Voice Based Summary Generation using LSTM

Ruchi Sharma[1], Nikhil Varshney[2], Manya Sharma[3], Prof. Ruchi Paliwal[4]

[1, 2, 3, 4]*Department of Electronics & Communication, JSSATEN-NOIDA Noida, India*

Abstract: Summary generation is utilized to have a fast outline of a total meeting without wasting much time. In this research paper, we are making a pipeline to create a summary of a voice recording. We will change the voice recording into text and afterward, we will utilize a deep learning model to produce an outline of that voice recording. We have 10 min voice recording as information and a rundown as result. Rundown age is essential of two sorts: Abstractive synopsis and extractive outline. Abstractive rundown creates new sentences with comparable significance as a synopsis of huge sentences and extractive outline produces sentences in the wake of pulling just significant expressions and catchphrases of enormous sentences.
Keywords: Text Summary, Extractive, Abstractive, Natural Language processing, Machine Learning
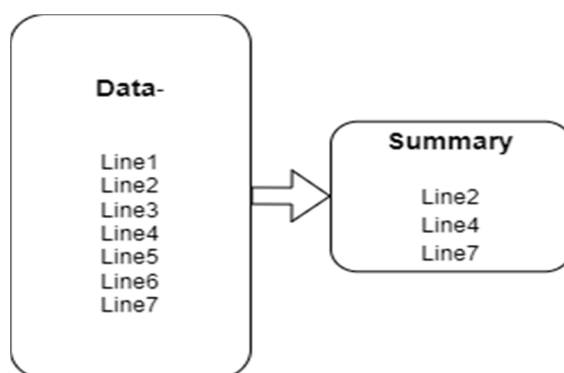
I. INTRODUCTION

The text summary is broadly used to save time. We have a ton of online stages where we can utilize message synopsis highlights to abbreviate the message into more limited sentences. In an exploration paper distributed by the International Research Journal of Engineering and Technology (IRJET) and composed by Kasimahanthi Divya, Kambala Sneha, Bassetti Sowmya, G Sankara Rao, they utilized encoder-decoder design named as a succession to arrangement model of Long Short Term Memory (LSTM) which goes under profound learning. They depicted techniques named as spotless information, construct information, assemble dict, and tokenize to tidy up the dataset and set up the dataset. CNN_dailymail dataset is utilized and for approval, they involved ROUGE which means Recall-Oriented Understudy for Gisting Evaluation.[5][10][20]

Another paper Natural Language Processing (NLP) based Text Summarization - A Survey which is distributed by IEEE Xplore utilized directed, solo and support strategies to recognize the different procedures to create the outline. They characterized the two sorts of synopsis age under these above composed three classifications. Like they characterized an abstractive and extractive rundown age in the administered classification first then in solo and finally in the support category.[21]-[25]

An research paper distributed by the International Journal of Advanced Research in Computer and Communication Engineering and named A Review Paper on Text Summarisation utilized a design based approach and semantic-based approach for abstractive synopsis [2][11]

In this research paper, we will create a voice-based synopsis. Fundamentally, the voice is recorded and afterward, that is utilized as a contribution to creating its rundown. We utilized python to change over sound into text and afterward applied the encoder-decoder engineering of LSTM to produce the outline. All research papers depend on text-based outline age. Yet, here we are creating a synopsis from sound as info. We utilized LSTM which is a deep learning calculation and has encoder-decoder engineering with an inserting and thick layer. As referenced in the theoretical, we have two wide arrangements of synopsis age, so here is a portrayal for that to explain conceptual outline and extractive summarization..

A. Extractive Summarization Depiction

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

B. Abstractive Summarization Depiction
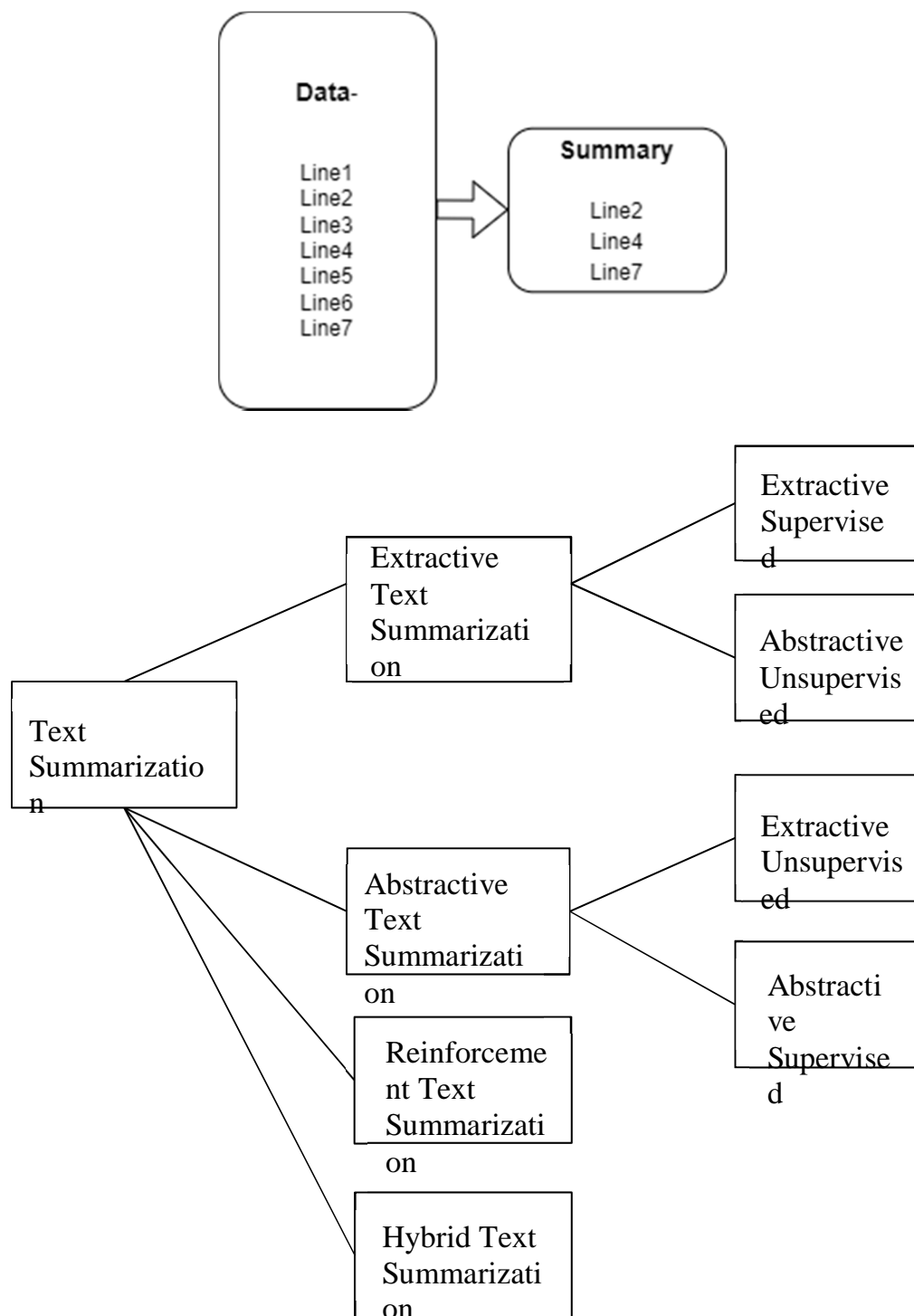
Fig. 1. Classification for Text summary generation process[25]

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

| Dataset	Description
CNN/Daily Mail	It comprises both articles and summaries of long news articles.
Gigaword	It comprises nearly ten million documents, articles, and their headlines (over four billion words) of the original English Gigaword Fifth Edition.
NYT	The New York Times dataset has the full text and metadata of NYT articles from 1987 to 2007.
DUC	The Document Understanding Conference (DUC) archives and synopses have been assessed by the National Institute of Standards and Technology (NIST) since 2001.
20NG	It consists of 19,997 papers compiled in 20 separate categories from the Usenet newsgroup archive. 80 percent of the documents in 20NG were used for MNB preparation for assessment purposes and the remaining 20 percent for classification assessment.
TIDSUMM	TIDSUMM contains Darknet utilization information with 6831 documents of 26 distinct classifications crawled over the onion web or Tor network.
TTNews	A Chinese news synopsis corpus, made for the common outline task at NLPCC 2017
SummMac	Sumac contains records about computer science gathered from ACL-sponsored conferences.

Table 1.1 Dataset for Summarization[8][9][11]

II. METHODOLOGY AND WORKFLOW

A. Methodology

We utilized a deep learning based technique for an intermittent brain organization to produce an outline of information text. An individual sound is recorded to portray the gathering voice recording. The recorded sound is changed over into text utilizing the speech_recognization library of python and none of the python libraries support the transformation of long sounds appropriately so we utilized an additional contention of record work i.e length to store demanded time spans in the division of its greatest cutoff. In the event that you won't specify this boundary then just 1 moment of sound is changed over into text as a default. This change over text is our trying information essentially which we will use to test our deep learning model. We utilized the news rundown dataset to prepare the LSTM model. It gives many lines of information, which is a great measure of information to prepare a deep learning model with good accuracy and validation statistics. This trained LSTM model is additionally used to produce our very own outline input text which we have made by switching recorded sound over completely to message.

Everything in this cycle is composed advance-wise in the work process area underneath.

B. Workflow
1) Gathering input information
2) Conversion into text
3) Dataset assortment for preprocessing
4) Making a model of profound learning.
5) Testing of the model on our own feedback text

a) *Input:* Voice recording is base info that will be changed over completely to text and afterward that text will go about as information that will be passed into the model design for a summary generation .

b) *Convert Audio Into Text:* We utilized the speech_recognization library of python to change the sound documents into text and it gives a contention named span in its capacity of record which can be utilized to record or store just that specific term of source sound into a variable. So that's what we utilized and changed over 10 minutes of sound into text by separating them into pieces to get more precise text out of the sound information. Recognize_google is utilized to perceive that sound and convert it into text. Then store that text in a solitary text record.

c) *Dataset Preparation For Model Training:* The news summary dataset is utilized from Kaggle. Dataset has segments named writer, date, read_more, text, and c_text. Furthermore, another CSV record has two sections named text and title.
The title section is utilized as an outline and text as information. The dataset contains two CSV records news_summary.csv and news_summary_more.csv.
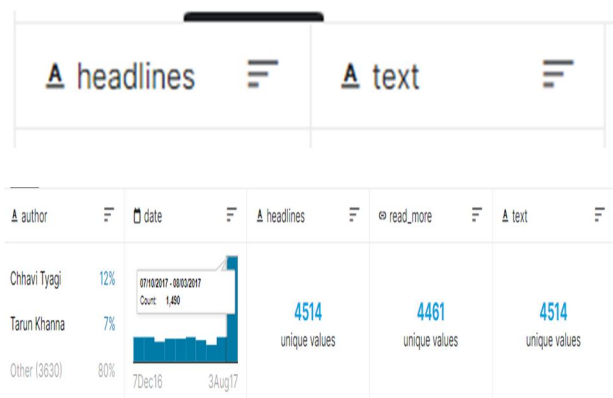
The first CSV dataset comprises 4515 models and contains Author_name, Headlines, Url of Article, Short text, and Complete Article.

For dataset cleaning-Remove get away from characters, eliminate exceptional characters, and convert all sentences to bring down case. Eliminate single characters and various Spaces. Regex is utilized to do this cleaning in lines of sections.

Dataset readiness Merged all segments in the text section of the recently made information outline with the exception of the title segment and afterward made two segments in the information outline for text and rundown.

```
preText1 =  raw_data.iloc[:,0:2].copy()
# pre1['head + text'] = pre1['headlines'].str.cat(pre1['te
xt'], sep =" ")
# print(pre1)
#pre1 is concatanation of headline and text
#pre2 is concatanation of headline, authorname, date,url,
read more, ctext, text
preText2 = summary_data.iloc[:,0:6].copy()
preText2['text'] = preText2['author'].str.cat(preText2['d
ate'].str.cat(preText2['read_more'].str.cat(preText2['tex
t'].str.cat(preText2['ctext'], sep = " "), sep =" "),sep=
" "), sep = " ")
# print(pre2['text'][0])
```

```
preText = pd.DataFrame()
preText['text'] = pd.concat([preText1['text'], preText2
['text']], ignore_index=True)
preText['summary'] = pd.concat([preText1['headlines'],pre
Text2['headlines']],ignore_index = True)
```

d) *LSTM Model:* Sequence to arrangement model is utilized with encoder-decoder design. Three layers of encoder LSTM and one layer of decoder LSTM are utilized. Inserting and thick layers are additionally utilized in model creation.
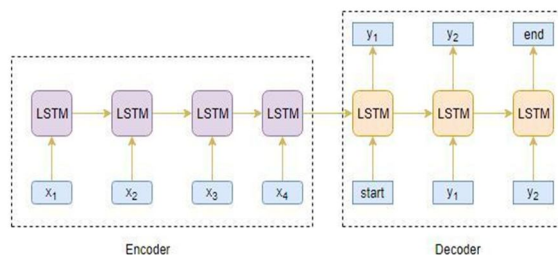
<start> and <end> are the special tokens that are added to the target grouping prior to taking care of it into the decoder.

The model is prepared on 88517 examples, the model is approved on 9836 examples.

The encoder model of the Long Short Term Memory model (LSTM) peruses input single word by single word, just a single word is taken care of into the encoder model at a time.

The decoder LSTM network peruses the whole objective succession word-by-word and predicts the grouping.

e) *Testing of the model:* The encoder model and decoder model are made. A capacity named decode_seq is composed to disentangle token-based groupings into words. Two variables are made to store the record of words from the beginning and from the final word or we can say to get the file of the expression of the switched sentence. Padding was done to adjust the length of sentences. The Decode_seq capacity will return the synopsis of info cushioned text information.

```
print("Predicted summary:",decode_seq(test_data4_pad))

Predicted summary:  indian plates pressure i
ndian plates   india's by     by   and to th
e the the the
```

III. CONCLUSION

These days time is viewed as the main part of life. Individuals attempt to deal with their time in a ton of ways. In this way having a method for getting a fast outline of huge settings or enormous gatherings is truly significant in everyday life. Through this research paper, we have made a method for getting an extractive synopsis of essentially a 10-minute voice recording. We can create a rundown of huge gathering accounts also along these lines. There isn't any time limitation here in view of the time span of the gathering, all you really want is to break that changed over message into little chunks in light of the fact that LSTM is a repetitive organization and in some cases, it doesn't work precisely over enormous message due to an absence of network and exactness in the memory cell of LSTM.

REFERENCES

[1] U. Hahn, I. Mani, "Automatic Researchers are investigating summarization tools and methods that", IEEE Computer 33. 11, pp. 29-36, November 2000.

[2] Reeve Lawrence H., Han Hyoil, Nagori Saya V., Yang Jonathan C.,Schwimmer Tamara A., Brooks Ari D, "Concept Frequency Distribution in Biomedical Text Summarization", ACM 15th Conference on Information and Knowledge Management (CIKM),Arlington, VA, USA,2006.

[3] K. Spärck Jones, "Automatic summarizing: The state of the art", Information Processing & Management, vol. 43, pp. 1449-1481, nov 2007

[4] E. Lloret, M. Palomar, "Text summarization in progress: a literature review" in Springer, Springer, pp. 1-41, 2012

[5] M. Kamal, "Text Summarization in Bioinformatics", 15th International Conference on Computer and Information Technology, pp. 592-597, 2012

[6] Suneetha Manne, Zaheer Parvez Shaik Mohd. , Dr. S. Sameen Fatima, "Extraction Based Automatic Text Summarization System with HMM Tagger", Proceedings of the International Conference on Information Systems Design and Intelligent Applications, 2012, vol. 132, P.P 421-428.

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

[8] Saranya Mol C S, Sindhu L, "A Survey on Automatic Text Summarization", International Journal of Computer Science and Information Technologies, 2014,Vol. 5 Issue 6.

[9] Khan Atif, Salim Naomie, "A review on abstractive summarization Methods", Journal of Theoretical and Applied Information Technology, 2014, Vol. 59 No. 1.

[10] 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

[11] Deepali K. Gaikwad and C. Namrata Mahender, "A Review Paper on Text Summarization", International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 3, March 2016

[12] Padmakumar, A., & Saran, A. (2016). "Unsupervised Text Summarization Using Sentence Embeddings (pp. 1-9). Technical Report", the University of Texas at Austin.

[13] Reeta Rani and Sawal Tandon, "Literature Review On Automatic Text Summarization," International Journal of Current Advanced Research Volume 7; Issue 2(C); February 2018.

[14] Dohare, S., Gupta, V., & Karnick, H. (2018, July). "Unsupervised semantic abstractive summarization", In Proceedings of ACL 2018, Student Research Workshop (pp. 74-83).

[15] Schumann, R. (2018), "Unsupervised abstractive sentence summarization using length controlled variational autoencoder", arXiv preprint arXiv:1809.05233.

[16] Wang, Y. S., & Lee, H. Y. (2018), "Learning to encode text as human-readable summaries using generative adversarial networks", arXiv preprint arXiv:1810.02851.

[17] Raphal, Nithin, Hemanta Duwarah, and Philemon Daniel. n.d. "Survey on Abstractive Text Summarization." International Conference on Communication and Signal Processing, April 3-5, 2018, India.

[18] Prabhudas Janjanam, CH Pradeep Reddy, "Text Summarization: An Essential Study" Second International Conference on Computational Intelligence in Data Science (ICCIDS-2019)

[19] Chu, E., & Liu, P. (2019, May), "MeanSum: a neural model for unsupervised multi-document abstractive summarization" International Conference on Machine Learning (pp. 1223-1232).

[20] Kasimahanthi Divya, Kambala Sneha, Baisetti Sowmya, G Sankara Rao, "Text Summarization using Deep Learning," International Research Journal of Engineering and Technology (IRJET) Volume: 07 Issue: 05 | May 2020.

[21] Gonçalves, Luís. 2020. "Automatic Text Summarization with Machine Learning — An overview summarization wit machine learning an overview 68ded5717a2

[22] Zhang, X., Zhang, R., Zaheer, M., & Ahmed, A. (2020). "Unsupervised Abstractive Dialogue Summarization for Tete-a-Tetes", arXiv preprint arXiv:2009.06851

[23] Zheng, C., Wang, H. J., Zhang, K., & Fan, L. (2020) "A Baseline Analysis for Podcast Abstractive Summarization", arXiv preprint arXiv:2008.10648.

[24] Yang, Z., Zhu, C., Gmyr, R., Zeng, M., Huang, X., & Darve, E. (2020), "TED: A Pretrained Unsupervised Summarization Model with Theme Modeling and Denoising", arXiv preprint arXiv:2001.00725.

[25] Ishitva Awasthi, Kuntal Gupta, Prabjot Singh Bhogal, Sahejpreet Singh Anand, Prof. Piyush Kumar Soni, "Natural Language Processing (NLP) based Text Summarization - A Survey" Proceedings of the Sixth International Conference on Inventive Computation Technologies [ICICT 2021]

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ◎ (24*7 Support on Whatsapp)