International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 11 | Nov - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Youtube Transcript Summerizer

Prof. Shivani Karhale[1], Isha Temkar[2], Disha Salve[3] , Shubham Giri[4] , Sakshi Mane[5]

[1]Prof. Shivani Karhale, Information Technology, P.G.Moze College Of Engineering
[2]Isha Temkar, Information Technology, P.G.Moze College Of Engineering
[3]Disha Salve, Information Technology, P.G.Moze College Of Engineering
[4]Shubham Giri, Information Technology, P.G.Moze College Of Engineering
[5]Sakshi Mane, Information Technology, P.G.Moze College Of Engineering

---

Abstract - In the modern world, a video transcript summarizer offers a wide range of applications. It draws attention to the crucial points in the video. YouTube videos are watched in large quantities by people for a variety of reasons, including enjoyment, education, obtaining crucial information, and pursuing personal hobbies. Finding a video with any meaningful information about a subject is incredibly tough because most videos are just buffers for unimportant information. The main goal in the majority of situations is to get useful information from the video. In order to generalize YouTube video transcripts for abstractive text summary without sacrificing the essential details and substance, this research has developed a video summarization system based on Natural Language Processing and Machine Learning. The goal of this project is to make the writing for the videos shorter.

***Key Words*:** NLP, Machine Learning, Abstractive Text Summarization.

1.INTRODUCTION

In today's fast-paced world, consuming information efficiently is crucial. With the exponential growth of online video content, extracting meaningful insights from these videos can be time-consuming and overwhelming. That's where our project comes in.
In today's fast-paced world, consuming information efficiently is crucial. With the exponential growth of online video content, extracting meaningful insights from these videos can be time-consuming and overwhelming. That's where our project comes in.

Our project aims to develop a sophisticated tool that automatically summarizes YouTube video transcripts, making it easier for users to understand and retain the key points without having to watch the entire video. By leveraging advanced natural language processing (NLP) techniques, this tool extracts the most important information from video transcripts and presents it in a concise and easily digestible format.

It helps users save time by providing quick access to essential content, making it ideal for busy professionals, students, and anyone who wants to stay informed without watching lengthy videos.
For educational and training videos, a summary can reinforce learning and provide quick revision material.

The use of YouTube transcript summarizers has gained attention from researchers in recent years due to the increasing amount of video content available on the platform. This section presents a literature survey of some of the previous works on YouTube transcript summarizers.
'Automated Video Summarization Using Speech Transcript' by Cuneyt M. Task Iran, Aronon Amir, Dulce B. Ponce Leon, Edward J. Delph describes the compact representations of video data can enable efficient video browsing. They propose the method which summarizes the long video automatically. Their representations provide the user relevant information about the content with particular sequence examined while preserving the essentials of the content.

2.LITERATURE SURVEY

Bhandare (2022) This paper proposes an automatic videotape summarization algorithm using NLP grounded algorithms. With an increase in internet vids on the videotape depository platforms like YouTube. there's an increase in demand for good summarization algorithms to epitomize colorful vids. This paper aims to produce short videotape summary that summarizes colorful YouTube vids. The proposed fashion first summarizes the YouTube videotape reiterations grounded on which epitomized videotape is generated. A web operation that takes input as a YouTube videotape link and the needed summary duration from the stoner is also developed. Epitomized videotape affair is generated and displayed on the web runner, after successful processing.

Prasad (2023) YouTube is a massive platform that hosts a vast amount of video content. However, finding relevant information from these videos can be time-consuming and challenging, especially when one wants to understand the key points of a video quickly. This problem can be addressed by automatically summarizing the transcript of a video into a concise and informative summary. The TF-IDF (Term

Frequency-Inverse Document Frequency) algorithm is used to summarize the transcript of YouTube videos. The TF-IDF algorithm is a popular information retrieval technique that measures the importance of a word in a document. The algorithm calculates the term frequency (TF) of each word in a transcript and measures the inverse document frequency (IDF) of the words across a large corpus of documents. The TF-IDF algorithm is applied to the transcript of each video to determine the most important words in the transcript. The summary is then generated by selecting a subset of the most important sentences that contain these important words. The generated summary effectively condenses the transcript into a concise and informative summary that can be quickly consumed. The results of the project indicate that the TF-IDF algorithm is an effective approach for summarizing the transcript of YouTube videos. The generated summaries accurately capture the key points of the video, making it easier for viewers to quickly understand the content of the video. This approach can be useful for a variety of applications, including content discovery, information retrieval, and education.

Inamdar (2023) The goal of this project is to construct a chrome extension that will send a request to a backend REST API, conduct NLP, and return a summary version of a YouTube transcript to improve the surfing experience without distracting from it. This procedure combines both transcript production and text summarizing. To generate the transcript, first convert the video to audio using the PyTube package, which extracts audio in mp3 format. A toolkit called hugging sound is used for text creation from audio. Spacy, an NLP library, is used for text summarizing. Moreover, an API based on Flask is being developed to allow users to communicate with the backend, as well as a Chrome extension to give a user-friendly browsing experience.

3.METHODOLOGY

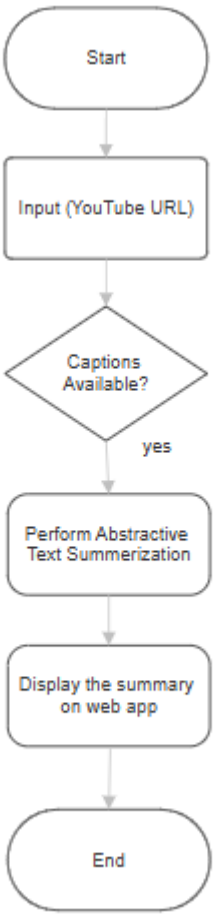
A. Machine Learning

Without explicit programming, a subset of artificial intelligence known as "machine learning" enables machines to learn from experience and improve over time. It is a method of teaching computers to spot patterns in data and draw conclusions from those patterns. In other words, machine learning algorithms analyze data using statistical models and then forecast or decide based on their findings. Reinforcement learning, unsupervised learning, and supervised learning are only a few of the several types of machine learning. In supervised learning, the computer uses labelled data to make predictions about the results of fresh, uncontaminated data. Contrarily, when using unlabeled data in unsupervised learning, the computer must figure out patterns and structure on its own. The computer learns by doing things and receiving feedback or rewards when employing reinforcement learning for decision- making.

B. Natural Language Processing (NLP)

"Natural language processing" (NLP), a subfield of artificial intelligence (AI), investigates how computers and human language interact. With NLP approaches, computers are now able to process, decipher, and analyze huge amounts of text, voice, and other types of written or drawn human language data. Language translation, sentiment analysis, Chatbots, speech recognition, and other uses for NLP are just a few examples. The processing of human language by computers can be a challenging task because natural language is often ambiguous, has multiple meanings, and varies based on context. However, with the advancement of technology and NLP and machine learning have advanced significantly in recent years. The core components of NLP include tokenization, stemming, stop word removal, part-of-speech tagging, named entity recognition, and sentiment analysis.

FLOW CHART



4.IMPLIMENTATION

This project makes use of well-known Python libraries that have practical applications, such as Flask, YouTube Transcript API, the Hugging- Face library of Transformers, the Speech Recognition API, the Google Translate API. There are three steps to perform the project:

1. Enter the user's legitimate YouTube video URL: Obtaining the video link that the user wishes to describe is the first step. The video must have a current video id and be accessible on YouTube. Before using this web application, users should bear the following in mind.

2. Obtaining the video's transcripts: The next step is to obtain the video transcripts after getting the user's video URL. Now it will determine whether the provided video has available subtitles or not. If the provided video contains subtitles, we can easily extract the transcripts from it using the Python module known as YouTube Transcript API. If the given video doesn't have subtitles, it doesn't provide output and occurs an error.

A. Extractive summarization

Extractive summarization is a technique in natural language processing (NLP) that involves extracting the most important sentences or phrases from a document or a set of documents to create a condensed version that captures the essence of the original text. It is called "extractive" because it involves selecting and extracting text directly from the source document, rather than generating new text. The process of extractive summarization typically involves several steps, including: 1. Text pre-processing: This involves removing stop words, punctuation, and other unnecessary elements from the text, as well as tokenization and sentence segmentation to break the text into smaller, more manageable chunks. 2. Sentence or phrase scoring: This entails giving each sentence or phrase in the text a score or weight based on how important it is to understand the document's overall meaning. This can be accomplished using a variety of methods, such as machine learning algorithms, semantic analysis, and keyword frequency analysis. 3.Sentence or phrase selection: To do this, choose the sentences or phrases thatreceived the highest scores and put them together into a summary that includes the most crucial details from the original text.

B. Abstractive Summarization:

Abstractive summarization is a natural language processing (NLP) technique that generates a summary of a given text by understanding its meaning and generating a new summary that may not contain the same exact words or phrases as the original text. Unlike extractive summarization, which selects sentences or phrases directly from the input text to form a summary, abstractive summarization involves creating a summary that is more concise and fluent, while retaining the essential meaning and information. Abstractive summarization is a more challenging task for NLP systems, as it requires the ability to understand the nuances of language and generate natural-sounding text. Abstractive summarization systems use techniques such as neural networks, natural language generation (NLG), and language models such as BERT and GPT to analyze the input text, identify the most important information, and generate a summary in a way that sounds like it was written by a human.

ADVANTAGES

1. Timesaving: NLP-based YouTube transcript summarizers can automatically process and summarize lengthy video transcripts. This saves time for users who want to quickly extract the key information from a video without watching or reading the entire transcript.

2. Accessibility: Summarizers make video content more accessible to individuals with hearing impairments or those who prefer reading over watching videos. The summarized transcripts provide a concise overview of the video's content, making it easier for a broader audience to engage with the material.

3. Efficient Content Consumption: With the help of summarizers, users can quickly scan through multiple video summaries to identify relevant information or topics of interest. This is particularly useful for researchers, students, or professionals who need to review a large amount of video content efficiently.

4. Multi-Language Support: NLP allows for the development of transcript summarizers that can process and summarize video transcripts in different languages. This expands the accessibility and usability of the summarization tool for users worldwide.

5. Scalability: NLP techniques enable the automation of the summarization process, making it possible to handle large volumes of video content. This scalability is beneficial for content creators, media organizations, and platforms like YouTube, as it facilitates the creation of concise summaries for numerous videos.

DISADVANTAGES

1. Loss of Contextual Information: Summarization algorithms may struggle to capture the full context and nuances of a video in a concise summary. Important details or subtle cues present in the audiovisual content may be overlooked, leading to a potential loss of information or misinterpretation.

2. Ambiguity and Inaccuracies: NLP-based summarizers can encounter challenges in disambiguating certain phrases, expressions, or specialized jargon used in the video. This can result in inaccurate or misleading summaries if the algorithms fail to grasp the intended meaning correctly.

3. Dependency on Transcript Quality: The accuracy and effectiveness of transcript summarizers heavily rely on the

quality of the video's automated transcript. If the transcript contains errors, inaccuracies, or omissions, it can negatively impact the quality of the generated summary.

4. Lack of Subjectivity and Creativity: NLP algorithms are generally focused on extracting information and may struggle to capture subjective opinions, emotions, or creative elements presented in a video. As a result, the summarized transcripts may lack the human touch or subjective insights.

5. Limited Summarization Length: Summarization algorithms often generate concise summaries that are restricted to a specific length or word limit. This limitation can be a challenge when summarizing complex or lengthy videos, as important details may be truncated or omitted. It's worth noting that the performance and accuracy of YouTube transcript summarizers using NLP can vary depending on the specific algorithms, training data, and advancements in the field. Ongoing research and development in NLP aim to address some of these limitations and improve the overall quality of automated summarization systems.

5. CONCLUSIONS

On the Internet, more and more video recordings are produced and distributed. It has become quite challenging to dedicate time to watching them that might run longer than anticipated, and if we are unable to glean any meaningful information from them, our efforts might be in useless. These videos can be summarized automatically, saving us time and effort by allowing us to quickly identify the most important information without having to watch the full thing. Recently, there has been a lot of interest among researchers in video summarizing, leading to the development of numerous methodologies and strategies. This project seeks to develop a web application or Chrome extension that can be used to summarize YouTube video content and pull relevant data from those using cutting-edge NLP techniques for categorization and abstractive text summarization. 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 web application that can be used to summarize YouTube video footage and extract crucial information from those patterns using cutting- edge NLP techniques for categorization and abstractive text summary.

To improve the performance of this project, here are some suggestions:
1. Data Preprocessing
2. Feature Engineering
3. Model Selection

6.REFERENCES

[1] Hank Liao, Erik McDermott, and Andrew Senior "Large scale deep neural network acoustic modelling with semisupervised training data for YouTube video transcription" January 2014

[2] "Summarizer for Easy Video Assessment" by Gaurav Sharma, Shaba Parveen Khan, Shivanshu Sharma, and Syed Ubed Ali. Vol. 3, No. 4, April 2021

[3] Rushikesh Padaki and Krishna Kulkarni's "Video Based Transcript Summarizer for Online Courses Using Natural Language Processing" Problem: December 18, 2021
[4] "Video Summarization Using Deep Neural Networks: A Survey" by EvalampoisApostolidis, Eleni Adamantidou, and Vasileios Mezaris, pages 1838–1863. Problem: November 13, 2021

[5] M. PRASHA MEENA and G. PRIYANKA Issue: "Survey and Evaluation on Video Summarization Techniques" Date: 28 May 2020

[6] Y. N. Chung, R. Sharma, and A. Workie Volume 09 Issue: January 2020 of "Digital Video Summarization Techniques"

[7] A. Dilawari and M. Usman Khan, "ASoVS: Abstractive Summarization of Video Sequences" Pages 29253–29263 Issue: 11 March 2019

[8] "AutomatedVideo Summarization Using Speech Transcripts" by Cüneyt M. Taskiran, Arnon Amir, Dulce B. Ponceleon, and Edward J. Delp.

[9] Abstract summarization of video sequences, AniquaDilawari and Muhammad Usmanghani Khan, IEEE Access, 2019.

[10]https://developer.chrome.com/docs/ext ensions/mv2/

[11]https://developer.mozilla.org/enUS/docs/Mozilla/Addons/ WebExtensions/manifest.json/page action

[12]https://developer.mozilla.org/enUS/docs/Web/API/XML HttpRequest/User_XM LHttpRequest

[13] Developer Chrome Documents Extensions Mv2 Messaging

[14] I. Awasthi, K. Gupta, P. S. Bhogal, S. S. Anand and P. K. Soni, "Natural Language Processing (NLP) based Text Summarization - A Survey," 2021 6th International Conference on Inventive Computation Technologies (ICICT), 2021, pp. 1310-1317, doi: 10.1109/ICICT50816.2021.9358703.