



MULTILINGUAL YOUTUBE TRANSCRIPT SUMMARIZER

Anusuya A¹, Monika R², Maheswari M³, Dr. Roselin Mary S⁴

Student, Computer science and engineering, Anand Institute of Higher Technology, Chennai, India¹

Student, Computer Science and Engineering, Anand Institute of Higher Technology, Chennai, India²

Assistant Professor, Computer Science and Engineering, Anand Institute of Higher Technology, Chennai, India³

Head of Department, Computer Science and Engineering, Anand Institute of Higher Technology, Chennai, India⁴

Abstract: The Multilingual YouTube Transcript Summarizer project utilizes machine learning and natural language processing techniques to provide an automated summary of the transcripts of YouTube videos. The project involves extracting the text from the video transcripts, analyzing the language patterns, and using algorithms to generate a concise summary of the content. This technology can assist users in quickly identifying key points and important information within lengthy videos. The project utilizes machine learning techniques to continually improve the quality and accuracy of the summaries, ensuring that they are relevant and useful for viewers. Natural language processing is also utilized to identify important keywords and phrases within the text and to analyze the tone and sentiment of the content. The YouTube Transcript Summarizer project represents an exciting advancement in the field of artificial intelligence, offering a powerful tool for anyone looking to save time and streamline their video-watching experience.

Keywords: Natural language processing (NLP), Machine learning, Text mining, Language detection, Keyword extraction, Speech recognition, Text classification

I. INTRODUCTION

Every day, a enormous quantity of video recordings are made and allotted on-line. nearly 2.three billion people have been energetic YouTube customers in 2020, and that wide variety has been growing quick. On common, three hundred hours of video are published to YouTube each minute. consistent with a Google survey, over 48 hours are spent consistent with month at the internet site by way of over one-third of YouTube visitors in India who access films on their cellular devices. It has come to be quite difficult to spend time watching those films, which can be longer than expected. If we are unable to find the records, we are seeking out in them, our efforts can also once in a while be in useless. locating films that incorporate the information we are honestly seeking out may be time-consuming and irritating. as an example, there are various motion pictures on line wherein the speaker discusses a certain challenge in-depth, but unless we watch the entire video, it may be difficult to parent the message the speaker is trying to recover from to the audience. there are many beneficial applications for Python. currently, the Python library's API makes it less difficult to access YouTube fabric along with video transcripts and other stuff. With this benefit, we may attain the video transcripts immediately and summarise them for the viewer to view.

Hugging face transformer, one of the textual content summarising techniques may be used to do that. Chrome extensions may be used to enhance the consumer interface for the capability segment. A summarise button in this Chrome extension displays the text of the currently playing YouTube video on the Google Chrome internet browser whilst clicked. The embracing face transformer bundle created the textual content that has been condensed right here. normally, thumbnails and manual descriptions are used to provide a precis of the YouTube motion pictures. the second one-maximum considered website in the international is YouTube.

The style of movies to be had on YouTube includes vlogs, stay streaming, song movies, feature movies, documentaries, audio recordings, and trailers for films which have been backed with the aid of businesses. every day, a couple of billion hours of video are watched on YouTube. To provide an insightful and pertinent summary of the movie, this challenge shows using a transformer package to summarise the video's transcripts. T5 is an encoder-decoder model for which every undertaking is translated into a text-to-textual content layout after it's been pre-trained on a set of supervised and unsupervised duties. We rent a pre-educated summarization technique due to the fact that is what we care approximately maximum, summarizing the statistics.



II. RELATED WORKS

Ishitwa Awasthi, Kuntal Gupta, Prajbot Singh Bhojal, Anand, and Piyush Kumar published textual content Summarization - A examine in herbal Language Processing (NLP) in 2021. Extract and abstract methods are used for text summarization. benefits: The results of the sentence in linguistics and decided. with the aid of the use of statistical homes. The disadvantage is that each summary approach has unique contexts which can be beneficial. It isn't always possible to determine which technique is extra effective [1]. Adhika Pramita, Supriyadi Rustad, Abdul Shukur and Effendi created "A review of automatic textual content Summarization techniques and methods". In 2020, it turned into published. text summarization and systematic evaluation strategies have been used. Fuzzy-primarily based techniques have the drawback of being inadequate for semantic issues. Many drawbacks in extraction techniques need to be closed [2]. Parth Rajesh Dedhia, Hardik Pradeep and Meghana Naik created "Survey of abstract textual content Summarization strategies". published in 2020. Seq2Seq, encoder-decoder and indicator mechanism are used. the standard is that it's going to no longer paintings if the version is given too many files [3]. Anika Dilawari and Muhammad Usman Ghani Khan are the creators of "Video series summary Compilation". They used RCNN deep neural network model and multi-line video description. the error indicates how correct the precis is. Time and reminiscence constraints are neglected[4].

III. EXISTING SYSTEM

Maximum of the prevailing YouTube transcript summarizer is completely based on the technique that entails pulling transcripts from the person-furnished video link with the employment of Pytube, Ffmpeg, Speech recognition model, Librosa, Hugging Sound, and Hugging Face Transformers, and pipelining the effects.

This approach makes the whole process of obtaining the transcript of a video and summarizing to be more complicated as it downloads the video in mp4 layout and converts the mp4 into mp3 format further extracting the transcripts from the audio and summarizing with the hugging face transformers.

IV. PROPOSED SYSTEM

Right here in this undertaking, we will be decreasing the complexity and time consumption by way of without delay the usage of the YouTube transcript API which has the potential the accumulate transcript of a YouTube video within seconds and we can additionally be growing the accuracy of the existing works by using additionally using professional summarization models together with Pegasus and Bart.

An added feature includes multilingual summarization using the HiBart and mT5 fashions. by means of including this feature, time consistency can be decreased to folks that will study the transcript summary in place of studying the real whole transcript summary and they are able to read the summary of their preferred languages.



Fig. 1 System Architecture Diagram



V. IMPLEMENTATION

This system is split into four sections. importing the YouTube transcript API is the first module and we are able to be uploading the YouTube transcript API with a view to accumulate the transcripts of the video as consistent with the consumer's requirements.

Then we can be transferring on to the second one section wherein we are able to be getting the video's transcript based totally on the consumer's desired language. after you have the transcripts primarily based at the user's desired language, we are able to be loading the summarizing models through developing checkpoints for every version and sooner or later, we are able to be summarizing the transcripts in the very last phase.

the first module of the device entails uploading the YouTube transcript API, which lets in the machine to obtain the transcripts of motion pictures based totally on consumer necessities.this could involve specifying the video's URL or identity, or looking for films primarily based on certain keywords or standards. once the API has been imported, the device can start retrieving the transcripts of the movies that the person is interested in.

the second one segment of the system specializes in acquiring the video's transcript inside the consumer's desired language. this may involve the use of machine translation to transform the transcript from one language to every other, or it may contain retrieving transcripts that have already been translated by means of YouTube or different resources. The intention of this section is to ensure that the person can recognize the video's content despite the fact that it is in a language they do now not communicate.

The 1/3 phase of the machine entails loading summarizing fashions, which can be algorithms or packages that could condense massive quantities of text into shorter summaries. The system creates checkpoints for every summarizing model, which lets in it to music the progress of the summarization method and make certain that it is accurate and dependable. This section is important because it permits the machine to generate summaries which might be concise and informative, that can shop the person time and effort.

The very last phase of the device entails summarizing the transcripts obtained inside the preceding sections. The gadget makes use of the loaded summarizing fashions to create summaries of the transcripts, which may be supplied to the person in a diffusion of codecs (e.g., text, audio, video). these summaries are designed to provide the user with a short and clean way to apprehend the main factors of the video, while not having to watch or examine the entire transcript. ordinary, the system is designed to make it simpler for customers to get admission to and recognize video content material, irrespective of the language it's far in or the amount of time they have to be had.

RESULTS:

ENGLISH SUMMARY:

```
[ ]: pegasus = summarize(tokenizer1, model1, transcript_en)
```

```
[ ]: pegasus[0]
```

```
[ ]: bart = summarize(tokenizer3, model3, transcript_en)
```

```
[ ]: bart[0]
```



ENGLISH SUMMARY OUTPUT:

```
[19] pegasus = summarize(tokenizer1, model1, transcript_en)
```

```
/usr/local/lib/python3.9/dist-packages/transformers/generation/utils.py:1288: UserWarning: Using `max_length`'s default (256) to control the generation length. This behaviour is deprecated. warnings.warn(
```

```
[20] pegasus[0]
```

```
'In an effort to create a culture within my classroom where students feel safe sharing the intimacies of their own silences, I have four core principles posted on the board that sits in the front of my class, which every student signs at the beginning of the year: read, write consciously, speak clearly, tell your truth. I spent so much of my life telling people the things they wanted to hear instead of the things they needed to hear, told myself I wasn't meant to be anyone's conscience because I still had to figure out being my own, so sometimes I just wouldn't say anything, appeasing ignorance with my silence, unaware that validation doesn't need words to endorse its existence.'
```

```
[21] bart = summarize(tokenizer3, model3, transcript_en)
```

```
/usr/local/lib/python3.9/dist-packages/transformers/generation/utils.py:1288: UserWarning: Using `max_length`'s default (142) to control the generation length. This behaviour is deprecated. warnings.warn(
```

```
[22] bart[0]
```

```
' Dr. Martin Luther King, Jr., in a 1968 speech, states, "In the end, we will remember not the words of our enemies but the silence of our friends" "Silence is the residue of fear. It is the air retreating from your chest because it doesn't feel safe in your lungs"'
```

TAMIL SUMMARY:

```
[ ]: mt5 = summarize(tokenizer2, model2, transcript_ta)
```

```
[ ]: mt5[0]
```

```
[ ]: HiBart = summarize(tokenizer4, model4, transcript_ta)
```

```
[ ]: HiBart[0]
```



TAMIL SUMMARY OUTPUT:

```

mt5 = summarize(tokenizer2, model2, transcript_ta)

/usr/local/lib/python3.9/dist-packages/transformers/generation/utils.py:1288: UserWarning: Using `max_length`'s default (84) to control the generation length. This behaviour is deprecated.
warnings.warn(

[24] mt5[0]

'mெளம் என்பது ஒரு அவமானம். ஆனால், அது ஒரு கொடுமான செயல்.'

[25] HiBart = summarize(tokenizer4, model4, transcript_ta)

/usr/local/lib/python3.9/dist-packages/transformers/generation/utils.py:1288: UserWarning: Using `max_length`'s default (20) to control the generation length. This behaviour is deprecated.
warnings.warn(

[26] HiBart[0]

'க மாக்ஃப்பாராக்:'

```

V. RESULTS AND DISCUSSION

Multilingual YouTube transcript summarization includes generating a concise and informative precis of a YouTube video's transcript in more than one languages. This project is challenging because of the diversity of languages, accents, and dialects used in YouTube movies, as well as the need to accurately seize the primary factors and ideas expressed within the video. latest studies in this area has targeted on developing machine studying and natural language processing techniques to mechanically summarize video transcripts in a couple of languages.

One of the strategies used consist of:

Multilingual pre-skilled language fashions consisting of BERT which uses a transformer architecture with a multi-layer encoder whereas BART the one which we used on this venture makes use of a transformer structure with a multi-layer encoder-decoder.

VI. CONCLUSION

This mission has supplied the perfect way to summarize a YouTube video with its transcript and the pleasant part about this mission is that the user can summarize the summary from the transcript up to their choices with customizations inclusive of summarizing the specified wide variety of characters from the transcripts and favored languages too which makes it a great deal less difficult than all of the other previous transcript summarizers.

REFERENCES

- [1] 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.
- [2] AdhikaPramitaWidyassari, SupriadiRustad, GuruhFajarShidik, Edi Noersasongko, Abdul Syukur, Affandy De Rosal Ignatius Moses Setiadi, Review of automatic text summarization techniques &methods, Journal of King Saud University-Computer and Information Sciences,2020,ISSN1319-1578, <https://doi.org/10.1016/j.jksuci.2020.05.006>.
- [3] P. R. Dedhia, H. P. Pachgade, A. P. Malani, N. Raul and M. Naik, "Study on Abstractive Text Summarization Techniques," 2020 International Conference on Emerging Trends in Information Technology
- [4] <https://ai.googleblog.com/2020/06/pegasus-state-of-art-model-for.html>
- [5] <https://huggingface.co/docs/tokenizers/index>
- [6] <https://pypi.org/project/youtube-transcript-api/>
- [7] https://huggingface.co/docs/transformers/model_doc/bart
- [8] https://huggingface.co/docs/transformers/model_doc/mt5