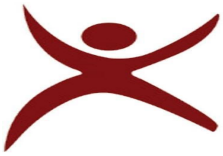


DEPARTMENT OF COMPUTER SCIENCE ENGINEERING

Rajiv Gandhi University of Knowledge Technologies– Nuzvid,

Eluru, Andhra Pradesh – 521202.

Youtube Transcript Summarizer

using NLP

A Project Report

Submitted in partial fulfillment for the degree of

BACHELOR OF TECHNOLOGY
in
COMPUTER SCIENCE AND ENGINEERING

Submitted by

N190375	K.JyothiSai
N190370	B.Vemesh
N190245	A.V.V.amsi
N191035	S.Sravani
N190495	Sk.Jasmin

Under the Esteem Guidance of

Mrs. D. Srilakshmi

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING

Rajiv Gandhi University of Knowledge Technologies – Nuzvid,

Eluru, Andhra Pradesh – 521202.

CERTIFICATE OF COMPLETION

This is to certify that the work entitled, "**Youtube Transcript Summarizer using NLP**"is the bonafide Work of **K.JyothiSai (IDNo:N19o375) ; B.Vemesh (IDNo:N190370); A.V.Vamsi (IDNo: N190245); S.Sravani (IDNo:N191035) ;Sk.Jasmin(IDNo: N190495) ;** carried out Under my guidance and supervision for 3rd year minor project of Bachelor of Technology in the department of Computer Science and Engineering under RGUKT IIIT , Nuzvid.This work is done during the academic session February 2024 – June 2024, under our guidance.

Mrs. D. Srilakshmi

Assistant professor,
Department of CSE,
RGUKT, Nuzvid.

Dr . D.V. Nagarjana Devi

Head of the Department,
Department of CSE,
RGUKT, Nuzvid.

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING

Rajiv Gandhi University of Knowledge Technologies – Nuzvid,

Eluru, Andhra Pradesh – 521202.

CERTIFICATE OF EXAMINATION

This is to certify that the work entitled "**Youtube Transcript Summarizer using NLP**"is the bonafide Work of **K.JyothiSai (IDNo:N19o375) ; B.Vemesh (IDNo:N190370); A.V.Vamsi (IDNo: N190245); S.Sravani (IDNo:N191035) ;Sk.Jasmin(IDNo: N190495) ;**And here by accord our approval of it as a study carried out and presented in a manner required for its acceptance in 3rd Year of Bachelor of Technology for which it has been submitted .This approval does not necessarily endorse or accept every statement made ,opinion expressed ,or conclusion drawn, as recorded in this thesis.It only signifies the acceptance of the thesis for the purpose for which it has submitted.

Mrs. D. Srilakshmi

Assistant professor,
Department of CSE,
RGUKT, Nuzvid.

Project Examiner

Assistant Professor,
Department of CSE,
RGUKT, Nuzvid.

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING

Rajiv Gandhi University of Knowledge Technologies – Nuzvid,

Krishna, Andhra Pradesh – 521202.

DECLARATION

We **K.JyothiSai (IDNo:N19o375) ; B.Vemesh (IDNo:N190370); A.V.Vamsi (IDNo: N190245); S.Sravani (IDNo:N191035) ;Sk.Jasmin(IDNo: N190495) ;** hereby declare that the project report entitled "**Youtube Transcript Summarizer using NLP**"\done by us under the guidance of Mrs.D.Srilakshmi , Assistant Professor, is submitted for the fulfillment of a minor project during the academic session February 2024-June 2024 at RGUKT-Nuzvid.

We also declare that this project is a result of our own effort and has not been copied or imitated from any source. Citations from any websites are mentioned in the references.The results embodied in this project report have not been submitted to any other university or institute for the award of any degree or diploma.

Date: 30-06-2024
Place: Nuzvid

N190375	K.JyothiSai
N190370	B.Vemesh
N190245	A.V.amsi
N191035	S.Sravani
N190495	Sk.Jasmin

ACKNOWLEDGEMENT

We would like to express our profound gratitude and deep regards to our guide **Mrs. D.Srilakshmi** for her exemplary guidance, monitoring and constant encouragement to us throughout the B.Tech course. We shall always cherish the time spent with her during the course of this work due to the invaluable knowledge gained in the field of reliability engineering.

I'm extremely grateful for the confidence bestowed in me and entrusting our project entitled "**Youtube Transcript Summarizer using NLP**".

We express gratitude to Mrs. D. Srilakshmi (Ass.Professor) and other faculty members for being source of inspiration and constant encouragement which helped us in completing the project successfully.

Our sincere thanks to all the batch mates of 2019 CSE, who have made our stay at RGUKT-NUZVID, a memorable one.

Finally, yet importantly, we would like to express our heartfelt thanks to our beloved God and parents for their blessings, our friends for their help and wishes for the successful completion of this project

ABSTRACT

This project involves creating a YouTube transcript summarizer using Google Gemini Pro and Streamlit. The application extracts transcripts from YouTube videos and generates concise summaries using the Google Gemini Pro model. Users can input a YouTube video URL, and the app displays the video's thumbnail. The summary can be translated into multiple languages using the Google Translate API, catering to a diverse audience. The summarized content is available for download in the chosen language, making it convenient for users to save and share the information. This tool provides an efficient way to obtain and disseminate key points from YouTube videos.

.

TABLE OF CONTENTS

- Abstract
- Introduction & Real world applications
- Related works / Existing works
- Proposed Methodology
- Technologies and Libraries used
- API Key Generation
- NLP Technologies used
- Hardware and software requirements
- Results
- Conclusion and Future Scope
- References

CHAPTER 1

Introduction

1.1 MOTIVATION FOR THE WORK

- This project aims to simplify the process of extracting and summarizing YouTube video content, making it accessible and comprehensible for diverse audiences.By integrating advanced AI models for summarization and translation, users can quickly obtain key insights from videos in their preferred language and complexity level. This tool is particularly useful for students, researchers, and professionals who need to efficiently digest large amounts of video content**.**

1.2 Real-World Applications of Placement Prediction

- Education: Helps students and educators quickly grasp key concepts from educational videos, such as summarizing a physics lecture into key formulas and principles.

- Content Creation: Aids content creators in generating concise summaries for video descriptions or social media posts. For instance, summarizing a cooking tutorial video into key steps and ingredients for a recipe post.

- Research: Facilitates researchers in extracting relevant information from academic or instructional videos efficiently. For example, summarizing a scientific conference presentation into key findings and methodologies for research papers.

.

CHAPTER 2

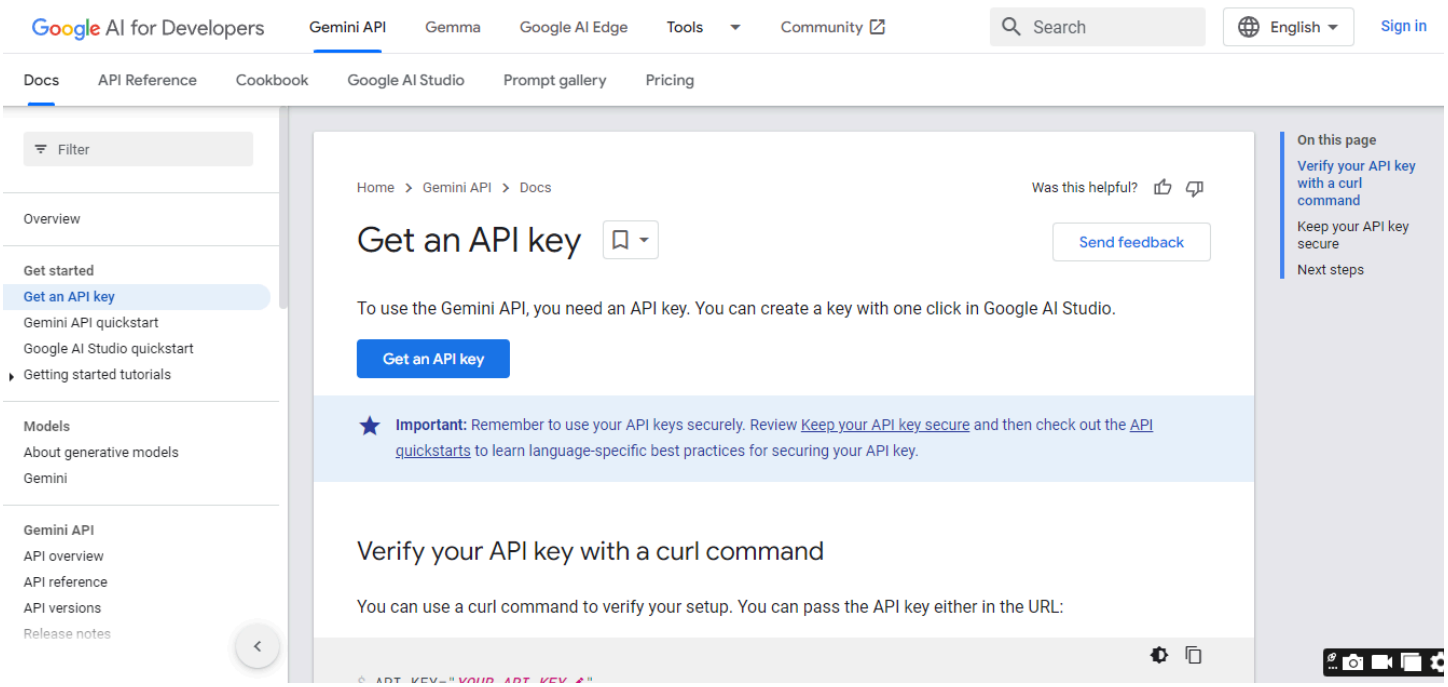
2.1Technologies and Libraries used

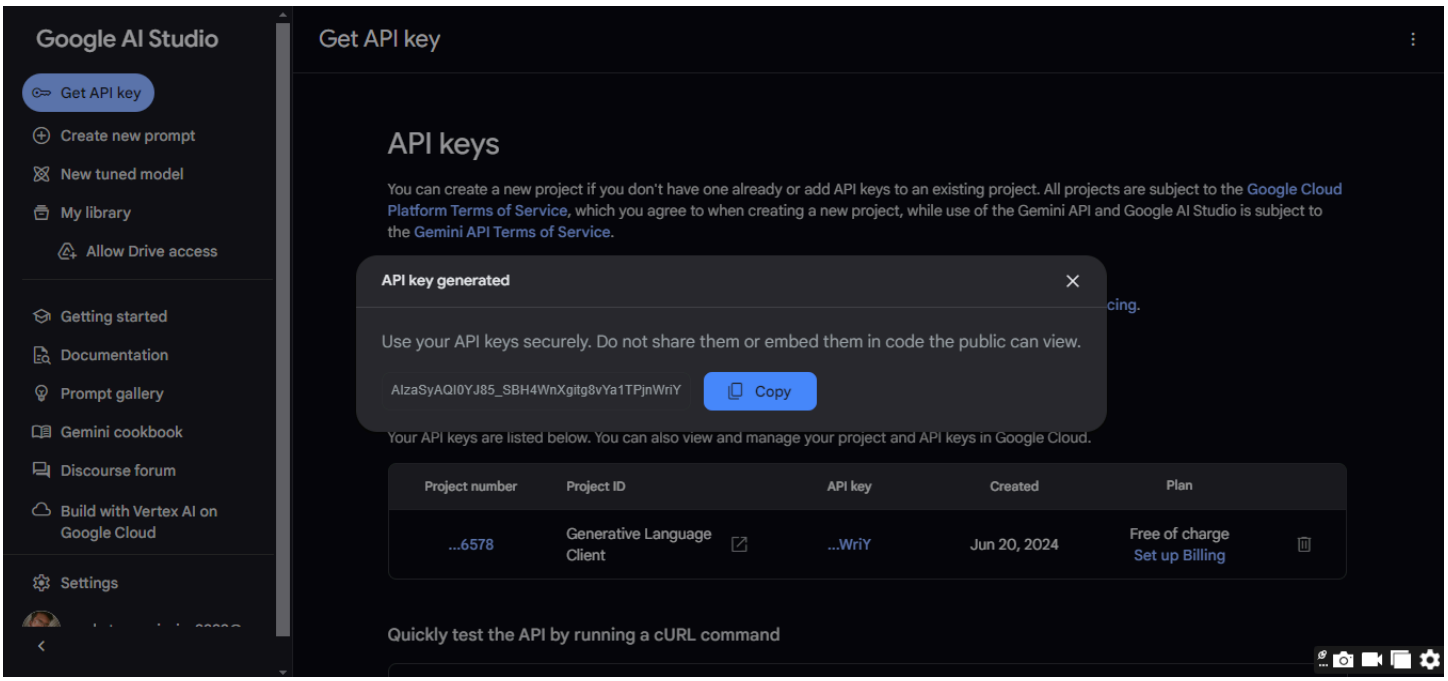
- Python: The primary programming language used for the entire application.
- Streamlit: Streamlit is used to create the web application interface for interacting with the user. It simplifies the process of building interactive web apps for data science and machine learning.
- dotenv: The dotenv library is used to load environment variables from a .env file. This is typically used to store sensitive information like API keys securely**.**
- googletrans: This library provides a Python wrapper around Google Translate API. It allows for language translation, which in this case, is used to translate the summary text into different languages based on user selection.
- youtube_transcript_api: This library allows fetching transcripts of YouTube videos. It's used to extract the transcript text from a YouTube video URL provided by the user.
- google.generativeai (presumably GenAI): This library seems to be part of a service (possibly GenAI) that provides access to generative models like Gemini Pro. It is used for generating summaries or content based on given prompts and input text.

2.2NLP Techniques Used:

- Text Summarization: Using a generative model (Gemini Pro) to summarize video transcripts.
- Machine Translation: Translating the summary into different languages using Google Translate API.
- Text Extraction: Extracting textual data (transcripts) from YouTube videos.
- Content Generation: Generating textual content based on a prompt using a deep learning model.
- Environment Management: Securely managing and accessing API keys and other sensitive information**.**

API Key Generation:



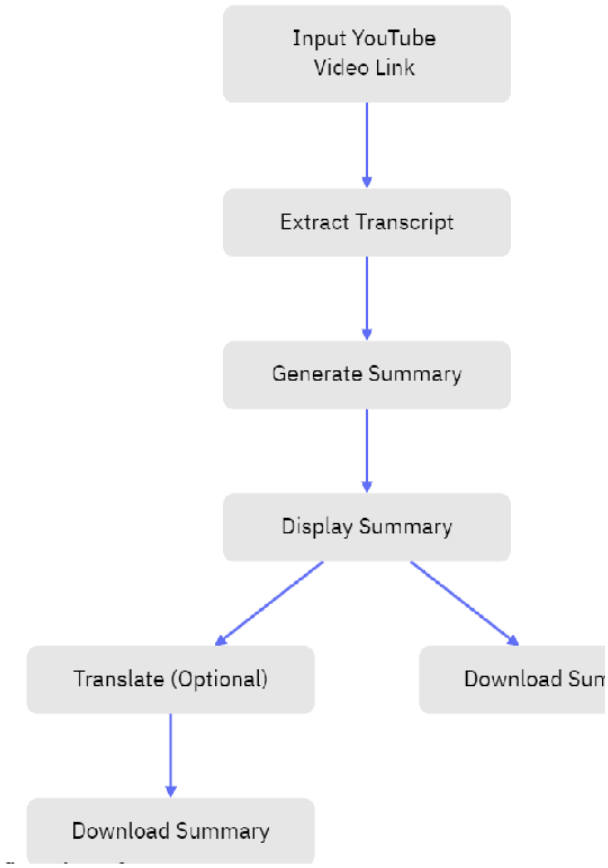


CHAPTER 3

3.1 Proposed Method

- Users input a YouTube video link in a Streamlit interface.
- YouTubeTranscriptApi extracts the video's transcript.
- Google's Gemini Pro model generates a brief summary based on a predefined prompt.
- Users can view the summary in the original language or translate it using Google Translate API.
- Summarized content can be downloaded for offline use, ensuring convenient access to distilled information from YouTube videos.

3.2 WORK FLOW OF THE SYSTEM:

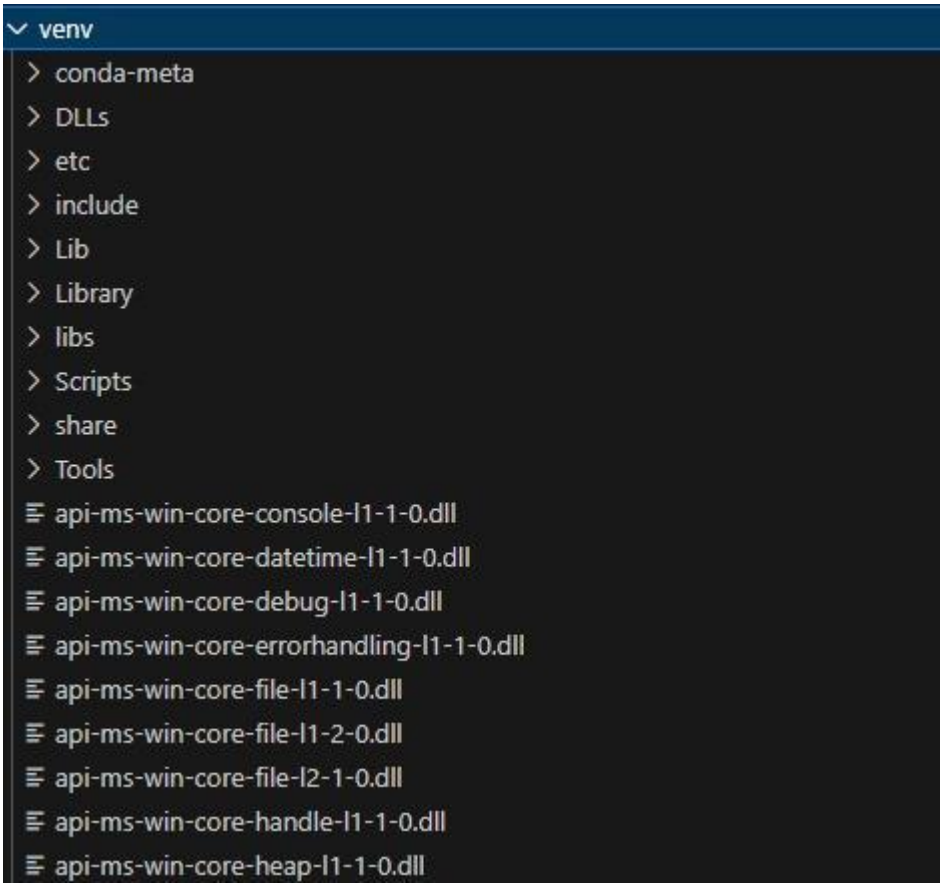


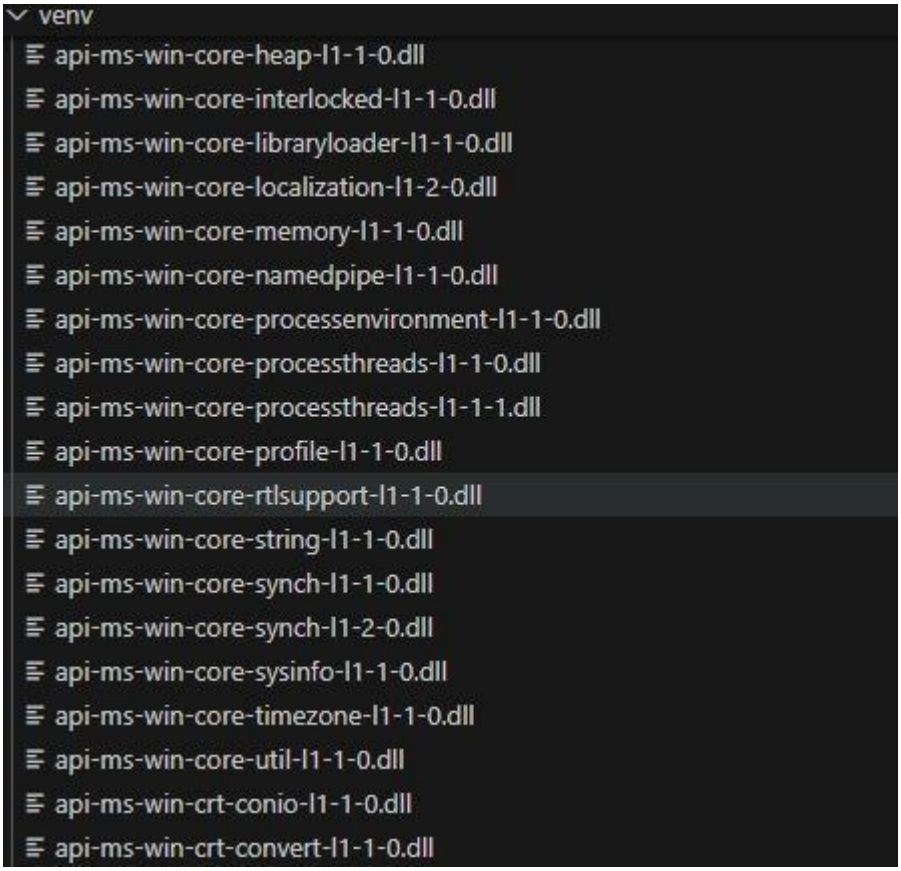
4. IMPLEMENTATION

4.1 REQUIRED LIBRARIES AND PACKAGES:

requirements.txt

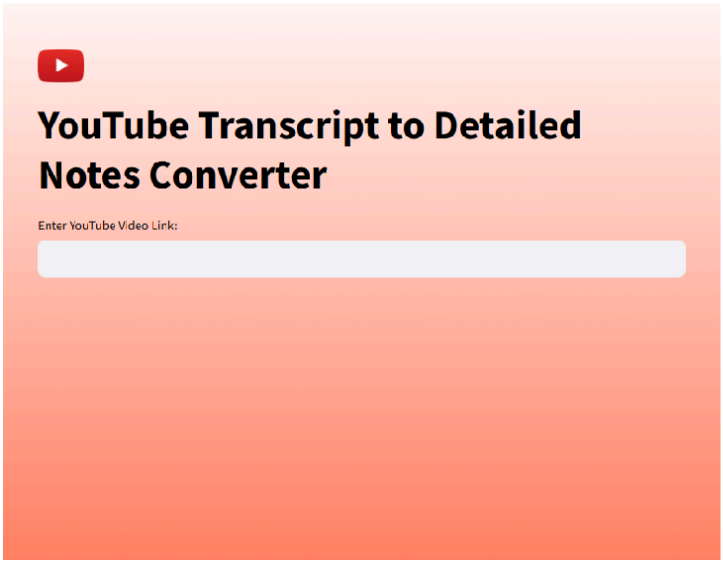
```
youtube_transcript_api
streamlit
google-generativeai
python-dotenv
pathlib
googletrans==4.0.0-rc1
google-cloud-translate
```

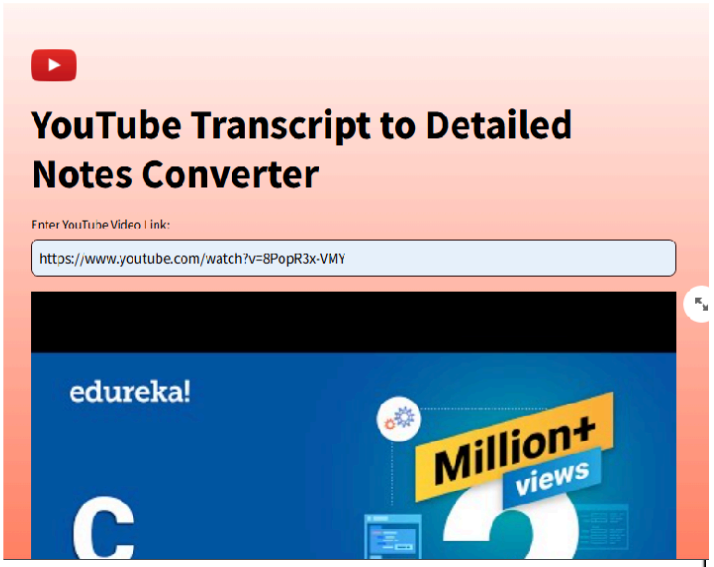


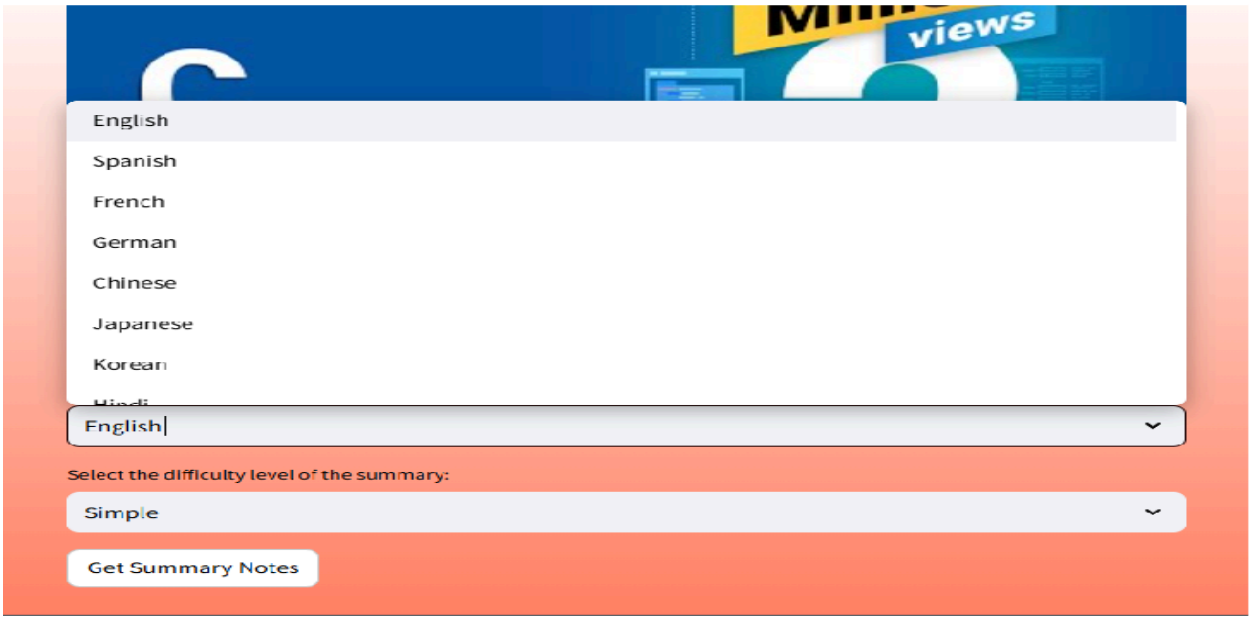


Google API Key

4.4 WORKING OF STREAMLIT WEB INTERFACE







Simple

Medium

Hard

Simple

Get Summary Notes

CHAPTER 5

1.1 Conclusion

- In conclusion, This Streamlit application efficiently integrates APIs to offer users a streamlined experience for summarizing and translating YouTube video transcripts. It combines YouTubeTranscriptApi for fetching transcripts, Google's Gemini Pro for generating summaries, and Google Translate for translating summaries into multiple languages. With a user-friendly interface and a downloadable summary feature, the application simplifies the process of accessing concise video summaries changed to user preferences.

1.2 Future Scope

- Implement speech-to-text for video summarization instead of using transcribers.
- Develop it as a website accessible to all users.
- Add a feature that allows users to find videos by searching keywords or video titles.
- Enable customization of our user interface by allowing users to define their preferred languages.

1.3 REFERENCES

- GitHub Repository YTtranscriber (GitHub)
- Video summarization: A conceptual framework and survey of the state of the art. Journal of Visual Communication and Image Representation. Volume 19, Issue 2,2008. Pages 12 l Arthur G. Money. Harry Agios. 143. ISSN 1047-3203.
- Youtube Video :End To End Youtube Video Transcribe Summarizer LLM App With Google Gemini Pro