YouTube Transcript Summarizer

Surabhi Sunil Maydeo Master of Science in Information Technology

*K.C. College, HSNC University, Mumbai 400 020, India* [surabhi.maydeo@gmail.com](mailto:surabhi.maydeo@gmail.com)

*Dr. Rakhi Gupta*

*Head of the Department*

*I.T Department*

*K.C. College, HSNC University, Mumbai 400 020, India* [*rakhi.gupta@kccollege.edu.in*](mailto:rakhi.gupta@kccollege.edu.in)

*Nashrah gowalkar*

*Asst professorI.T Department*

* 1. *College, HSNC University, Mumbai 400 020, India* [*nashrah.gowalker@kccollege.edu.in*](mailto:nashrah.gowalker@kccollege.edu.in)

***Abstract* — Currently in today’s world andgeneration we use a lot of YouTube videos for knowledge, reference and learnings. There is ample ofmaterial available on YouTube for every topic that we can search. When we hit search for a particular topic multiple results dropdowns relating to our search. But our main challenge is to find the right video with correctcontext. We have to browse through a lot of videos before finding the correct video, and this can be time consuming and unproductive effort. To avoid this, I would like to propose a YouTube Transcript Summarizer that could summarize and provide the key points from a video which would help the user understand what this video will cover and if the topic doesn’t interest the user, then the user can view other videos. Here I, will be converting speech to text. For this I will be using NLP processing for extracting text and BERT Summarizer to summarize the context, Also I will be using Machine learning algorithms to derive correct output from the speech and obtain maximum perfect result. This will provide a text description and abstract summary for the searched YouTube video. This will help the users to decide between relevant andirrelevant information that is being displayed. Hence further, my research can also be used to interpret text or summaries from different data sources like calls, meetings, zoom calls or online lectures.**

***Keywords— NLP, Bert Summarizer, Machine learning algorithms, speech to text generation, AI algorithms***

* + 1. INTRODUCTION

NLP is a subset of Machine Learning. Thistechnology gives computers or machines to interpret human language. It will help the machine to comprehendhuman language. Its primary usage is to convert human speech to text. Another branch of NLP is video summarization. This will summarize contents from the video and convert it. Here we have to capture the key summary from video. Video summarization will involve a combination of techniques such as text extraction, audio analysis, and image processing, among others. Our main idea is to able to find the short summary of YouTube video and present it in a textual format. This will be very helpful to the users as it will help them to find concise text in very short time. In the field of NLP we generate summaries of transcripts and produce human readable outputs

* + - 1. OBJECTIVE:

This YouTube transcript summarizer aims to condense and simplify the content of a YouTube video's transcript into a concise summary. The main objectiveof my research is stated below:

**Saving User’s time:** By providing a summarized transcript of the searched videos,users can quickly grasp the highlights and keypoints of a video without having to watch the entire video or content.

**Information reuse or retrieval:** This summarized content can help users find specific information within the video moreefficiently.

**Accessibility:** For people with some hearing impairments or limited time, a summary canmake the video's content more accessible.

**Education and learning of students as well as professionals:** Summaries can be used as study aids or to get the key concepts from educational videos.

**Content analysis and data study or Contentcategory allocation:** These summaries can be used for content analysis, such as identifying common themes or trends or patterns in a collection of videos.

Overall, my research on YouTube transcript summarizer will aim to improve the efficiency and accessibility of YouTube video content.

* + - 1. IMPORTANCE OF THE STUDY:

This Research Paper will help the users to generate summarized and short concise summary from YouTube content using Machine learning, Bert Summarizer and NLP- Natural Language Processing.

* + 1. LITERATURE REVIEW:

The research paper introduced Dr. T. Bala Murali Krishna talks about Abstractive Multi-Document Summarization via Phrase Selection and Merging. This paper proposed an abstraction-based multi- document summarization framework that can constructnew sentences by exploring more fine-grained syntactic units than sentences, namely, noun/verb phrases. Different from existing abstraction-based approaches, our method first constructs a pool of concepts and facts represented by phrases from the input documents. Thennew sentences are generated by selecting and merging informative phrases to maximize the salience of phrasesand mean while satisfy the sentence construction constraints. Here they employed integer linear optimization for conducting phrase selection

and merging simultaneously in order to achieve the global optimal solution for a summary. Experimental results on benchmark data set TAC 2011 show that our framework outperforms the state-of-the-art models under automated pyramid evaluation metric, and achieves reasonably well results on manual linguistic model.

* + 1. PROPOSED METHODOLOGY

My paper mainly focusses on getting input from the YouTube video and providing a clear and crisp summarized output.

*Getting URL*

A Uniform Resource Locator (URL) is a reference to an internet resource that provides the location of that resource on a computer network. It serves as a means to retrieve information from the internet. The user willenter a YouTube URL in the search box and click enter, then the system processes the URL in the background to obtain the necessary information. Thencheck whether it is valid URL or NOT, if it is valid URL then shortened the link process will happened in the backend. Then the required YouTube video will get from the link. It will process for the Video to Text Extraction. In this user will enter the URL, we will check whether the URL is valid or not. If the entered URL is not valid then the user will get an error message saying the entered URL is invalid or incorrect. If the URL is valid, then the correct URL will go forward in next step and it will pass the URL toNLP model.

*Video to text extraction*

After the required YouTube video is attained, then I will pass the URL from NLP model file to utubeextract.py file. Then by using the YouTube transcript API I can get the transcript for required YouTube video.

*NLP processing*

This Process is for Text Summarization. NLP contains many methods for text summarization. Here I will talk about the text summarization is process of shortening the big paragraph into short summary. If the paragraph contains lots of lines, it will need more time to cover the content. We will be able to convert the large text into small text by removing unimportant information. The process of cutting down huge chunks of text and data into understandable paragraphs is known as NLP Text processing or summarization.

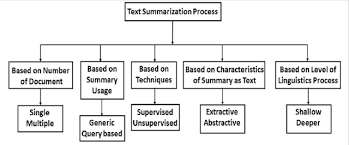
*BERT summarizer*

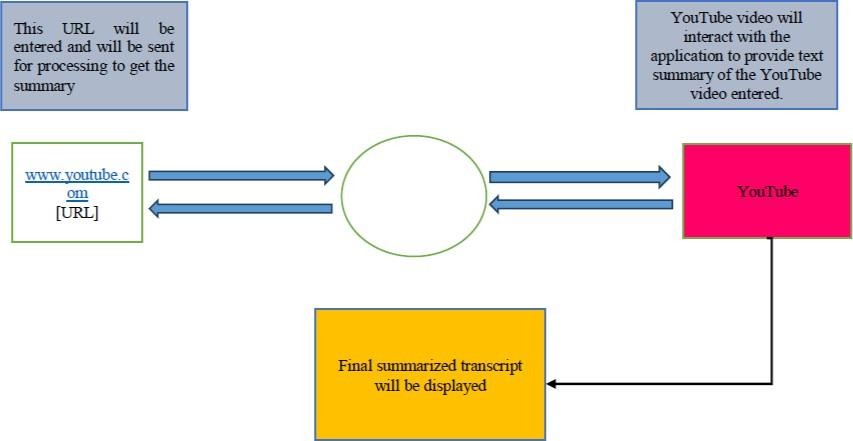
BERT is Bidirectional Encoder Representations from Transformers. This BERT introduces new advanced approach to deliver NLP tasks. BERT is one of the important algorithms of Natural Language Processing Models. The important information can be retained and

extracting the relevant information by Extractive Text Summarization. This extractive Summarization is more challenging. In this progress I will be using superior embedding’s that provided by encoder models like BERT. By taking two supervised approaches and use the BERT sentence embedding’s to build an extractive summarizer.

*Language translation*

The main objective is to create a Language Translator which would help us translate a word, sentence or evena paragraph to another language. I will try to incorporate as many languages as possible. Can use T-kinter Module to build our GUI for the project andgoogle-trans library to present us with a number of languages.





1. TESTING

Accuracy table:

Currently the limitation of this model are

**Accuracy:** The accuracy of a summary depends on the video's complexity and the algorithm used by the tool. For example, YouTube's automatic captions are only about 60- 70% accurate, and the accuracy can be worse in videos with background noise, accents, or multi-syllable words.

**Data quality:** The quality of the data fed into the summarizer can be an issue. For example, many YouTube uploaders don't include captions, and auto-generated translations aren't always perfect.

Bert Text Summarizer will provide with maximum accuracy.

|  |  |  |
| --- | --- | --- |
| Testing | Algorithm | Accuracy |
| Black box testing | Latent Semantic  Analysis (LSA) | 96% |
| Black box testing | Transformer  Model | 96.66% |
| Black box testing | Text Rank  Algorithm | 97.3% |
| Black box testing | Frequency Based  Approach | 96.95% |
| Black box testing | Bert Text  Summarizer | 98.2% |

1. POTENTIAL APPLICATIONS

**Educational environment:** Summarizers canbe used to create study material for students, helping them efficiently review and understand key concepts from lectures, tutorials, and other educational videos.

**Content Creators:** Content creators can use this summarizer to generate summaries of their videos, making it easier for viewers to find and consume their content.

**Research and Analysis:** My Summarizers canbe used for large-scale content analysis, enabling researchers to identify trends, patterns, and emerging topics within vast datasets of YouTube videos.

**Office Meetings or Client calls:** From thistool professionals can get key points or highlights of the meeting at the end of the meeting **Accessibility:** Summarizers can enhance the accessibility of YouTube content for individuals with hearing impairments or limited time, providing a valuable tool for inclusive content consumption.

1. CONCLUSION

In conclusion I can saymy model can be used to save time of students or other professionals. Instead of watching an entire video they can glance through the video and get a summary of a particular video. My model will provide multi language summarization and will contribute in saving time or a lot of people. Professionals can use this model in future during meetings as well to get the minutes of the meeting. I am confident that this paper will contribute in betterment is already published papers.

This summarizer also represents a significant advancement in information processing and accessibility. By automating the task of condensing or shortening lengthy transcripts into short precise summaries, my tool will offer numerous benefits to users, including time-saving, improved information retrieval, and increased accessibility for individuals with disabilities, students or working professionals.

1. LIMITATION
2. FUTURE SCOPE

In future my research can also be used to interpret text or summaries from different data sources like calls, meetings, zoom calls or online lectures etc. It can also be used as browser extensions in the future to increase efficiency.

1. REFERENCES
2. Abstractive Multi-Document Summarization via Phrase Selection and Merging∗ Lidong Bing§ Piji Li\ Yi Liao\ Wai Lam\ Weiwei Guo† Rebecca J. Passonneau‡

§Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA USA\Department of Systems Engineering and Engineering Management, The Chinese Universityof Hong Kong †Yahoo Labs, Sunnyvale, CA, USA ‡Center for Computational Learning Systems, Columbia University, New York, NY, USA.

1. [Https://www.lumina.sh/c5bbe32b-4fb7-476a-81aa-](https://www.lumina.sh/c5bbe32b-4fb7-476a-81aa-)

[3] https://doi.org/10.55041/ijsrem29057

1. Ghadage, Yogita H. and Sushama Shelke. “Speech to text conversion for multilingual languages.” 2016 International Conference on Communication and Signal Processing (ICCSP) (2016).
2. Pravin Khandare, Sanket Gaikwad, Aditya Kukade, Rohit Panicker, Swaraj Thamke, “Audio Data Summarization system using Natural Language Processing, ” International Research Journal of Engineering and Technology (IRJET),
3. (2019). 6 S.M.Mahedy Hasan, Md. Fazle Rabbi, Arifa Islam Champa, Md. Asif Zaman, “An Effective Diabetes Prediction System Using Machine Learning Techniques”, 2nd International Conference on Advanced Information and Communication Technology (ICAICT), (2020).
4. Prof. S. A. Aher , Hajari Ashwini M , Hase MeghaS, Jadhav Snehal B, Pawar Snehal S, “Generating Subtitles Automatically For Sound in Videos, ” International Journal of Modern Trends in Engineering and Research (IJMTER), 3, 3 (2016).
5. [https://www.researchgate.net/publication/3804627](https://www.researchgate.net/publication/380462773_Trends_and_challenges_in_sentiment_summarization_a_systematic_review_of_aspect_extraction_techniques?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6Il9kaXJlY3QiLCJwYWdlIjoiX2RpcmVjdCJ9fQ) [73\_Trends\_and\_challenges\_in\_sentiment\_summar](https://www.researchgate.net/publication/380462773_Trends_and_challenges_in_sentiment_summarization_a_systematic_review_of_aspect_extraction_techniques?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6Il9kaXJlY3QiLCJwYWdlIjoiX2RpcmVjdCJ9fQ)

[ization\_a\_systematic\_review\_of\_aspect\_extraction](https://www.researchgate.net/publication/380462773_Trends_and_challenges_in_sentiment_summarization_a_systematic_review_of_aspect_extraction_techniques?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6Il9kaXJlY3QiLCJwYWdlIjoiX2RpcmVjdCJ9fQ)

\_techniques?\_tp=eyJjb250ZXh0Ijp7ImZpcnN0UG [FnZSI6Il9kaXJlY3QiLCJwYWdlIjoiX2RpcmVjd](https://www.researchgate.net/publication/380462773_Trends_and_challenges_in_sentiment_summarization_a_systematic_review_of_aspect_extraction_techniques?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6Il9kaXJlY3QiLCJwYWdlIjoiX2RpcmVjdCJ9fQ) [CJ9fQ](https://www.researchgate.net/publication/380462773_Trends_and_challenges_in_sentiment_summarization_a_systematic_review_of_aspect_extraction_techniques?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6Il9kaXJlY3QiLCJwYWdlIjoiX2RpcmVjdCJ9fQ)

1. [Automatic customer review summarization using](https://www.researchgate.net/publication/380376077_Automatic_customer_review_summarization_using_deep_learning-_based_hybrid_sentiment_analysis?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6Il9kaXJlY3QiLCJwYWdlIjoiX2RpcmVjdCJ9fQ) [deep learning- based hybrid sentiment analysis](https://www.researchgate.net/publication/380376077_Automatic_customer_review_summarization_using_deep_learning-_based_hybrid_sentiment_analysis?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6Il9kaXJlY3QiLCJwYWdlIjoiX2RpcmVjdCJ9fQ)
2. https:/[/www.researchgate.net/publication/3642450](http://www.researchgate.net/publication/3642450) 17\_Opinion\_texts\_summarization\_based\_on\_texts

\_concepts\_with\_multi-opt