

YouTube Transcript Summarizer Analysis using Machine Learning Techniques

¹Ankitha Sowmya, ²Yashaswini G, ³Prof. Chandrani Chakravorty

^{1,2}Department of Master of Computer Applications, RV College of Engineering@ Bengaluru -560059, Karnataka, India

³Assistant Professor, Department of Master of Computer Application, RV College of Engineering@ Bengaluru -560059, Karnataka, India

Authors Email: [1ankithas.mca20@rvce.edu.in](mailto:ankithas.mca20@rvce.edu.in) [2yashaswinig.mca20@rvce.edu.in](mailto:yashaswinig.mca20@rvce.edu.in), [3chandrani@rvce.edu.in](mailto:chandrani@rvce.edu.in)

Abstract - Nowadays, on YouTube, there are lots of videos uploaded every single day over the world which is very long, mostly educational, entertainment, documentaries, news clips, and more. If you want an important source from the videos, it is impossible. And to get a particular content from a video, a whole video has to be watched which is a complete waste of time and the source is also not a piece of relevant information out of it and also effort for watching multiple videos to extract the useful information. To solve that problem, summarizing transcripts are useful and are one way go. It uses the 'Natural Language Processing' technique to summarize the abstractive text.

Keywords: Natural Language Processing, Abstractive Summarization, Hugging Face, Machine Learning, YouTube, Keyword Extraction.

I. INTRODUCTION

Natural Language Processing (NLP) is a field of linguistics and machine learning to understand everything related to human language. Generating summarizing of video transcripts is the process of creating a short, involves outlining the text's major points and a fluent summary of a longer text document. Here, translating a text into another language, with the help of NLP summarizing a text. Text identification, interpretation, summary generation, and analysis of the generated summary are some of the challenges faced in the process of text summarization. The rephrasing tasks in extraction-based summarization identified the key phrases in the document and used them to find out relevant information to be included in the summary. Text Summarization helps in summarizing and shortening the text in the user feedback [2]. It can be done with the help of an algorithm that can help reduce the text bodies while keeping their original text.

Text summarization is done by using a Python API (Application Programming Interface) Hugging Face Transformers. Backend API displays summarized text through chrome extension.

II. LITERATURE SURVEY

Evaluates three techniques for automatic creation of summaries for online audio-video presentations. These Techniques exploit information in the audio signal (e.g., pitch and pause information), knowledge of slide transition points in the presentation, and information about access patterns of previous users [1].

Presents two tools for text summarization. The first text summarization tool performs extractive summarization on the input articles using TF-IDF and Textrank. The extractive summarization allows the extraction of any number of key sentences from the original articles [2].

The conceptualizes that transformers facilitate users to access large-scale pretrained models, to build and experiment on top of them, and to deploy them in downstream tasks with state-of-the-art performance. Transformers has gained significant organic traction since its release and is set up to continue to provide core infrastructure while helping to facilitate access to new models [3].

In Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, California USA, pages 3075-3081 is abstractive training of the extractive summaries alone, eliminating the need for sentence-level extractive labels [4].

An attempt to summarize and present the view of text summarization from every aspect from its beginning till date. The two major approaches are extractive and abstractive summarization [5].

An approach for automatic movie summarization is presented by exploring a new type of user-generated data, which is bullet screen comments, which allow the audience to comment on the movie in a real-time manner. The number of comments on a movie segment shows the exciting degree of the audience, while the content of the comments includes the concepts that interest the audience [6].

A novel technique for summarizing news articles using neural networks is presented. A neural network is trained to learn the relevant features of sentences that should be included in the summary of the article [7].

Automatic text Summarization is an old challenge but now the research direction is learning from extractive text summarization to abstractive text summarization. In abstractive summarization methods, it produces cohesive, highly coherent, information-rich, and less redundant summaries. Abstractive text summarization is a challenging area because of the complexity of natural language processing [8].

In this, it provides a systematic introduction to the field, explaining basic definitions, the strategies used by human summarizers, and automatic methods that leverage linguistic and statistical knowledge to produce extracts and abstracts. Drawing from a wealth of research in artificial intelligence, natural language processing, and information retrieval it also includes detailed assessments of evaluation methods and new topics such as multi-document and multimedia summarization [9].

Uses concepts from text summarization, applied to transcripts derived using automatic speech recognition. Also, use temporal analysis of pauses between words to detect sentence boundaries. IBM Cue Video system is used to perform media processing. Also have developed an experimental design for a user study to judge the quality of the generated summaries using quizzes based on questions about original program content that the subjects answered after watching summaries generated from the program [10].

III. METHODOLOGY

APIs changed the way of building applications. There are countless examples of APIs in the world, and many ways to structure or set up APIs. This system depicts how to create a back-end application directory and structure it to work with the required files.

Python API is utilized which allows users to get the transcripts/subtitles for a given YouTube video. It also works for automatically generated subtitles, supports translating subtitles and it does not require a headless browser.

Text summarization is the task of shortening long pieces of text into a concise summary that preserves key information content and overall meaning [2]. Hugging Face’s transformers library in Python is used to perform abstractive text summarization on the transcript obtained from previous milestone.

The next step is defining the resources that will be exposed by this backend service. This is an extremely simple application, so the only resource will be the summarized text.

Extensions are small software programs that customize the browsing experience. It enables users to tailor Chrome functionality and behaviour to individual preferences. It are built on web technologies such as HTML, CSS and JavaScript. Milestone goes through how to create a recommended Chrome extension application directory and structure it to work with the required files.

A user interface is developed so that the user can interact with the popups which are one of several types of user interface that a Chrome extension can provide. It usually appears upon clicking the extension icon in the browser toolbar.

A basic UI is provided to enable users to interact and display the summarized text but there are some missing links which must be addressed. Functionality is added to allow the extension to interact with the backend server using HTTP REST API Calls.

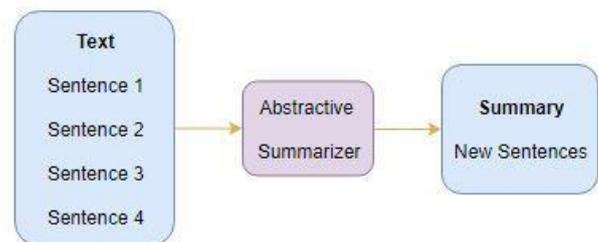


Figure 1: Design of the Abstractive Summarizer

IV. PROPOSED SYSTEM

Step 1: APIs enable the application building. API enables the creation of a backend application directory and builds it to work with the required files.

Step 2: Python API allows getting subtitles/Transcripts for a particular YouTube video.

Step 3: Long pieces of text are shortened into a concise summary with its original key information content. This is done by using Hugging Face’s Transformation PyTorch library in Python.

Step 4: Extension will customize browsing experience, by making tailor chrome functionality and behavior.

Step 5: A user interface enables interaction with the popups, just by clicking the extension icon in the browser toolbar.

Step 6: Using “HTTP REST API” calls a basic UI that enables users to interact and display the summarized text.



Figure 2: Chrome Extension

V. WORKING AND IMPLEMENTATION

Input

- A YouTube video id is entered in the URL
- Checks for id pattern using regular expressions
- If it matches, summarizing process begins or else it returns an error



Figure 3: YouTube Transcript Summarizer

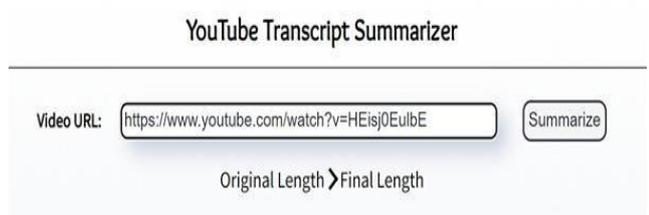


Figure 4: video URL or id

Processing

- Important words and phrases is the original text are summarized to make it easy to understand the content
- Synonyms are not used and there is no rephrasing
- Words are rearranged slightly and given a proper structure

Output

- The summarized text of the YouTube video is displayed in different languages such as English, Hindi, Gujarati, Braille
- The users are also allowed to download the summarized text in a .txt format



Figure 5: Summarized text in English



Figure 6: Summarized text in Hindi

VI. CONCLUSION

The system automatically summarizes video programs on YouTube. Transcripts are derived using the subtitles and temporal analysis of pauses and are used between words to detect sentence boundaries. A YouTube transcript summarization system ensures to make the transcript of the video for mainly educational purposes. It is simple, time-saving which eliminates watching the full video, and gives summarizing the text. It helps in time-saving for the student by watching of video for their knowledge and if there is no proper contained in the video so before watching any video they can use this transcript summarization.

Generally, two different methods are proposed to generate summary and important keywords from the given YouTube video – extractive and abstractive summarization. A simple user interface is made through which users can easily get their summaries through the abstractive method and surely find it easy to interact with the user interface and get what they want. This system will satisfy the users and solve all the problems that it is supposed to tackle which is saving time and effort, by providing only the useful information about the topic which is

interests them so that they do not have to watch those long videos and the time saved can be used in gaining more knowledge. The accuracy by this system approach it was 10% transcripts now, it is 80% - 85% of transcript of summary in educational field.

- To summarize very long transcripts, extractive summarizations techniques are added.
- To summarize transcripts from a non-English video, and display it in the English language.
- To adjust the maximum length of the summarized text.
- To support transcript summarization from a video with no subtitles.

REFERENCES

- [1] L. He, E. Sanocki, A. Gupta, and J. Grudin, "Auto-summarization of audio-video presentations," Proceedings of the 19th ACM International Multimedia Conference, 30 October - 5 November 2018, Orlando, FL, pp. 489-498.
- [2] Ankit Kumar, Zixin Luo & Ming Xu, "Text Summarization using Natural Language Processing". WPI Juniper Networks. (March 2018).
- [3] Victor Sanh, "HuggingFace's Transformers: State-of-the-art Natural Language Processing", 9 Oct 2019.
- [4] Ramesh Nallapati, Feifei Zhai, and Bowen Zhou. 2017. SummaRuNNer: A recurrent neural network-based sequence model for extractive summarization of documents. In Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, California USA, pages 3075-3081.
- [5] Deepali K. Gaikwad and C. Namrata Mahender. 2016. A review paper on text summarization, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5, Issue 3. Pp. 154-160.
- [6] L. Agnihotri, K. V. Devera, T. McGee, and N. Dimitrova, "Summarization of video programs based on closed captions," Proceedings of SPIE Conference on Storage and Retrieval for Media Databases 2016, vol. 4315, January 2016, San Jose, CA, pp. 599-607.
- [7] Khosrow Kaikhah. 2016. Automatic Text Summarization with Neural Networks, In Proceedings of second international conference on intelligent systems, IEEE, 40-44, Texas, USA.
- [8] S. Saranya Mol, L. Sindhu. 2014. A Survey on Automatic Text Summarization. International Journal of Computer Science and Information Technologies, Vol. 5 Issue 6.
- [9] Inderjeet Mani. 2001. Automatic Summarization. Natural language processing. John Benjamins Publishing Company.
- [10] J. Oh and K. A. Hua, "An efficient technique for summarizing videos using visual contents," Proceedings of the IEEE International Conference on Multimedia and Expo (ICME'2000), July 30-August 2 2000, New York, NY.
- [11] J. Kupiec, J. Pedersen, and F. Chen. 1995. A trainable document summarizer. In Proceedings of the 18th ACM SIGIR Conference, pages 68-73.
- [12] Yihong Gong and Xin Liu. 2017. "Generic text summarization using relevance measure and latent semantic analysis". In Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval.
- [13] H. Jing, R. Barzilay, K. McKeown, and M. Elhadad, "Summarization evaluation methods: Experiments and analysis," Proceedings of the AAAI Symposium on Intelligent Summarization, March 23 - 25 1998, Palo Alto, CA.
- [14] M. G. Christel, M. A. Smith, R. Taylor, and D. B. Winker, "Evolving video skims into useful multimedia abstractions," Proceedings of the ACM Computer-Human Interface Conference (CHI'98), April 18-23 1998, Los Angeles, CA, pp. 171-178.
- [15] L. He, E. Sanocki, A. Gupta, and J. Grudin, "Comparing presentation summaries: Slides vs. reading vs. listening," Conference on Human Factors in Computing Systems (CHI 2000), April 1-6 2000, The Hague, Netherlands, pp. 177-184.
- [16] Atif Khan and Naomie Salim. 2014. "A review on abstractive summarization Methods Journal of Theoretical and Applied Information Technology", 2014, Vol. 59 No. 1.
- [17] A. Bookstein, S. T. Klein, and T. Raita. 1995. "Detecting content-bearing words by serial clustering" In Proceedings of the 18th ACM-SIGIR Conference, pages 319-327.
- [18] D. DeMenthon, V. Kobla, and D. Doerman, "Video summarization by curve simplification," Proceedings of the ACM Multimedia Conference, September 12- 16 1998, Bristol, England, pp. 211-218.
- [19] S. Srinivasan, D. Ponceleon, A. Amir, B. Blanchard, and D. Petkovic, "Engineering the web for multimedia," Proceeding of the Web Engineering Workshop (WEBE), WWW-9, May 15-19 2000, Amsterdam, Netherlands.
- [20] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1349-1380, December 2000.
- [21] Y. Taniguchi, A. Akutsu, and Y. Tonomura, "Panorama excerpts: Extracting and packing

- panoramas for video browsing.” Proceedings of the ACM Multimedia, November 9-13 1997, Seattle, WA, pp. 427–436.
- [22] N. Vasconcelos and A. Lippman, “Bayesian modeling of video editing and structure: Semantic features for video summarization and browsing,” Proceedings of IEEE International Conference on Image Processing, October 4- 7 1998, Chicago, IL.
- [23] C. Taskiran, J. Chen, C. A. Bouman, and E. J. Delp, “A compressed video database structured for active browsing and search,” Proceedings of the IEEE International Conference on Image Processing, October 4-7 1998, Chicago, IL.
- [24] D. DeMenthon, V. Kobla, and D. Doerman, “Video summarization by curve simplification,” Proceedings of the ACM Multimedia Conference, September 12- 16 1998, Bristol, England, pp. 211–218.
- [25] S. Uchihashi, J. Foote, A. Girgenson, and J. Boreczky, “Video manga: Generating semantically meaningful video summaries,” Proceedings of ACM Multimedia’99, October 30 - November 5 1999, Orlando, FL, pp. 383–392.

Citation of this Article:

Ankitha Sowmya, Yashaswini G, Prof. Chandrani Chakravorty, “YouTube Transcript Summarizer Analysis using Machine Learning Techniques” Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 6, Issue 2, pp 62-66, February 2022. Article DOI <https://doi.org/10.47001/IRJIET/2022.602011>
