

# Youtube Video Summary Generator

Ms. Sumalata Bandri, Mr. Abhishek Pandey, Mr. Bhushan Mahadule,

Mr. Om Satpute, Mr. Vaibhav Jawade

Department of Computer Science and Engineering,  
G H Raisonni University Amravati, Maharashtra, India

**Abstract** This project introduces the YouTube Video Transcribe Summarizer, a tool designed to automatically extract transcripts and generate concise summaries from YouTube videos. By leveraging the YouTube Transcript API, the system retrieves accurate video transcripts and utilizes Google Gemini Pro's advanced text-based model to create coherent summaries. Users can input a YouTube video URL, which displays the video thumbnail for context. The application features a customizable prompt template to tailor the summary generation process, ensuring relevance to individual needs. Built on a user-friendly Streamlit interface, this tool aims to enhance content accessibility and engagement. Additionally, the project explores the possibility of executing local models for improved performance and user control. By streamlining the summarization of video content, the YouTube Video Transcribe Summarizer facilitates more efficient information consumption, empowering users to navigate the vast landscape of online video more effectively.

**Index Terms-** Large Language Model(LLM), Artificial Intelligence(AI), General Pre-Trained Transforms, Natural Language Model(NLP), Application Programming Interface(API), Uniform Resource Locator(URL).

## I. INTRODUCTION

As digital media continues to proliferate, platforms like YouTube have become vital sources of information, entertainment, and education. With millions of videos available, users often struggle to extract meaningful insights without dedicating extensive time to watching entire clips. This challenge highlights the need for efficient tools that can distill content into digestible formats.

The YouTube Video Transcribe Summarizer aims to address this need by automating the extraction of transcripts and generating concise summaries from YouTube videos. Utilizing the YouTube Transcript API, this project retrieves accurate transcripts, which are then processed through advanced text-based models like Google Gemini Pro to create coherent and informative summaries. Built on a user-friendly Streamlit interface, the application allows users to easily input video URLs, providing immediate visual context through thumbnail displays.

Additionally, it offers customizable prompt templates, enabling users to tailor the summary output to their specific interests and preferences. By streamlining the process of summarizing video content, this tool not only enhances user engagement with digital media but also empowers individuals to navigate the vast array of available information more effectively. Ultimately, the YouTube Video Transcribe

Summarizer represents a significant step towards making online content more accessible and actionable for a diverse audience.

## II. LITERATURE REVIEW

Yixin Liu [1] Recent studies have found that summaries generated by large language models (LLMs) are favored by human annotators over the original reference summaries in commonly used summarization datasets. Therefore, we study an LLM-as-reference learning setting for smaller text summarization models to investigate whether their performance can be substantially improved. To this end, we use LLMs as both oracle summary generators for standard supervised fine-tuning and oracle summary evaluators for efficient contrastive learning that leverages the LLMs' supervision signals. Sulochana Devi [2] The paper goal is to design a user interface where the user can get the summary of the requested YouTube video using Natural Language Processing (NLP) and Machine Learning. Enormous number of videos are uploaded to YouTube daily. It has become difficult to find the relevant content that we are looking for, sometimes it may take longer than expected, and our efforts become futile if we are unable to extract meaningful information from it. Rand Abdulwahid Albeer [3] Automatic summarization is a technique for quickly introducing key information by abbreviating large sections of material. Summarization may apply to text and video with a different method to display the abstract of the subject. Natural language

processing is employed in automated text summarization in this research, which applies to YouTube videos by transcribing and applying the summary stages in this study. Based on the number of words and sentences in the text, the method term frequency-inverse document frequency (TF-IDF) was used to extract the important keywords for the summary. Some videos are long and boring or take more time to display the information that sometimes finds in a few minutes.

Md. Ahsan Habib [4] a new text summarization method has been proposed: combining the Extractive and Abstractive Text Summarization technique. In the extractive-based method, the model generates a summary using Sentence Ranking Algorithm and passes this generated summary through an abstractive method. When using the sentence ranking algorithm, after rearranging the sentences, the relationship between one sentence and another sentence is destroyed. To overcome this situation, Pronoun to Noun conversion has been proposed with the new system. Shengli Song [5] Abstractive Text Summarization (ATS), which is the task of constructing summary sentences by merging facts from different source sentences and condensing them into a shorter representation while preserving information content and overall meaning. It is very difficult and time consuming for human beings to manually summarize large documents of text. In this paper, we propose an LSTM-CNN based ATS framework (ATSDDL) that can construct new sentences by exploring more fine-grained fragments than sentences, namely, semantic phrases.

Deepa Anand [6] The availability of legal judgment documents in digital form offers numerous opportunities for information extraction and application. Automatic summarization of these legal texts is a crucial and a challenging task due to the unusual structure and high complexity of these documents. Previous approaches in this direction have relied on huge labelled datasets, using hand engineered features, leveraging on domain knowledge and focused their attention on a narrow sub-domain for increased effectiveness. In this paper, we propose simple generic techniques using neural network for the summarization task for Indian legal judgment documents.

### III PROPOSED WORK

Fig.1 represents the block diagram of the proposed YouTube Video Summary Generator workflow. The process starts with the user inputting a YouTube video link. The application then fetches the video. Next, it checks for transcript availability. If a transcript is available, the system generates a concise summary using Google Generative AI. Finally, the generated summary is displayed to the user. This tool aims to enhance efficient information processing for educators, students, and professionals, enabling them to quickly understand key video content without watching the entire video.

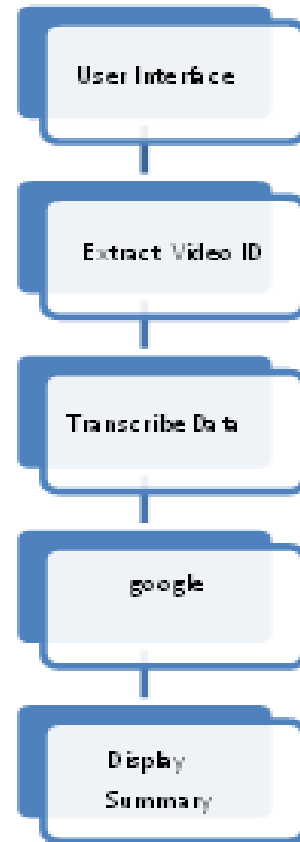


Figure.1: Block Diagram

Fig. 2 illustrates the proposed work for the YouTube Video Summary Generator. In this architecture, users input the URL of a YouTube video, which the system fetches and analyzes. The Google Gemini Pro API processes the video's content to extract key information. This extracted data is then passed to a Large Language Model (LLM), which generates a coherent summary of the video. The architecture encompasses components for video fetching, content extraction, and text generation, ensuring an efficient workflow from user input to the final summary output.

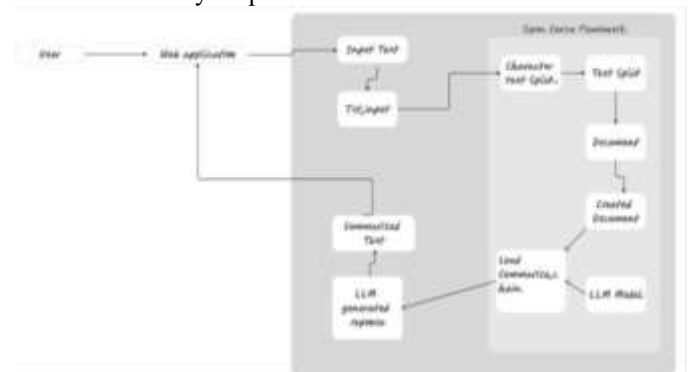


Figure.2 Architecture Diagram

## IV. RESULT & DISCUSSION

**Module 1: Accepting YouTube Links on Streamlit Framework**  
 In this module, we successfully integrated a feature that allows users to input YouTube links directly into the Streamlit interface.



Figure.3 Enter YouTube link

This enhancement streamlines the process of generating video summaries by eliminating the need for users to download and re-upload video files. Upon submitting a link, the application retrieves the video content and processes it through our summarization algorithm. This functionality significantly improves accessibility and convenience, enabling users to summarize a broader range of video content directly from the web.

**Module 2: Generating Summaries Using a Large Language Model**

This module focuses on generating concise and meaningful summaries through the utilization of a large language model in conjunction with the Google Gemini Pro API. After processing the video content, the model analyzes the data to extract key points and themes effectively. This integration enhances the quality of the summaries, providing users with valuable insights into the video content and improving their overall experience with the application.

## V. CONCLUSION & FUTURE SCOPE

The development of a YouTube video summary generator using the Streamlit framework and advanced generative AI technologies, such as Google Gemini Pro, showcases significant potential in enhancing content accessibility and user engagement. By leveraging API integrations and Python programming, the application effectively condenses lengthy video content into concise summaries, allowing users to quickly grasp key points without watching entire videos. This

not only saves time but also aids in content discovery, demonstrating the power of modern AI in transforming video consumption. Looking ahead, the project can expand with features like multilingual support, customizable summary options, enhanced user interaction through keyword highlighting or sentiment analysis, and integration with other platforms. Additionally, developing a mobile app version and utilizing machine learning techniques for continuous improvement could further elevate the user experience, making the generator a comprehensive tool for various demographics and platforms.

## REFERENCES

1. Xufeng He, Yang Hua, Tao Song, Zongpu Zhang. Unsupervised Video Summarization with Attentive Conditional Generative Adversarial Networks.(2019).
2. Tsu-Jui Fu, Shao-Heng Tai, Hwann-Tzong chen. Attentive and Adversarial Learning for Video Summarization (2019).
3. Li Yuan, Francis Eng Hock Tay, Ping Li, Jiashi Feng. Unsupervised Video Summarization with Adversarial LSTM Networks Behrooz Mahasseni, Michael Lam and Sinisa Todorovic Oregon State University Corvallis (2017).
4. Yogendra Sing, Rishu Kumar, Soumya Kabdal, Prashant Upadhyay. YouTube Video Summarizer using NLP: A Review, International Journal of Performability Engineering (Dec 2023).
5. Chrysa Collyda, Konstantinos Apostolidis, Evlampios Apostolidis. A Web Service for Video Summarization (June 2020).
6. Mrigank Rochan, Yang Wang. Video Summarization by Learning From Unpaired Data (2019).
7. Kuo-Hao Zeng, Tseng-Hung Chen, Juan Carlos Niebles, min sun. Title Generation for User Generated Videos. (Sep 2016)..
8. Rossi, F., & Martinez, M. (2021) on AI- Enhanced Space Robotics for Future Lunar Missions.