

## YOUTUBE TRANSCRIPT SUMMARIZER

**Tryphena Brighty B<sup>1</sup>, Pooja V V<sup>2</sup>, Jaya Varsini V S<sup>3</sup>, Vishnu Priya N R<sup>4</sup>**

<sup>1,2,3</sup>Students, Department of Computer Science and Engineering, Agni College of Engineering, Chennai, Tamil Nadu India.

<sup>4</sup>Assistant Professor, Department of Computer Science and Engineering, Agni College of Engineering, Chennai, Tamil Nadu India.

DOI: <https://www.doi.org/10.58257/IJPREMS31301>

### ABSTRACT

The YouTube Transcript Summarizer is a tool designed to save users time by providing them an easy and efficient way to get the gist of any YouTube video they are interested in. The system leverages the power of machine learning to analyse either the transcripts of videos or converts the voice into text and generates accurate abstractive summaries that capture the essence of the content. The system is designed to be user-friendly and accessible, with Chrome extension available for easy access.

KEYWORD: Accurate abstractive text summarization, Chrome extension, Machine Learning and NLP.

### 1. INTRODUCTION

After more than 17 years on the internet, YouTube is still going strong. It is the second most visited website and the second-most often used search engine. Every day, YouTube posts around 720,000 hours of new video content. The amount of videos that are accessible on the web platform is always increasing. It's become easier and easier to view videos on YouTube for everything, including motivational, dance, cuisine, and other strange things. The content is widely accessible and mainly used for educational purposes. In contrast to photos, where information can be gleaned from a single frame, extracting information from a video presents the most obstacle because the user must watch the full film in order to understand the context. The video will be fuzzy and difficult to watch if the viewer has a slow network connection or any other technological limitations. Additionally, the commercial breaks are too annoying.

Therefore, it is beneficial and time-effective to cut out the clutter at the beginning and end of the relevant video, skip the advertisements, and get the summary to jump right to the section you are interested in. The goal of this project is to make the writing for the videos shorter. Such movies can be summarised automatically, saving time and effort by avoiding the need to manually go through all of the content. This allows one to swiftly scan the video for significant patterns. The most crucial aspect of this project will be its ability to string together all the necessary information and concentrate it into a small paragraph. Finding the most important parts of a video and producing an output video with content that accurately summarises the full input video is the technique of "video summarization." One benefit is that the video takes up less storage space. This project will provide an opportunity to put cutting-edge NLP techniques for abstractive text summarization into practise while implementing an intriguing concept that is ideal for intermediates and a revitalising side project for pros.

### 2. LITERATURE SURVEY

YouTube Video Summarizer tool developed by the Allen Institute for Artificial Intelligence (AI2). This system uses a combination of automatic speech recognition (ASR) and NLP techniques to transcribe and summarize the audio content of YouTube videos. The system uses a deep neural network to generate a summary of the video's main points, and also provides a list of key phrases and a timeline of the video's content.

Another system is the YouTube Caption and Title Downloader developed by Kapwing. This system allows users to download the captions and titles of YouTube videos and then use a summarization tool like TextTeaser to create a summary of the content.

Lofindo: Lofindo is a YouTube transcript summarization tool that uses an algorithm to identify the most important sentences and phrases in a video. The system then generates a summary of the video's content using these key sentences.

XSum: XSum is a system developed by researchers at the University of Edinburgh that uses a combination of extractive and abstractive summarization techniques to summarize news articles and YouTube videos. The system uses a deep neural network to identify important sentences in the transcript and then generates a summary using these sentences.

### 3. OPPORTUNITIES

The goal of the YouTube Transcript Summarizer project is to develop a tool that uses machine learning-based summarising techniques to automatically produce summaries of YouTube videos.

This tool's objective is to save users time and effort when they are attempting to glean the most crucial information from a video.

The project can be helpful in a variety of contexts, including for students, researchers, and professionals who wish to remain current on the newest news or trends without investing a lot of time in watching films.

Overall, the research has the potential to speed up the process by which consumers can absorb and understand the information included in YouTube videos.

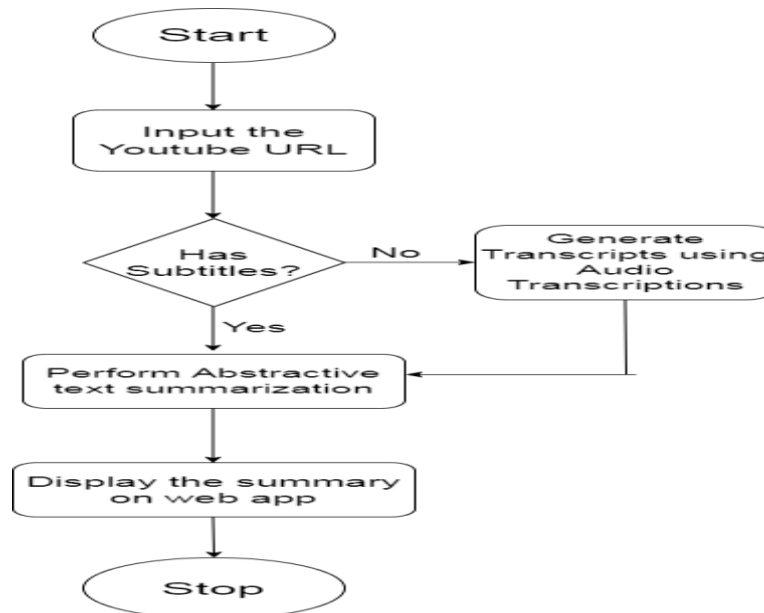
### 4. METHODOLOGIES

The main goal of this project is to give users a clear, succinct description of YouTube videos they shouldn't waste their time watching. This project makes use of well-known Python modules that have practical applications in various fields, including Flask, the YouTube transcript api, Transformers, Pipeline, and many others.

The project is divided into 3 modules :-

#### 1. Input the valid YouTube video URL

Obtaining the video link that the user wishes to summarise is the first step. The video must be recorded, have a current video id, and be accessible on YouTube. Before using this web application, users should bear the following in mind.



#### 2. Getting the transcripts from that video.

The next step is to obtain the video transcripts after getting the user's video URL. Now it will determine whether or not the provided video has available subtitles. If the provided video contains subtitles, we can easily extract the transcripts from it using the Python module known as YouTube\_Transcript\_Api. What if the given video doesn't contain subtitles? In this situation, we will first use the same Python api to transcribe the video's audio. This is how we obtain any video's transcription.

#### 3. Passing the Generated transcripts to the text summarizer.

Now that the project is in its crucial phase, everything depends on it. The text summarization is essentially included at this stage. Currently, there are primarily two categories of summarization methods:-

The text summarization approach known as "extractive summarization" basically takes the key phrases and sentences from the YouTube video, groups them together, and creates a brief description of the film. It won't result in any new sentences.

Abstractive summarising: This approach of text summarising is a fresh development in the field of art since it creates sentences from scratch. Essentially, it will copy the sentence, making it clearer and shorter than the original sentence.

### 5. CONCLUSION

Researchers have recently become quite interested in video summarization, and as a

result, many methods and techniques have been put forth. The goal of this project is to create a Chrome extension that can be used to summarise YouTube video content and extract crucial information from those patterns using cutting-edge NLP techniques for classification and abstractive text summarization.

## ACKNOWLEDGEMENT

We are thankful to Ms. N R Vishnu Priya,

CSE Department, Dr. L Sharmila, HOD-CSE, Agni College of Technology, OMR, Chennai for providing the support for this work.

## 6. REFERENCE

- [1] Hank Liao, Erik McDermott, Andrew Senior "Large scale deep neural network acoustic modeling with semisupervised training data for YouTube video transcription" December 2013.
- [2] Gaurav Sharma, Shaba Parveen Khan, Shivanshu Sharma, Syed Ubed Ali "Summarizer For Easy Video Assessment" Volume: 3 Issue: 04-April-2021
- [3] Atluri Naga, Laggiseti Valli, Jahnvi Duvru "Video Transcript Summarizer" Issue: 11 March 2022
- [4] Krishna Kulkarni, Rushikesh Padaki "Video Based Transcript Summarizer for Online Course using Natural Language Processing" Issue: 18 December 2021
- [5] Evalampos Apostolidis, Eleni Adamantidou, Vasileios Mezaris "Video Summarization Using Deep Neural Networks: A Survey" Pages 1838-1863 Issue: 13 November 2021
- [6] G. PRIYANKA, M. PRASHA MEENA "Survey and Evaluation on Video Summarization Techniques" Issue: 28 May 2020
- [7] Workie, R. Sharma, Y. N. Chung "Digital Video Summarization Techniques" Volume 09 Issue: 01 January 2020
- [8] Cüneyt M. Taskiran, Arnon Amir, Dulce B. Ponceleon, Edward J. Delp. "Automated Video Summarization Using Speech Transcripts".
- [9] Aniqua Dilawari, Muhammad usmanghani khan. "Abstractive Summarization of Video Sequences" IEEE Access, 2019.
- [10] Dilawari, M. Usman Khan, "ASoVS: Abstractive Summarization of Video Sequences" Pages 29253- 29263 Issue: 11 March 2019