

SUMMARIZE PODCAST TRANSCRIPTS

Harshitha M¹, Deepika M N², Lamees Fathima³, Sachin K K⁴, Spoorthi R⁵

¹Harshitha M, ISE, Vidya Vikas Institute of Engineering & Technology Mysore

²Deepika M N, ISE, Vidya Vikas Institute of Engineering & Technology Mysore

³Lamees Fathima, ISE, Vidya Vikas Institute of Engineering & Technology Mysore

⁴Sachin K K, ISE, Vidya Vikas Institute of Engineering & Technology Mysore

⁵Spoorthi R, ISE, Vidya Vikas Institute of Engineering & Technology Mysore

Abstract - In the modern world, a significant amount of videos with information are uploaded every day. Selecting the appropriate video and comprehending the proper content. Despite the abundance of videos available, some of them will contain pointless material, and even if there is flawless content, it should still be needed of us. If the correct one is not located, all of your time and effort will be wasted trying to extract the useful information. We offer a novel concept that makes use of NLP. This gives consumers the ability to distinguish between pertinent and irrelevant information based on their needs by providing the video's key material in text description and abstractive summary. Additionally, our tests demonstrate that the joint model can provide high-quality, succinct, and understandable multi-line video summaries and descriptions in a human evaluation text extraction through processing and text summarization through BERT summarization.

Key Words: Video Summarization, Deep Neural Networks, Supervised Learning, Summarization Datasets, Evaluation Protocols.

1.INTRODUCTION

Natural Language Processing(NLP) is a branch of Artificial Intelligence that focuses on machine-human language interaction. Enabling machines to comprehend, interpret, and produce natural language speech or text is its main goal. video synthesis, The objective is to provide succinct, logical summaries that highlight the most important details from the original video. When there is a need for a brief synopsis of the

video material or when time is of the essence, this technology can be useful. Watch Video Usually, summarization combines a number of methods, including image processing, audio analysis, and text extraction. Finding a brief synopsis of a YouTube video and presenting it in written language is our major goal. Users will find brief summaries of text approaches quite beneficial as they allow us to quickly read the relevant content.

Our paper's primary objective is to help users work more efficiently and save time. Many people click through to YouTube videos based just on their attention-grabbing image and captivating title, wasting time watching pointless content. Students frequently. Before tests, they might look up YouTube videos; however, because of time limits, they would view them twice as quickly, which could cause them to become confused about the material. Having access to recorded sessions and transcripts of meetings can be helpful in obtaining a summary of the video content, saving time and effort. The main focus of our paper is to extract the most important information from the transcript and present it in a concise paragraph. Our objective is to save users' time by providing them with relevant and useful information on their desired topic.

To make this procedure easier, automatic text summarizing algorithms have been created recently. Text summary is the process of reducing text to a concise form that people can understand. Though it can be difficult, text summarization task because of the limits of robots in comprehending human language and knowledge, but it has several advantages, including the ability to organize content, summarize it, retrieve data, and respond to questions. More effective and quick text

summary techniques have been made possible by technological improvements, whereas previous research mostly concentrated on condensing and summarizing texts from individual documents.

2. METHODOLOGY

The main goal of this essay is to give concise, accurate, and clear summaries of YouTube videos so that readers won't have to waste their time. For every purpose in this study, more well used Python libraries are used. For extracting transcripts, these are the YouTube Transcript API.

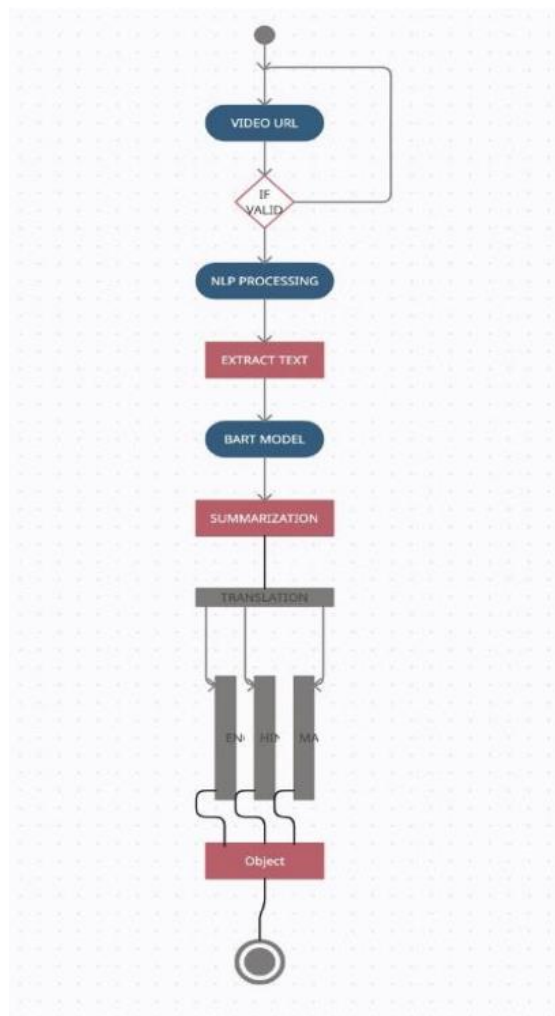


Fig 1. System Architecture

2.1 Getting URL

A reference to an online resource that gives the location of that resource on a computer network is known as a uniform resource locator, or URL. It functions as a tool for obtaining data from the internet. The system processes a user's entry of a YouTube URL in the search field. The URL in the background to get the data you need. Next, determine if the URL is legitimate or not. In the backend, the link shortening process will take place if the URL is legitimate. The needed YouTube video will then load from the provided link. It will handle the extraction of text from videos. When the user enters the URL, we will determine if it is legitimate or not. The URL will return if it is invalid.

2.2 Video to text Extraction

The URL from the NLP model file will be passed to the `utubeextract.py` file once the necessary YouTube video has been obtained. Next, we will obtain the transcript for the necessary YouTube video by utilizing the `youtube_transcript_api`. We are removing videos without transcripts from our collection. After utilizing an API to retrieve a video's transcript, we parse the text in preparation for further steps. This entails eliminating punctuation and commas, like as full stops and marks, are crucial for defining sentence boundaries. You can do this operation by utilizing the "punctuator" Python package. Applying the text preprocessing procedure for the extracted transcript is the next step.

Since commas, punctuation, and full stops are crucial for identifying sentence boundaries, they must be removed. You can do this operation by utilizing the "punctuator" Python package. The text will then be applied as our following step, preprocessing technique for transcript extraction. The aforementioned technique is similar to eliminating stopwords, exclamatory phrases, stemming, and punctuation for scanning reasons.

The mentioned task is to condense a lengthy phrase into a succinct overview. There will be crucial information in that synopsis from the video. The Extractive Summarization is the method most often used in NLP text summarization.

The summary provided by this extractive summarization only includes the crucial words and sentences. The website will always be checking the information on the page that the user has provided, and should the quantity mentioned change. A prompt alert

is transmitted to the user or client. Our website will extract the necessary data from the product's webpage as soon as the user enters the essential information. Our website uses the gorgeous soup module to scrape the data.

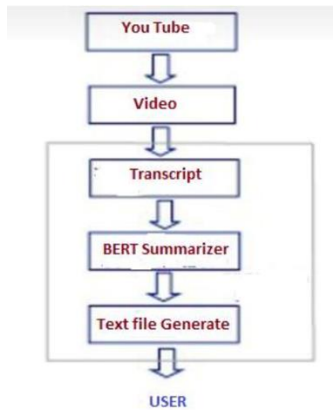


Fig 2. Text Extraction

2.3 NLP Processing

This procedure is for summarizing text. There are numerous methods for text summarizing in NLP. We will discuss text summarizing, which is the act of condensing a lengthy paragraph into a concise overview. If there are several lines in the paragraph, we will need additional time to cover the material; due to time constraints, we only need the text's core report. We are able to reduce the size of the lengthy text by eliminating unnecessary material. NLP Text Summarization is the process of dividing long texts into manageable paragraphs or sentences. The crucial data will be retrieved via the aforementioned procedure.

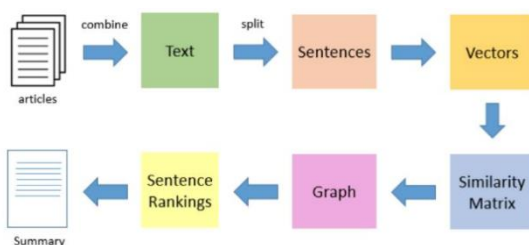


Fig 3.

NLP Processing

2.4 BERT Summarizer

Bidirectional Encoder Representations from Transformers, or BERT for short. This BERT presents a novel, cutting-edge method for completing NLP tasks. Among the crucial algorithms used in Natural

Language Processing Models is BERT. It is possible to remember the crucial information and using extractive text summarization to extract pertinent information. The Extractive Summarization presented here is more difficult. We are employing better embeddings in this advancement, which are supplied by encoder models such as BERT.

Using two supervised methods, construct an extractive summarizer using the BERT sentence embeddings. Only embeddings and their derivatives are taken into account in the first. We can infer from the article's internal structure that. A skilled summarizer can identify appropriate, meaningful sentences and parse meaning. The standard method is the Unsupervised Text-Rank model. In addition to these methodologies, fresh corpuses benefit from the sequential information and established advantages. In actuality, this tactic has been used by numerous publications. Lead serves as the second approach's baseline. The Rouge-1 and Rouge-L F1 metrics are crucial factors in the performance of supervised models.

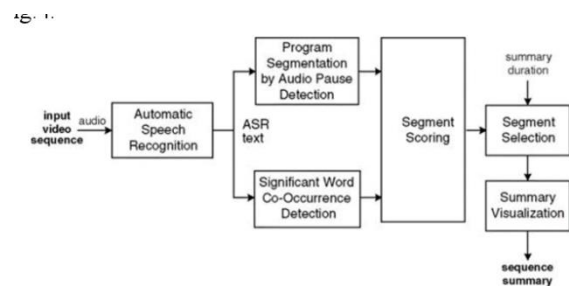


Fig 4. BERT Summarization Diagram

2.5 Language Translation

A language translator is a highly useful tool that converts our native tongue into the target language, enabling us to communicate in other languages. Speaking with each other used to be quite tough in the days before language translation apps with individuals traveling from all around the world. This allows us to use Python to develop our own language translation project. Our goal is to develop a language translator that will enable us to translate words, sentences, and even entire paragraphs across other languages. We will make an effort to include as many languages as we can. The Google Trans library will provide us with a variety of languages, and the T-kinter Module will be used to create the project's graphical user interface that are part of it.

3.FUTURE WORKS

We plan to continue working in the area of extensions in the future. Every browser, social media platform, and other video website (such as YouTube, Share-Chat, etc.) must support this extension. The user's main goal is to add the extension and choose the necessary footage .It will automatically download the URL and show the necessary condensed synopsis. browser, artificial intelligence, and machine learning technologies.

4.CONCLUSIONS

In conclusion, the user can save time by using our website. Instead of seeing the entire buffer waste content of the video, we will first view the YouTube video's core content to determine which one is best for us. The user's time and effort are saved. Through the use of their work in finding the ideal YouTube video will be lessened thanks to our website. Additionally, multilingual summarization and text-to-speech functionality are available on our website. We are sure that by saving consumers' time and effort, our paper will successfully fulfill their needs. Our method is to do away with the necessity for users to view lengthy videos by giving them access to only the pertinent and helpful information on the subjects that interest them. The time that is saved can be used for learning more and investigation.

REFERENCES

- 1 Alrumiah, S. S., Al-Shargabi, A. A. Educational Videos Subtitles' Summarization Using Latent Dirichlet Allocation and Length Enhancement. CMC-Computers, Materials & Continua, 70, 3 (2022).
- 2 Sangwoo Cho, Franck Démoncourt, Tim Ganter, Trung Bui, Nedim Lipka, Walter Chang, Hailin Jin, Jonathan Brandt, Hassan Foroosh, Fei Liu, "StreamHover: Livestream Transcript Summarization and Annotation", (2022).
- 3 S. Chopra, M. Auli, and A. M. Rush, "Abstractive sentence summarization with attentive recurrent neural networks," in Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics Hum. Lang. Technol., (2016).
- 4 Ghadage, Yogita H. and Sushama Shelke. "Speech to text conversion for multilingual languages." 2016 International Conference on Communication and Signal Processing (ICCSP) (2016).
- 5 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), 6, 9 (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).
- 7 Prof. S. A. Aher, Hajari Ashwini M, Hase Megha S, 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).
- 8 Aiswarya K R, " Automatic Multiple Language Subtitle Generation for Videos, " International Research Journal of Engineering and Technology (IRJET), 7, 5 (2020).
- 9 Savelieva, Alexandra & Au-Yeung, Bryan & Ramani, Vasanth. Abstractive Summarization of Spoken and Written Instructions with BERT (2020).
- 10 Patil, S. et al. "Multilingual Speech and Text Recognition and Translation using Image." International journal of engineering research and technology 5 (2016).
- 11 S. Sah, S. Kulhare, A. Gray, S. Venugopalan, E. Prud'Hommeaux and R. Ptucha, "Semantic Text Summarization of Long Videos," IEEE Winter Conference on Applications of Computer Vision (WACV), (2017).
- 12 A. Dilawari and M. U. G. Khan, "ASoVS: Abstractive Summarization of Video Sequences," in IEEE Access, 7 (2019).
- 13 Lin, Chin-Yew, "ROUGE: A Package for Automatic Evaluation of Summaries," In Proceedings of 2004, Association for Computational Linguistics, Barcelona, Spain.
- 14 A. E. B. Ajmal and R. P. Haroon, "Maximal marginal relevance based malayalam text summarization with successive thresholds," International Journal on Cybernetics and Informatics, 5, 2 (2016).
- 15 M. Allahyari et al., "Text summarization techniques: A brief survey," International Journal of Advanced Computer Science and Applications, 8, 10 (2017)