TIME SQUEEZE: QUICK RECAPS FROM YOUTUBE TRANSCRIPTS

[1] Ameena Fatima, [2] Imtiyaz khan, [3] Nazia Tazeen

[1] PG Scholar, Department of Information Technology, Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086. ameenafatima1213@gmail.com

[2] Professor, Department of Information Technology, Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086. Imtiyaz.khan.7@gmail.com

[3] Professor, Department of Information Technology, Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086. naziabaseer@gmail.com

Abstract: The growing amount of digital video material, especially on websites such as YouTube, prevented consumers' ability to effectively obtain useful information from extended films. This project is a Time Squeeze, a chrome, complemented by the API with the Backend Rest API, which autonomously creates a brief summary of video transcripts on YouTube. The system uses the methods of sophisticated natural language processing (NLP) to convert unrefined transcription data into organized, context -conscious summary. The YouTube Transcript API acquires video transcripts that are further refined, divided and condensed by the Backend system using the mining and personal algorithms. Users can get a customized summary according to the audience's inserted personas - such as technical, powerful or occasional - improve the relevance and availability of material. Modular design allows real -time processing, dynamic selection of summarization strategies and easy browser interface. This approach significantly reduces information overload and increases the efficiency of content consumption. This study defines the system design, including the implementation process, NLP algorithms, UML architecture, test protocols and prospective development, such as multilingual support and integration with AR.

"Index Terms - *YouTube, transcript summarization, Chrome extension, REST API, NLP, extractive summarization, user personalization, real-time processing".*

1. INTRODUCTION

The exponential increase in the consumption of digital media has transformed how consumers get information, while websites such as YouTube become primary sources for educational material, expertise and pleasure. Since viewers are increasingly depending on long video material, the difficulty of collecting valuable information without the use of unnecessary time has intensified significantly. While youtube transcripts provide a text approach to video material, sometimes they lack organizations and can be stunning, preventing viewers' ability to quickly identify and understand the necessary information, especially in time limitations.

Recent improvements to the natural language processing (NLP) have facilitated a summary of the text using mining and abstract techniques [1] [2]. Extraction methods recognize and assemble the most important phrases from the text, while abstract techniques, often using transformer -based models such as Bart and T5, produce a summary in a human paraphrased style. [4]. Regardless of progress, most summarization systems provide general outputs, regardless of the consider for user context, history or goal. This gap underscores the growing need for a summary focused on intentions adapted to individualized information consumption [5].

Suggested system, Timesqueeze: Fast recapitulation from YouTube transcripts, implements user -personalized summarization architecture that produces a summary adapted to people, including technical, powerful or occasional. The system uses fast engineering to adapt outputs from a single pre -trained transformer model and therefore increases the scalability and efficiency of calculation rather than training several models [6]. The transcript is first loaded using API, such as YouTube-Transcript-API, and then processed using a modular backend that adapts the summons to selected persons and therefore increases the understanding and connection of users.

This approach serves a diverse range of users, including students who long for academic clarification, experts who need a brief summary and a wide audience who want accessible information. Timesqueeze shows how users focused on users and current NLP can improve the availability of digital content in practical environments through light but smart solutions [7].

2. RELATED WORK

The Summary Fields, especially for lengthy video materials, such as YouTube videos, have attracted considerable interest in recent years due to the rapid expansion of online multimedia and concurrently accelerated and adapted information consumption. Numerous research has examined the ways of converting raw transcripts into a coherent summary using several NLP techniques, including conventional extraction strategies and advanced deep learning methodologies.

Liu et al. [8] He proposed a method unattended for automated keyword extraction from the fulfillment of transcripts and emphasized the importance of detection of relevant phrases before summary. Their research emphasized the need for preliminary processing and filtering in transcription data, similar to the first phase of the Timesqueeze system, where the necessary cleaning and transcript segmentation are necessary to ensure high quality inputs. The emphasis on meetings in [8] corresponds to a wider meaning summary of transcript outside entertainment, including business, academic and personal applications.

Mihalcea and Tarau [9] introduced a renowned Textrank algorithm, a graph -based evaluation model influenced by a pagerank that is effectively used for extraction summarization. Their methodology determined the basis for selecting a sentence on the content by assessing significance. Textrank continues to be a basic method in mining summary and is often used in hybrid systems that combine extraction and abstract elements. Although Timesqueeze mostly uses abstract models such as Bart, the impact of Textrank is evident in the first filtering and content prioritization mechanisms.

Liu and Guo [10] have designed the design of the neural network that integrates two-way short-term memory (BI-STM), attention mechanisms and convolution layers to increase text categorization. Their research illustrates the efficacy of deep

learning to capture long-range relationships and contextual information-editional attributes that also use abstract models such as Bart. Incorporating attention processes makes it easier to understand the transcription information, resulting in a summary. Similarly, Timesqueeze uses the BART -based design, allowing a summary of a controlled personality that is tuned to the context and purpose of the user.

Song et al. [11] They examined the deep learning to summarize the LSTM-CNN hybrid architecture to create an abstract summary. Their approach emphasized the efficiency of deep neural networks in the conversion of input sequences to a clear output. The hybrid characteristics of their paradigm facilitated improved sequence control and semantic structures. On the contrary, Timesqueeze effectively efficiently effectively implement the implementation of rapid engineering instead of a unique model training, which represents a more efficient method while maintaining efficiency.

Panthagani et al. [12] Introduced have introduced a summary summary of YouTube and emphasized the growing need for a solution that improves the availability and understanding for limited users. Their study shows that YouTube, as a multimedia platform, has different problems, including non -functional voice transcripts and inconsistent speakers. The authors used Bert and other models of deep learning to analyze transcripts and provide summary. Their approach has several conceptual parallels with Timesqueeze; However, it does not expect a priority to adapt to personality, which represents the basic innovation of this study.

Chauhan et al. [13] They introduced an alternative summary of transcripts designed to increase the availability and discovery of content from videos on YouTube. Their methodology emphasized the expansion of the user interface by increasing the readability and brevity of transcripts. Although their methodology was effective for a wide summary, it had a lack of personalization. Timesqueeze exceeds conventional adaptation methods Summary based on user profiles, providing technical depth for students, business perspective for managers and simpler tone for occasional users.

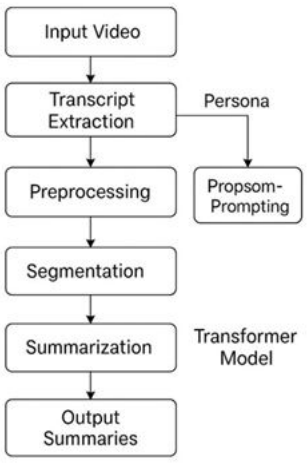
Kirwai et al. [14] have resolved accessibility using YouTube transcripts to increase access to information for individuals with disorders and other user demographies. Their system emphasized the social consequences of the method summary of the transcript. This corresponds to the goal of Timesqueeze Democratization of access to video material, especially for users who are limited or have different preferences of understanding. The concept of creating inclusive and user -friendly summaries is essential for both systems, while Timeseeze includes another dimension through personality customization.

Streck and Reddy-Best [15] examined the use of YouTube in discourse confirmed sex and analyzed the usefulness of the DIY lessons for different cultures. Although their study is not focused on the methodology summary, the umbrella's social significance emphasizes YouTube transcripts and the need for improved tools for assimilation of such material. Their research suggests that technologies such as Timesqueeze can increase content inclusiveness by facilitating easier access to a specialized community -oriented video content through an intelligent summary.

3. MATERIALS AND METHODS

The proposed system, Timesqueeze, represents an efficient and customizable method for summary of video transcriptions on YouTube using a persona-based challenge. Unlike conventional methods that

create a uniform summary, Timesqueeze allows you to choose between three personalities - technical, powerful or occasional - affect the tone and complexity summary. The system uses preliminary transformer models such as Bart, to provide a summary without the need for retraining or fine - tuning, and therefore minimizes computing requirements. The transcription data is obtained via the YouTube Transcript API API, then preliminarily processed and directed via the input pipe with increased strengthening to provide adapted outputs. This method increases the summary of the suitability and involvement of users and at the same time guarantees considerable scalability and deployment. Unlike previous systems that prefer a wide summary [12] [13], Timesequeeze is a unique layer of user-comprehensive adaptation, increasing its practical importance for students, professionals and daily users. Drawing from recent progress in transcriptional instruments aimed at accessibility [14] This study emphasizes that personalization in NLP can be achieved by simple and effective input adjustments instead of complex architectural changes.



"Fig 1 Proposed Architecture"

The solution uses the client-server modular architecture in which the client transmits the URL videos of YouTube and chooses the personality to the backend. Backend gets a transcript using the API YouTube API, performs preliminary processing and uses the transformer hugging model to provide a summary summary for persona. The summary is then sent to the customer. The versatility of the design makes it easier to integrate future improvements, such as multilingual summarization, domain optimization or deployment through light frames such as Fastapi or streamlining for scalable online access.

i) Transcript Extraction and Preprocessing:

The first step of the Timesqueeze system is the extraction of raw transcripts from YouTube videos using YouTube-Transcript-API. This API facilitates the extraction of accessible subtitles-a user-adapted or automatically generated-in-text format. Since automatic generated transcripts often lack punctuation and sentences, pre -workment is essential for increasing the readability and preparation of material for summary. The pipeline includes a lower incidence, eliminate special characters and restore punctuation using technologies such as Deepsegin or Heuristics NLP based on rules. The material is then divided into coherent segments such as paragraphs to maintain grammatical consistency. This systematic method facilitates increased downstream summarization and corresponds to the optimal standards in the pre - contact text for NLP tasks. Effective pre -workment converts noisy raw data into organized, high quality input suitable for abstract models, directly affecting clarity and coherence summary [8].

ii) Persona-Based Prompt Engineering:

The remarkable characteristic of Timesqueeze is its implementation of rapid engineering based on personality, which allows personalization without the need to retrain the model. At the beginning of the system, users choose persona - technical,

powerful or occasional - and evoke a summary output to their preferences. The technical question can get complex answers and specialized terminology, while an informal challenge focuses on conversational style. This challenge precedes the refined transcript and serves as a directive for modeling for tone, structure and depth. This lightweight adaptation method uses the ability to monitor transformer models to provide unique summary styles. The system maintains scalability and efficiency on several platforms using rapid engineering rather than computingly demanding model adaptation. Recent studies suggest that speed -based input modification can successfully control the model's behavior in different NLP tasks, providing an effective alternative to retraining [13].

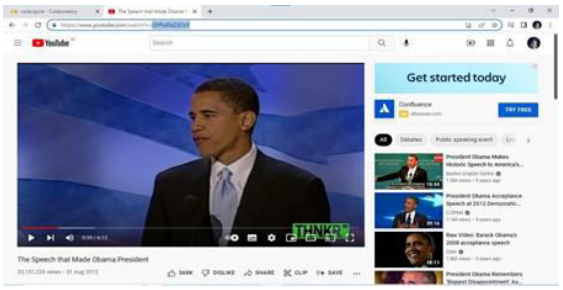
iii) Transformer-Based Abstractive Summarization:

Timesqueeze uses transformer abstraction summarizing models, including BART (Bidirectional and Auto-Regressive Transformers) and T5 (Text-to-Text Transfer Transformer) to create high quality summons. These models are pre -trained on extensive data sets summaries and are recognized for generating cohesive and context -aware output. Unlike mining approaches, abstract models can paraphrase and recontext the material and create a summary of a person that, more precisely, encapsulate the spirit of the original content. The system eliminates the need for retraining of the model; Rather, it uses inputs with fast engineering to generate outputs specific to style according to user -selected persons. This approach guarantees different output styles and at the same time significantly reduces the computing load, which makes the system efficient and user -friendly. The library of transformers from hugging the face facilitates implementation by providing pre -trained checkpoints and an efficient API interface. The versatility and fluency of models such as Bart makes them very useful in practical summarization applications, especially when paired with human input [10].

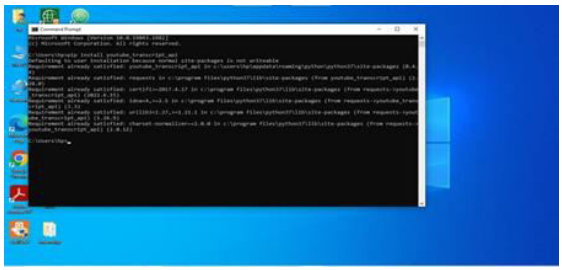
iv) Output Generation and Evaluation

The final phase includes the structuring and presentation of the created summary in accordance with the selected personality. The output is provided via a web interface or command line program that facilitates accessibility. Each summary is adapted to the tone - analytical for technical, concise for the executive and conversational for occasional - offers a personalized experience. The discretional assessment framework facilitates quality verification through quantitative and qualitative approaches. The Rouge score evaluates the lexical congruence between generated models and reference summons from human and provides objective performance assessment. In addition, the user's input is collected to evaluate clarity, tone and relevance. The score of persons alignment is used to determine whether the model properly embodies the selected style. This evaluation cycle makes it easy to identify areas for improvement and confirms the effectiveness of the system across different demographies of the user. Research Systems Summary shows that user input integration improves summary quality and involvement, especially in personality -based applications [14].

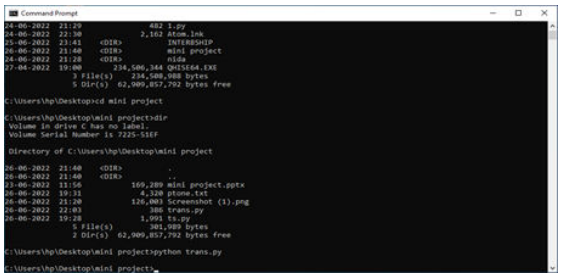
4. RESULTS AND DISCUSSION



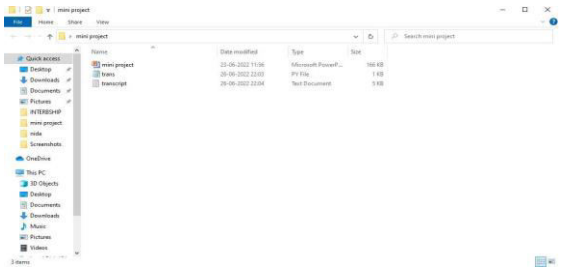
"Fig 2 Input"



"Fig 3 Input"

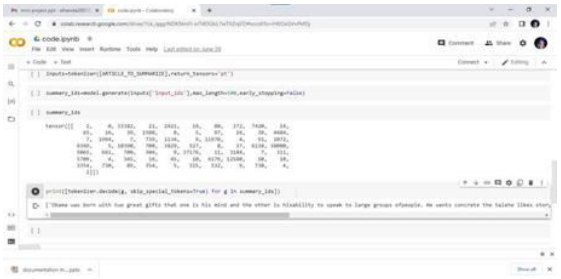


"Fig 4 Request for transcript"

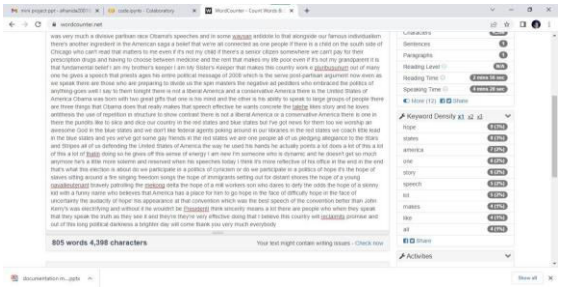


"Fig 5 Saved transcript in file"

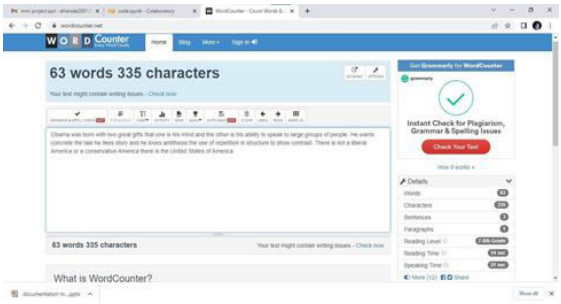
"Fig 6 Transcript of given video ID"



"Fig 7 Summary of the transcript"



"Fig 8 Number of words before summarization"



"Fig 9 Number of words after summarization"

5. CONCLUSION

In conclusion, Timesqueeze - rapid recapitulation from YouTube transcripts provides an effective and new cure for the escalatory edition of the overload of the video content. Using NLP models on YouTube Transcript, Fast Engineering and Transformers such as Bart and T5, the system provides a brief, persona specific summary adapted to different requirements of-technical, powerful and occasional user requirements. This tailor -made method increases the availability and understanding, while reducing the cognitive burden and the time it takes to get relevant long films. The solution

illustrates the successful achievement of user -controlled summarization without the need for retraining sophisticated models, which is lightweight, scalable and adaptive. Applications include education, business training, research and general entertainment where effective information extraction is necessary. The existing system can accommodate English language transcripts and a general summary, while potential enhancements may include multilingual support, time -based summons, visual context integration, and implementation through user -friendly graphical interfaces, or online platforms. Timesqueeze creates a robust framework to create intelligent tools for summary tools. It shows how the NLP can be used to develop intelligent and adaptable interfaces that synchronize the supply of materials with the user's goal, making it easier to make a more efficient and more adapted digital learning experience.

Future development for Timesqueeze includes improvement in language summary support, allowing real -time processing using Chrome extension and including multimodal inputs such as video images. Improvements include dynamic profiling of persons, voice orders and interactive summary will increase personalization and accessibility. Additional features such as Cloud Storage, Summary and AI generated chapters will increase usefulness. Models -specific model refining and semantic searches integration via vector storage will increase accuracy and contextual meaning and time is developing to a sophisticated and intelligent assistant summary of the video.

REFERENCES

[1] Anuj Gupta, Bodhisattwa Prasad Majumder, Harshit Surana and Sowmya Vajjala. “Practical natural language processing” A comprehensive guide to build real world NLP system. 1st edition. O’Reilly media, inc(june,2020)

[2] Ankit Kumar, Zixin Luo & Ming Xu, "Text Summarization using Natural Language Processing". WPI Juniper Networks. (March,2018).

[3] Gaurav Sharma, Shaba Parveen Khan, Shivanshu Sharma and Syed Ubed Ali (2021). "Summarizer For Easy Video Assessment (SEVA)" Vol.03. e-ISSN: 2582-5208. (April,2021)

[4] Hamza Shabbir Moiyadi, Harsh Desai, Dhairya Pawar, Geet Agrawal, and Nilesh M. Patil. (2016). "NLP Based Text Summarization Using Semantic Analysis". IJAEMS. ISSN: 2454- 1311 (Oct,2016)

[5] Korra Sathya Babu, Santosh Kumar Bharti and Sanjay Kumar Jena.2017 "Automatic Keyword Extraction for Text Summarization: A Survey" (Feb,2017).in this paper recent literature on automatic keyword extraction and text summarization are presented since text summarization process is highly depended on keyword extraction.

[6] Pratik K. Biswas, Aleksander Lakubovich. (2020)" Extractive Summarization on Call Transcript" Al & Data Science, Global Network and Technology, Verizon Communications, New Jersey, USA. This paper presents an indigenously developed method that combines topic modelling and sentence selection with punctuation restoration in condensing ill-punctuated or un-punctuated call transcripts.

[7] Cuneyt M. Taskiran, Arnon Amir, Dulce Ponceleon and Edward J. Delp. 2009. "Automated Video Summarization Using Speech Transcripts". Video and Image Processing Laboratory (VIPER)Can, 2009 compact representation of video data that enable efficient browsing such

representation provides the user with information about the content of particular sequence being examined while the actual content being preserved.

[8] Feifan Liu, Deana Pennell, Fei Liu and Yang Liu. 2009. "Unsupervised Approaches for Automatic Keyword Extraction Using Meeting Transcripts". Computer Science Department at the University of Texas at Dallas Richardson, USA. (June 2009)

[9] Rada Mihalcea, Paul Tarau. "TextRank: Bringing Order into Texts". Department of Computer Science University of North Texas. Textrank-a graph based ranking model can be successfully used in natural language applications. graph based ranking algorithms like kleinbergs algorithm or google page rank have been successfully used in citation analysis.

[10] Gang Liu, Jiabao Guo. 2019. Bidirectional STM with attention mechanism and convolutional layer for text classification. School of Computer Science, Hubei University of Technology, Wuhan China (Feb,2019). Neural network models have been widely used in field of natural language processing recurrent neural networks which have the ability to process sequence of arbitrary length are common methods for sequence modelling tasks.

[11] Shengli Song, Haitao Huang and Tongxiago Ruan.2018. "Abstractive Text summarization using LSTMCNN based deep learning. Springer Science+Business Media.

[12] Panthagani, V. B., Duggempudi, V. D. R., Kommera, N. L., & Vakkalagadda, N. (2024, January). Youtube Transcript Summarizer. In 2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI) (pp. 891-898). IEEE.

[13] Chauhan, V., Tiwari, M., Bharadwaj, H., Goswami, I., & Pareeka, S. (2024). YouTube Transcript Summarizer: Enhancing Accessibility and Content Discovery. Available at SSRN 4832193.

[14] Kirwai, N., Kashikar, S., Bali, A., Badhiye, S. S., Borkar, P., & Waghmare, K. (2024, November). Unveiling Voices: Empowering Accessibility Through YouTube Transcripts. In 2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS) (pp. 903-907). IEEE.

[15] Streck, K. G., & Reddy-Best, K. L. (2025). Openly-trans YouTubers and gender-affirming undergarments: Production, distribution, and consumption via DIY tutorials on YouTube. Clothing and Textiles Research Journal, 43(1), 65-80.