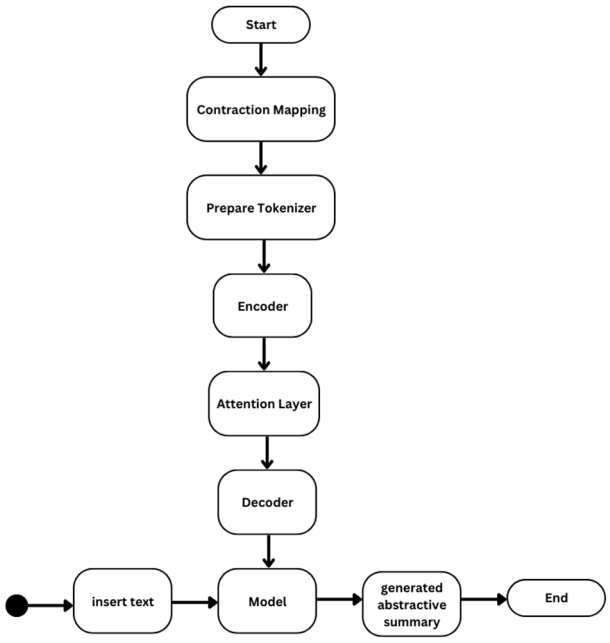
1) Convert everything to lowercase
2) Remove HTML tags
3) Contraction mapping

4) Remove (‘s)
5) Remove any text inside the parenthesis ( )
6) Eliminate punctuation and special characters
7) Remove stopwords
8) Remove short words

After that, we stored the text and summary cleaned through these parameters for further use. The next step is to Divide the data for training and testing. We kept 90% of the data for training the model and 10% of the data for testing it.

After that, we prepared a tokenizer for texts and summaries. A tokenizer creates the vocabulary and transforms word sequences into integer sequences.

We’re finally at the model-building part. Our model-building part is divided into two phases, the training phase, and the inference phase.



IV. RESULT

The predicted summaries are:

```
Review: best powder contain aluminum important want feed family
Original summary: baking
1/1 [==============================] - 0s 174ms/step
1/1 [==============================] - 0s 70ms/step
1/1 [==============================] - 0s 65ms/step
1/1 [==============================] - 0s 62ms/step
Predicted summary:  great product

Review: gluten free want crackers one also delicious second order
Original summary: cracker of cracker
1/1 [==============================] - 9s 9s/step
1/1 [==============================] - 2s 2s/step
1/1 [==============================] - 0s 54ms/step
1/1 [==============================] - 0s 62ms/step
1/1 [==============================] - 0s 63ms/step
1/1 [==============================] - 0s 64ms/step
Predicted summary:  great gluten free crackers
```

V. RELATED WORK

A. Experimental Research on Encoder-Decoder Architectures with Attention for Chatbots

Marta R. Costa-jussà1; Álvaro Nuez1; Carlos Segura2

To overcome the aforementioned drawback of the basic RNN-based encoder-decoder approach, decoders often employ an attention mechanism. In this case, instead of relying on the encoder's ability to compress the entire input sequence into the thought vector, the decoder computes a context vector consisting of the weighted sum of all of the hidden state vectors of the encoder for each generated word.

B. A Survey on Abstractive Text Summarization

Moratanch; Dr. S. Chitrakala

The study's author offers a complete evaluation of abstraction-based text summarizing techniques. Structured and semantic approaches to abstract abstraction are discussed in general terms in this article.

The author reviews a number of studies on two methods of abstraction.

C. Attention Mechanism for Neural Machine Translation: A survey

Weihua He; Yongyun Wu; Xiaohua Li

To identify images in the neural network region, use the attention mechanism to the recurrent neural network model. Bahdanau et al. employed the attention mechanism in machine translation reports to continually generate translation and alignment. Attention mechanisms have now been a common component of neural architectures and have been used for a variety of tasks, including the creation of picture captions, document classification, machine translation, motion detection, image-based analysis, speech recognition recommendation, and graph.

VI. FUTURE SCOPE

Search engines increasingly provide direct answers to some straightforward factoids as queries, going beyond the standard keyword-based document retrieval. However, there are still many questions that cannot be answered in a few sentences or in a single assertion. These non-factoid queries include, but are not limited to, definitions, arguments, steps, opinions, etc. To give accurate, comprehensive responses to these questions, it is required to combine and summarize the data from one or more studies. Due to the complexity of these issues, there haven't been many advances in recent years that have outperformed traditional information retrieval techniques. Future developments in discourse analysis and natural language processing are hoped for.

The majority of current techniques for identifying pertinent information still rely on occurrence frequency or surface-level features. There is still a sizable quality gap between automatic coming in streams and user-generated content of all kinds, including news texts and other types of information. Most of the approaches to general summarizing that have been discussed could be difficult to adapt for large-scale streaming data, which could lead to a loss of either efficiency or efficacy. A more specialized method is necessary due to the challenges of event recognition, dynamics modeling, contextual dependency, information fusion, and credibility assessment.

VII. CONCLUSION

The relevance of text summaries has increased in recent years due to the large amount of data available on the Internet. Text summarization techniques are classified as extractive or abstract. Based on linguistic and statistical aspects, the extracted text summary method provides a word and sentence summary of the original text, while the abstract text summary method rewrites the original text to generate produce a single sentence summary. This study analyzed the current algorithms that are based on deep learning for abstract text summarization. In addition, the difficulties encountered in applying this method, as well as their responses, were explored and analyzed.

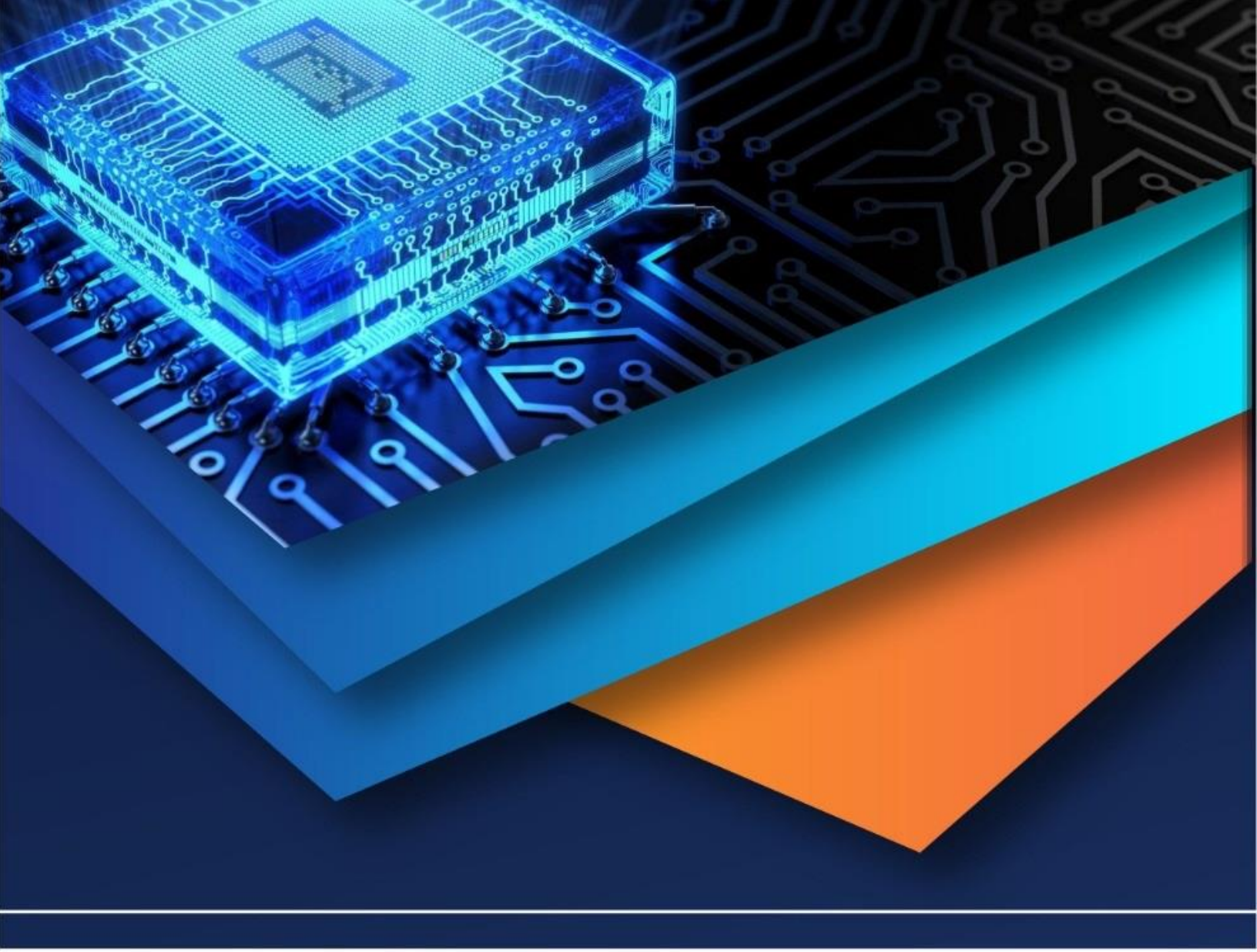
REFERENCES

[1] Opidi, A., 2019. A Genuine Introduction to Text Summary in Machine Learning. Blog, FloydHub, April, 15.
[2] Lloret, E., 2008. Text summary: an overview. Paper supported by the Spanish Government under the assignment TEXT-MESS (TIN2006-15265-C06-01).
[3] Kovačević, A. and Kečo, D., 2021, June. Bidirectional LSTM Networks for Abstractive Text Summarization. In International Symposium on Innovative and Interdisciplinary Applications of Advanced Technologies (pp. 281-293). Springer, Cham.
[4] Yang, L., 2016. Abstractive summarization for amazon reviews.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

[5] . A novel approach to workload prediction using attention-based LSTM encoder-decoder network in the cloud environment. EURASIP Report on Wireless Communications and Networking, 2019(1), pp.1-18.
[6] Fabbri, A.R., Kryściński, W., McCann, B., Xioing, C., Socher, R. and Radev, D., 2021. Summeval: Re-evaluating summary evaluation. Transactions of the Association for Computational Language, 9, pp.391-409.
[7] Bhati, V. and Kher, J., 2019. Survey for Amazon fine food reviews. Int. Res. J. Eng. Technol.(IRJET), 6(4).
[8] Syed, A.A., Gaol, F.L. and Matsuo, T., 2021. A survey of the state-of-the-art prototype in neural abstractive text summary. IEEE Access, 9, pp.13248-13265.
[9] Raphal, N., Duwarah, H. and Daniel, P., 2018, April. Survey on abstractive text summarization. In 2018 Global Summit on Communication and Signal Processing (ICCSP) (pp. 0513-0517). IEEE.
[10] Sherstinsky, A., 2020. Basics of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D: Nonlinear Phenomena, 404, p.132306.
[11] Shewalkar, A., 2019. Performance judgment of deep neural networks applied to voice recognition: RNN, LSTM, and GRU. Thesis of Artificial Intelligence and Soft Computing Research, 9(4), pp.235-245.
[12] He, W., Wu, Y. and Li, X., 2021, October. Attention Appliance for Neural Machine Translation: A survey. In 2021 IEEE 5th IT, Networking, Electronic and Automating Control Conference (ITNEC) (Vol. 5, pp. 1485-1489). IEEE.
[13] Song, S., Huang, H. and Ruan, T., 2019. Abstractive text summarization using LSTM-CNN-based deep learning. Multimedia Tools and Applications, 78(1), pp.857-875.
[14] Costa-jussà, M.R., Nuez, Á. and Segura, C., 2018. Experimental research on encoder-decoder architectures with attention to chatbots. Computación y Sistemas, 22(4), pp.1233-1239.
[15] Niu, Z., Zhong, G. and Yu, H., 2021. An analysis of the attention mechanism of deep learning. Neurocomputing, 452, pp.48-62.
[16] Xi, W.D., Huang, L., Wang, C.D., Zheng, Y.Y. and Lai, J., 2019, August. BPAM: Guidance grounded on BP Neural Network with Attention Mechanism. In IJCAI (pp. 3905-3911).

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)