



You-Tube Video Abstractor Using NLP

Prof. Agalave K. N., Aquib Hannure, Mahesh Mutkiri, Vijay Shelake

B.E. Computer, S B Patil College of Engineering, Indapur, Maharashtra, India

ABSTRACT

The YouTube Video Abstractor program is dedicated to changing the way users interact with the rich video content available on platforms like YouTube. With video uploads increasing exponentially, the need to summarize quality content has never been greater. Our systems use the power of natural language processing (NLP) and machine learning to deliver comprehensive solutions. The basis of our system is the automatic extraction and analysis of video subtitles to create balanced, mixed and video content. Thanks to advanced language processing techniques, we detect important content and segments in videos and extract important content to improve search performance. This approach not only saves your visitors' valuable time, but also simplifies the content search process. It provides a way for people with disabilities to follow the content of the movie without having to watch it in its entirety. Content creators can benefit from insights into audience engagement that allow them to refine their content ideas. Additionally, researchers and teachers have useful tools to browse YouTube's vast library to aid in information search for study and business purpose.

Keywords: Summarization, Natural Language Processing, Summary, Time Saving,

I. INTRODUCTION

The YouTube Video Abstractor program is dedicated to changing the way users interact with the rich video content available on platforms like YouTube. With video uploads increasing exponentially, the need to summarize quality content has never been greater. Our systems use the power of natural language processing (NLP) and machine learning to deliver comprehensive solution.

The basis of our system is the automatic extraction and analysis of video subtitles to create balanced, mixed and video content. Thanks to advanced language processing techniques, we detect important content and segments in videos and extract important content to improve search performance. This approach not only saves your visitors' valuable time, but also simplifies the content search process. Our YouTube Video Snippets have great potential for different user groups. It provides a way for people with disabilities to follow the content of the movie without having to watch it in its entirety. Content creators can benefit from insights into audience engagement that allow them to refine their content ideas. Additionally, researchers and teachers have useful tools to browse YouTube's vast library to aid in information search for study and business purposes. It will tell you in simple words what is happening in the movie. So you can decide now whether you want to watch or not, save time. But it's not just for tourists. Filmmakers can also use our tools to understand what people like about their videos. Therefore, they can create higher quality videos that everyone will like. It works like an assistant for everyone who watches and plays videos, making YouTube more user-friendly and useful for all of us.

II. RELATED WORKS

The following sentences are taken from different research papers related to text summarization.

As per Devi, S., Nadar, R., Nichat, T., & Lucas, A, the author describes a method that uses Natural Language Processing (NLP) to generate summaries of requested YouTube videos. [1]

Vayadande, K., Nemade, M., Parbhanikar, S., Rathod, S., Raut, A., & Thorat, R discusses the challenge of summarizing the vast amount of video content available on platforms like YouTube. [2]

Albeer, R. A., Al-Shahad, H. F., Aleqabie, H. J., & Al-shakarchy, N. D. aims to address this challenge by developing an automated summarization method. [3]

Puranik, M. G. M., Kamath, N., Dusane, G., & Akhadkar, N. has the goal of generating text summaries of YouTube video transcripts using the Flask and Hugging Face libraries to Transformers for text summarization. This would allow users to quickly access the most important information from videos without having to watch them entirely. [4]

Verma, P., & Verma, A. seeks to improve the quality of text summarization through various techniques that address challenges such as redundancy, irrelevancy, loss of coverage, non-readability, and less cohesive content. [5]

Kupiec, J., Pedersen, J., & Chen, F. describes the problem statement is to efficiently summarize and categorize unstructured news content from various sources to enable users to quickly comprehend and access relevant information in a rapidly evolving digital news landscape. [6]

Yu, H., Yue, C., & Wang, C. describes a text summarization model for news articles that generates a one-sentence summarization in the style of a news title. A ranked list of accounts is revised by annotators to obtain a reliable gold-standard dataset. [7]

Ilampiray, P., Thilagavathy, A., Nithin, A. S., & Raj, I. aims to generate an effective summary for large YouTube videos. [8]

Alhojely, S., & Kalita, J. contains text summarization involves producing a document summary by highlighting its most important content ideas, as described in a tenth paper. [9]

Sahoo, A., & Nayak, A. K. discusses how summarization is a useful tool for quickly understanding collections of text documents and has many real-life applications. [10]

III. PROPOSED SYSTEM

Here the proposed system Create a system that generates concise YouTube video summaries based on captions to improve user content discovery and consumption time.

A. Block Diagram

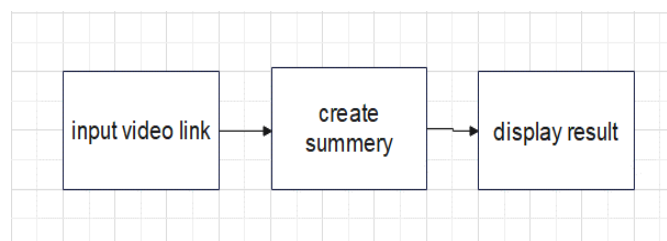


Figure 1: Block diagram

B. Mathematical Model

- Input: I= Youtube Video Link,

Where, Procedure(P)- P= I, using the I system get the video extract the caption, and generate a summary

- Output: O= Summary
- 1) Failures and Success conditions:
 - Failures: A huge caption can lead to more time consumption to get the summary. o Hardware failure. o Software failure.
 - Success: o YouTube videos have proper captions. o User gets results very fast according to their needs.
 - 2) Time Complexity:
 - The time complexity depends on the caption length. The time complexity is $O(n^2)$.
 - 3) Space Complexity:
 - Space complexity is $O(n^2)$.

C. Hardware and Software Requirements

- 1) Hardware Requirements:
 - Processor - intel i3/i5/i7 o Speed- 1.1 GHZ
 - RAM - 4 GB(min)
 - Hard Disk- 250 GB
- 2) Software Requirements:
 - Operating System - Windows o Browser - Chrome
 - Front End – HTML, CSS, JavaScript
 - Language – Python o IDE – collab

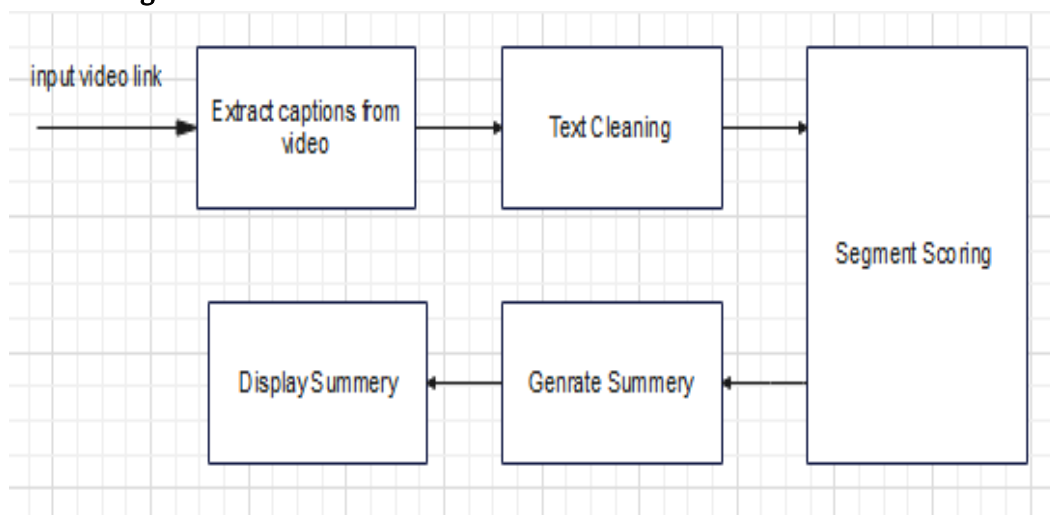
D. Architecture Diagram

Figure 2: Architecture Diagram.

IV.RESULTS

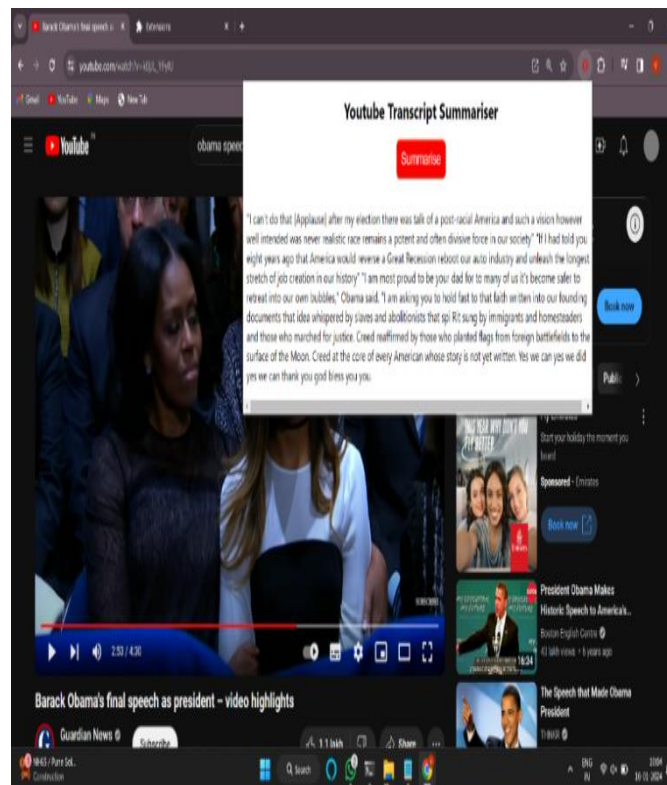
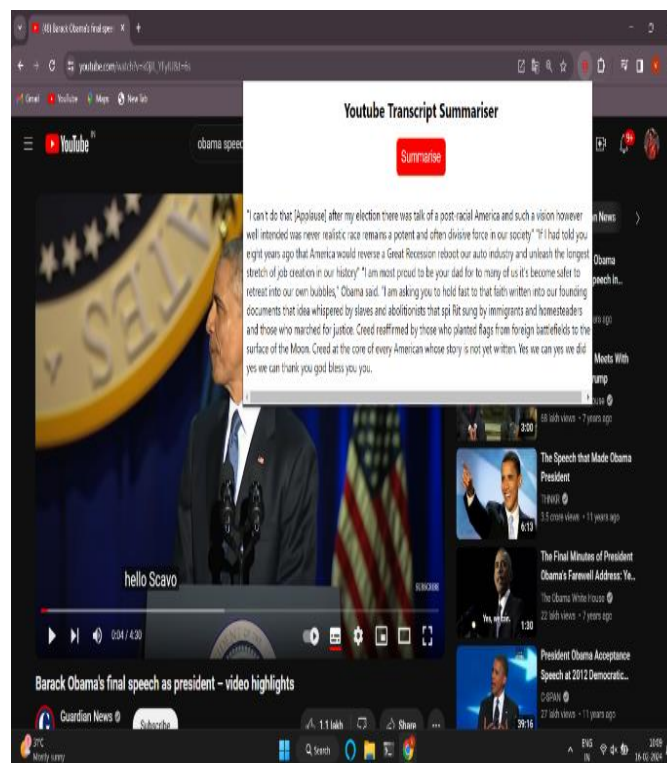


Figure 3:

Here speech of Obama on YouTube that we are going to Summarise the speech using our Youtube video Abstractor.



Here we are using the extension of our project You tube Video Abstractor.

V. CONCLUSION

Our project YouTube Video Abstractor project is a powerful tool that makes YouTube videos easier to understand and navigate. By looking at the video's captions, it figures out the main points and keywords, so you can quickly decide if a video is worth watching. This is a big time-saver. Our tool works for all kinds of videos and languages, so it's useful to people all around the world. Plus, you can adjust how long or short you want the summary to be, which gives you control. We're always looking to make it better based on what you and other users tell us. We also take your privacy and security very seriously. You can access our tool on the web, and it's designed to be user-friendly. In a nutshell, our Caption-Based YouTube Video Abstractor simplifies the process of finding and understanding YouTube videos, making it easier for you to get the content you want. We're committed to improving it and helping people make the most of online video content.

VI. REFERENCES

- [1]. Devi, S., Nadar, R., Nichat, T., & Lucas, A. (2023, May). Abstractive Summarizer for YouTube Videos. In International Conference on Applications of Machine Intelligence and Data Analytics (ICAMIDA 2022) (pp. 431-438). Atlantis Press.
- [2]. Vayadande, K., Nemade, M., Parbhanikar, S., Rathod, S., Raut, A., & Thorat, R. (2023, November). Efficient Content Exploration on YouTube: Automatic Speech Recognition-Based Video Summarization. In 2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1056-1062). IEEE.
- [3]. Albeer, R. A., Al-Shahad, H. F., Aleqabie, H. J., & Al-shakarchy, N. D. (2022). Automatic summarization of YouTube video transcription text using term frequency-inverse document frequency. Indonesian Journal of Electrical Engineering and Computer Science, 26(3), 1512-1519.
- [4]. Puranik, M. G. M., Kamath, N., Dusane, G., & Akhadkar, N. (2018). YOUTUBE VIDEO SUMMARIZER: A WEB BASED APPLICATION FOR CONCISE VISUAL AND TEXTUAL SUMMARY. Artificial Intelligence.
- [5]. Verma, P., & Verma, A. (2020). A review on text summarization techniques. Journal of scientific research, 64(1), 251-257.
- [6]. Kupiec, J., Pedersen, J., & Chen, F. (1995, July). A trainable document summarizer. In Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval (pp. 68-73).
- [7]. Yu, H., Yue, C., & Wang, C. (2017). News article summarization with attention-based deep recurrent neural networks. Stanford University [26] Young, T., Hazarika, D. Poria, S., Cambria, E., "Recent Trends in Deep Learning Based Natural Language Processing.
- [8]. Ilampiray, P., Thilagavathy, A., Nithin, A. S., & Raj, I. (2023). Video Transcript Summarizer. In E3S Web of Conferences (Vol. 399, p. 04015). EDP Sciences.
- [9]. Alhojely, S., & Kalita, J. (2020, December). Recent progress on text summarization. In 2020 International Conference on Computational Science and Computational Intelligence (CSCI) (pp. 1503-1509). IEEE.
- [10]. Sahoo, A., & Nayak, A. K. (2018). Review Paper on Extractive Text Summarization. International Journal of Engineering Research in Computer Science and Engineering (IJERCSE), 5(4).

- [11]. Aaglave, K. N., Shivanjali Santosh Jadhav, Amaan Firoj Khatib, and Rohini Laxman Khurangale. "A Survey on the Web Scraping: In the Search of Data." (2023).
- [12]. Karve, S. M, Kakad , S. ., Swapnaja Amol, Gavali, A. B. ., Gavali , S. B. ., & Shirkande, S. T. . (2024). An Identification and Analysis of Harmful URLs through the Application of Machine Learning Techniques. *International Journal of Intelligent Systems and Applications in Engineering*, 12(17s), 456–468. Retrieved from <https://www.ijisae.org/index.php/IJISAE/article/view/4905>
- [13]. Parlewar, P, Jagtap, V, Pujeri, U., Kulkarni, M. M. S, Shirkande, S. T., & Tripathi, A.(2023). An Efficient Low-Loss Data Transmission Model for Noisy Networks. *International Journal of Intelligent Systems and Applications in Engineering*, 11(9s), 267–276.
- [14]. Nalawade, V. S., Ashok, G. K., Hanumant, B. A., & Reshma, G. (2021). ENCRYPTION THEN COMPRESSION BASED SYSTEM USING GRAYSCALE BASED IMAGE ENCRYPTION FOR JPEG IMAGES.