Project Synopsis

#### on

YouTube Transcript Summarizer

Submitted as a part of course curriculum for

# Bachelor of Technology

in

# Computer Science



**Submitted by**

#### Harsh Bhardwaj (2000290120070)

Ishita Goswami (2000290120079)

#### Manish Tiwari (2000290120091)

**Under the Supervision of** Prof. Harsh Vardhan Assistant Professor

# KIET Group of Institutions, Ghaziabad Department of Computer Science

**Dr. A. P. J . Abdul Kalam Technical University**

**2022-2023**

# DECLARATION

We hereby declare that this submission is our work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgement has been made in the text.

Signature of Students

Name: Harsh Bhardwaj Roll No.: 2000290120070

Date: 11-11-2022

Name: Manish Tiwari Roll No.: 2000290120091

Date: 11-11-2022

Name: Ishita Goswami Roll No.: 2000290120079

Date: 11-11-2022

# CERTIFICATE

This is to certify that Project Report entitled “YouTube Transcript Summarizer” which is submitted by **Harsh Bhardwaj, Manish Tiwari, Ishita Goswami** in partial fulfilment of the requirement for the award of degree B. Tech. in Department of Computer Science of Dr A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

**Date: Supervisor Signature**

Prof. Harsh Vardhan (Assistant Professor)

# ACKNOWLEDGEMENT

#### It gives us a great sense of pleasure to present the synopsis of the B.Tech Mini Project undertaken during B.Tech. Third Year. We owe a special debt of gratitude to Prof. Harsh Vardhan (Assistant Professor), Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his/her constant support and guidance throughout the course of our work. His/Her sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his/her cognizant efforts that our endeavours have seen the light of the day.

We also take the opportunity to acknowledge the contribution of Dr. Ajay Kumar Shrivastava, Head of the Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

#### Last but not the least, we acknowledge our friends for their contribution to the completion of the project.

Signature: Date : Name : Roll No:

# ABSTRACT

#### Integrated video data presentations may allow active video browsing. Such presentations provide the user with information about the content of a particular sequence being tested while maintaining an important message. We suggest how to automatically make video summaries for longer videos. Our video access method involves two tasks: first, splitting the video into smaller, compatible parts and second, setting the levels into effects. Our proposed algorithm sections are based on analysis of word frequency in speech transcripts. After that the summary is made by selecting the parts with the highest scores depending on the length of time and these are illustrated. We created and conducted a user study to check the quality of the summaries made. Comparisons are made using our proposed algorithm and a random segment selection scheme based on mathematical analysis of user learning outcomes. Finally, we can see the summarized context of the video we want to know about.

Summarization of the video is done by the Python API and NLP (Natural Language Processing). An API, or Application Programming Interface, is a server you can use to receive and send data using code. APIs are widely used to retrieve data, and that will be the focus of this first study.

When we want to receive data from an API, we need to make a request. Applications are used across the web.

# TABLE OF CONTENTS

Page No.

TITLE PAGE .................................................................................................................... i

[DECLARATION .............................................................................................................. ii](#_TOC_250013)

[CERTIFICATE …........................................................................................................... iii](#_TOC_250012)

[ACKNOWLEDGEMENT.................................................................................................. iv](#_TOC_250011)

[ABSTRACT...................................................................................................................... v](#_TOC_250010)

[CHAPTER 1 INTRODUCTION 1](#_TOC_250009)

[1.1. Introduction ……………………................................................... 1](#_TOC_250008)

[1.2 Problem Statement.……………………....................................... 1](#_TOC_250007)

1.2. Objective………………………………………………………… 1

[CHAPTER 2 LITERATURE REVIEW…………………………………………….... 2](#_TOC_250006)

[CHAPTER 3 PROPOSED METHODOLOGY …………………………………........ 3](#_TOC_250005)

[3.1 Flowchart 4](#_TOC_250004)

[CHAPTER 4 TECHNOLOGY USED ………..………………………..………………. 5](#_TOC_250003)

[CHAPTER 5 DIAGRAMS …........................................................................................... 6](#_TOC_250002)

[CHAPTER 6 CONCLUSION …....................................................................................... 7](#_TOC_250001)

[REFERENCES….............................................................................................................. 8](#_TOC_250000)

**CHAPTER 1 INTRODUCTION**

### Introduction

The summarizer is a Chrome extension that works with YouTube to extract the key points of a video and make them accessible to the user. The summary is customizable per user's request, allowing varying extents of summarization. Key points from the summarization process, together with corresponding time-stamps, are then presented to the user through a small UI next to the video feed. This allows the user to navigate to more important sections of the video, to get to the key points more efficiently.

### Problem Statement

A large number of video recordings are made and shared online all day. It is very difficult to spend time watching such videos which may be longer than expected and sometimes our efforts may be in vain if we do not get the right information about it. Summarize the text of those videos automatically allows us to quickly look at important patterns in the video and helps us save time and effort in all the content of the video.

### Objective

In this project, we will be creating a Chrome Extension that will apply to the backend REST API where it will do NLP and respond with a summary of YouTube text.

## CHAPTER 2 LITERATURE REVIEW

The Field of NLP is a field under artificial intelligence, which is used to classify, and process text with greater efficiency and accuracy than humans. While NLP as a technology can be a great technology for the love of many businesses, NLP APIs are an easy way for companies, organization, group or a single human to integrate technology into business and many more processes. Integrating NLP APIs with existing business plan, software assisted companies to increase the efficiency and effectiveness of business processes. It’s all about teaching machines to understand how to understand human languages and extract meaning from text. It provides developers with extensive collection of NLP tools and libraries that enables developers to handle a great number of NLP related tasks such as document classification, topic modelling, part of speech (POS) tagging, word vectors, and sentiment analysis.

Modem technology can analyse more details about language than humans, without exhaustion and in a consistent, impartial way. Given the staggering amount of informal data generated on a daily basis, from medical records to social media platforms, automation will be essential to fully analyse text and speech data effectively.

Human language is amazingly complex and diverse. We express ourselves in endless ways, verbally and in writing. Not only are there thousands of languages and dialects, but within each language there is a different set of grammar and syntax rules, Niles and slang. When we write, we often mispronounce words or abbreviate words, or leave punctuation marks. When we speak, we have ways of speaking in the region, and we mix. help and borrow words in other languages. While both supervised and supervised learning, and especially in-depth reading, are now widely used in modelling the human language, there is also a need for a well-crafted understanding and background technology that is not present in these electronic learning methods.

## CHAPTER 3 PROPOSED METHODOLOGY

First, we need to get the subtitles or transcript for a given YouTube video id by using the python API known as YouTube transcript API. Since there are three types of transcript that we can extract - manually generated transcript, automatically generated transcript, and the videos that contain no transcript. We are not considering videos that do not have transcripts. Secondly, when we get the transcript of a given YouTube video since it does not contain any punctuations like comma(.), full stops(.) which is very important for us in finding the boundaries of a sentence, so we will restore punctuations from our extracted transcript by using the python library known as "punctuator". Now we will apply the text pre-processing methods to clean the extracted transcript by tokenizing the sentences as well as the words, lowercasing it, removing stop words like a, an, the, etc, removing punctuations, and stemming or lemmatization to generate the root form of inflected words. Performing text summarization: This task consists of shortening a large form of text into a precise summary that keeps all the necessary information intact and preserves the overall meaning. For this purpose in NLP for text summarization, there are two types of methods used: Extractive Summarization: In this type of text summarization, the output is only the important phrases and sentences that the model identifies from the original text.

For the purpose of extractive summarization, we have used the TF-IDF model with Text Rank Algorithm. TF-IDF(Term Frequency - Inverse Document Frequency) After the cleaning process, we have to convert the words into its vectorized form so that our algorithm will process it by using TF-IDF. This is a technique to measure the quantity of a word in documents, we compute a weight to each word which signifies the importance of the word in the document and corpus. TF(Term Frequency): TF calculates the frequency of a word in a document. TF No, of repetition of the word in the sentence / No. of words in a sentence IDF(Inverse Document Frequency): IDF is the inverse of the document frequency which measures the informativeness of term t. IDF = log(No. of sentences/No. of sentences containing words) After this, we will multiply both matrices to obtain the vectorized form which tells us which words are the most important.

### Flowchart

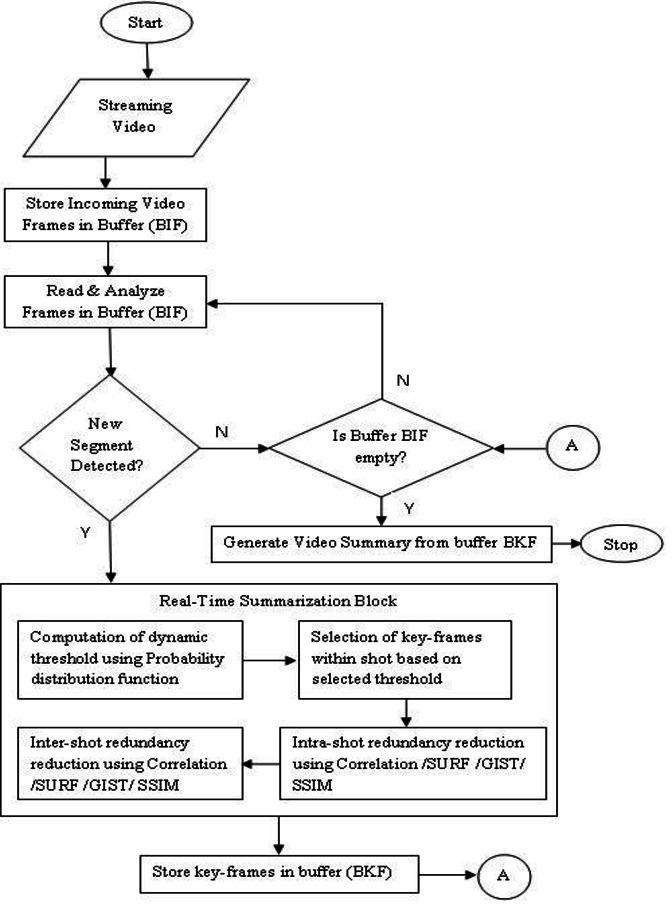


Fig. 2 Flowchart

## CHAPTER 4 TECHNOLOGY USED

* + - REST APIs for backend services
    - HTML, CSS, JavaScript for chrome extension
    - Hugging Face Transformer for summarization
    - Python API to get transcript of a given YouTube video

## CHAPTER 5 DIAGRAMS

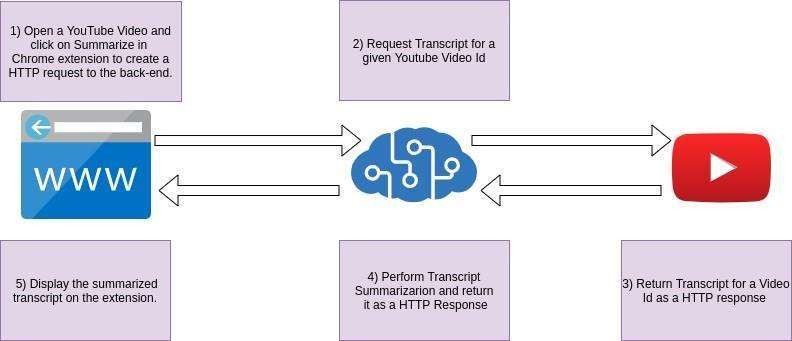


Fig.2 Project Stages

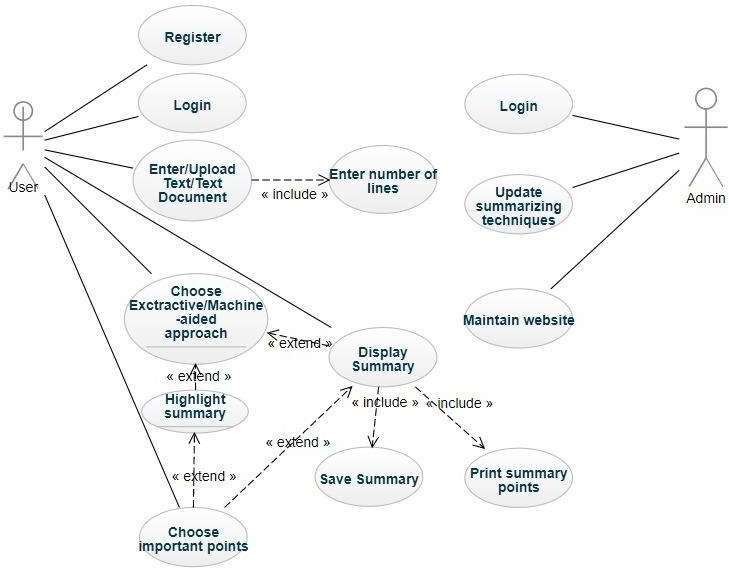


Fig. 3 Use Case Diagram

## CHAPTER 6 CONCLUSION

The increase in popularity of video content on the internet requires an efficient way of representing or managing the video. This can be done by representing the videos based on their summary.

Learning how to set up web services using API, create Google Chrome extensions and implement the Cloud Computing. In addition, we used HTML and CSS to develop Web Apps and write software packages in python.

We need to follow time management and fully grasp the difficulties which may occur and when to change the course of the project based to use our time more efficiently.

It is important to understand connections between different technology and to account for possible bugs when incorporating different software packages into the corpus of a final software product.

## REFERENCES

* A. N. S. S. Vybhavi, L. V. Saroja, J. Duvvuru and J. Bayana, "Video Transcript Summarizer," 2022 International Mobile and Embedded Technology Conference (MECON), 2022, pp. 461-465, doi: 10.1109/MECON53876.2022.9751991
* P. Choudhary, S. P. Munukutla, K. S. Rajesh and A. S. Shukla, "Real time video summarization on mobile platform," 2017 IEEE International Conference on Multimedia and Expo (ICME), 2017, pp. 1045-1050, doi: 10.1109/ICME.2017.8019530
* Ying Li, Shih-Hung Lee, Chia-Hung Yeh and C. . -C. J. Kuo, "Techniques for movie content analysis and skimming: tutorial and overview on video abstraction techniques," in IEEE Signal Processing Magazine, vol. 23, no. 2, pp. 79-89, March 2006, doi: 10.1109/MSP.2006.1621451.
* Yu-Fei Ma, Xian-Sheng Hua, Lie Lu and Hong-Jiang Zhang, "A generic framework of user attention model and its application in video summarization," in IEEE Transactions on Multimedia, vol. 7, no. 5, pp. 907-919, Oct. 2005, doi: 10.1109/TMM.2005.854410