SYNOPSIS ON
YouTube Transcript Summarizer

Submitted By
Name: Prachi Agrawal
Branch/Sec: CSE/C
University RollNo: 191500552

Submitted To
FacultyName:Mandeep Singh
Technical Trainer
Computer Engineering and Applications

INTRODUCTION

Enormous number of video recordings are being created and shared on the Internet throughout the day. It has become really difficult to spend time in watching such videos which may have a longer duration than expected and sometimes our efforts may become futile if we couldn't find relevant information out of it. Summarizing transcripts of such videos automatically allows us to quickly look out for the important patterns in the video and helps us to save time and efforts to go through the whole content of the video.

This project will give us an opportunity to have hands-on experience with state of the art NLP technique for abstract text summarization and implement an interesting idea suitable for intermediates and a refreshing hobby project for professionals.

EXISTING SYSTEM

It’s fairly easy to download a transcript from YouTube. Keep in mind that transcripts are only available for videos with closed captioning. YouTube has Google’s speech recognition software that can provide automatic transcription for most of the videos as soon as they are uploaded. Some owners of YouTube videos provide their own transcripts, which are of better quality than those produced automatically.

There’s also an option to transcribe a YouTube video without an available transcript. Google Docs has a built-in voice typing feature that listens to video or audio and automatically types what it hears.

Trint’s transcription software creates transcripts in no time at all.

But all these mentioned above are just providing transcription for the video only.

USE OF THE PROJECT

- Meetings and video-conferencing - A system that could turn voice to text and generate summaries from your team meetings.
- Patent research - A summarizer to extract the most salient claims across patents.

FEASIBILITY OF PROJECT

Project Overview

In this project, you will be a creating a Chrome Extension which will make a request to a backend REST API where it will perform NLP and respond with a summarized version of a YouTube transcript.

1. **Getting Started with the back-end**
Initialize the back-end portion of your application with the required boiler plate as well as the dependencies.
2. **Get transcript for a given video**
fetch the transcripts with the help of a function created which we will later utilize as a feed input for the NLP processor in the pipeline.

3. **Perform text summarization**
Verify that the model generates a completely new summarized text that is different from the original text.
4. **Create REST API endpoint**
create an endpoint to summarize YouTube video transcripts and test the response with different video URLs
5. **Getting started with Chrome Extension**
Create a recommended Chrome extension application directory and structure it to work with the required files
6. **Build a User Interface for Extension Popup**

7. **Display summarized transcript**

FUNCTIONAL SPECIFICATION

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

Software Specification

• **Technology Implemented** : Natural Language Processing, Flask
• **Language Used** : Python, HTML, CSS, JavaScript
• **Database** :
• **User Interface Design** :
• **Web Browser** : Google Chrome

Hardware Requirements

- **Processor** : AMD and Intel
- **Operating System** : Windows, Linux, Macs
- **RAM** : minimum 8GB
- **Hard disk** : 512GB
- **Display** : low-blue light and flicker-free screens

Future of project

Even though our current version of Youtube summarizer is able to provide users with valuable information regarding the videos they watch and can provide a compact summary of the video, we believe that this tools can be further developed to meet the needs of YouTubers. While we used advanced online services to provide a summary of the description, we believe that we can use a variety of text, audio and video analysis tool to provide valuable information and more accurate summaries for videos. As an example, we attempted to use the online Google Cloud AI services to provide transcripts for videos with no captions and use AWS services to analyze audio files and figure out important sections of the video based on its correlated audio file. Unfortunately, many of these services required more time in order to fully analyze the information and were not incorporated into the final product. In addition, we believe that using Youtube's video API, we will be able to provide more control for our users and give them more freedom when using this tool. Overall, we believe that video analysis and summarization is not only important to the individual, but also proves to be more important when considering big data analysis. Hence, we believe the youtube summarizer can be the start of a novel technology in informatics.