Video Synopsi

A Mini-Project Report Submitted
For
Partial Fulfillment of the Requirements of the
Degree of Bachelor of Engineering

In

COMPUTER ENGINEERING

(Semester VII)
By

Aditya Vyas 9238

Hitesh Sharma 9233

Atharva Pawar 9427

Under the guidance of

Prof. Supriya Kamoji

DEPARTMENT OF COMPUTER ENGINEERING
Fr. Conceicao Rodrigues College of Engineering
Bandra (W), Mumbai - 400050
University of Mumbai

2023-2024

This work is dedicated to my family.

I am very thankful for their motivation and support.

CERTIFICATE

This is to certify that the mini-project entitled **"Video Synopsi"** is a bonafide work of " Aditya Vyas (9238)**,** Hitesh Sharma (9233), Atharva Pawar (9427) **"** submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of **Bachelor of Engineering** in **Computer Engineering** (Semester-VII).

Prof. Supriya Kamoji
(Name and Sign)

Guide/ Supervisor

Dr. Sujata Deshmukh

Dr. S. S. Rathod

Approval Sheet

Mini Project Report Approval for B.E. (Semester-VII)

This mini-project report entitled Video Synopsi submitted by Aditya Vyas (9238)**,** Hitesh Sharma (9233), Atharva Pawar (9427) is approved for the degree of Bachelor of Engineering in **Computer Engineering** (Semester-VII).

Examiner 1._____

Examiner 2._____

Date:

Place:

Declaration

We declare that this written submission represents our ideas in our own words and where others' ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Date:

Aditya Vyas (9238)

Hitesh Sharma (9233)

Atharva Pawar (9427)

Abstract

With the exponential growth of online video content, the need for efficient video summarization techniques has become increasingly evident. This paper presents a web application that leverages state-of-the-art deep learning methods to automatically generate textual summaries of videos. The application offers a user-friendly interface for uploading videos, and within seconds, it delivers concise, coherent, and contextually relevant textual summaries that encapsulate the essential content of the video.

We employ cutting-edge pre-trained models and transfer learning to enhance the system's performance, making it adaptable to a wide range of video genres and languages..

The system's capability to capture key events, sentiments, and themes in the video content makes it valuable for a broad spectrum of applications, including content curation, video indexing, and accessibility for individuals with visual impairments.The application's user-friendly design and seamless integration with web-based platforms make it accessible to a wide audience.

This paper showcases the development and evaluation of our web application for automatic video summarization, highlighting its potential to simplify video content consumption and promote a more efficient and accessible digital media landscape.

Keywords:

Video Summarization, accessibility., web application, automatic summarization

Acknowledgments

We have great pleasure in presenting the report on “Video Synopsi”. I take this opportunity to express my sincere thanks towards the guide Prof. Supriya Kamoji, C.R.C.E, Bandra (W), Mumbai, for providing the technical guidelines, and the suggestions regarding the line of this work. We enjoyed discussing the work progress with him/her during our visits to the department.

We thank Dr. Sujata Deshmukh, Head of Computer Engineering department, Principal and the management of C.R.C.E., Mumbai for encouragement and providing necessary infrastructure for pursuing the project.

We also thank all non-teaching staff for their valuable support, to complete our project.

Date:

Aditya Vyas (9238)

Hitesh Sharma (9233)

Atharva Pawar (9427)

Table of Content

Chapter 1

INTRODUCTION

In the age of digital information, the abundance of video content has transformed the way we consume media. However, this wealth of visual information often presents a daunting challenge, requiring users to wade through extended video content to find relevant information. In response to this challenge, video summarization has emerged as a vital tool for distilling video content into manageable, informative snippets.

This paper introduces an innovative approach to video summarization that leverages advanced machine learning techniques, along with APIs for text conversion. Our web application seamlessly integrates these technologies to provide users with concise textual summaries of videos, transcending language and content type barriers. This fusion of API-driven text conversion and large language models empowers the system to offer comprehensive video summarization, efficiently transforming visual content into textual information.

Central to our methodology are large language models, renowned for their natural language understanding and generation abilities. These models have demonstrated remarkable capabilities in extracting and generating text, and, by employing their deep neural networks, our system excels at capturing the nuances, sentiments, and essential themes embedded within video content.

The web application's user-friendliness stands out as a defining feature, allowing users from diverse backgrounds to easily upload videos and obtain coherent, concise textual summaries.

The seamless integration of APIs for text conversion, large language models, and machine learning in our web application positions it as an innovative solution in the field of video summarization, with the promise to transform the way we engage with and understand video content in the digital age.

In the following sections, we delve into the technical details of our approach, the experimental results, and the potential applications of our web application. The integration of APIs and advanced machine learning techniques represents a significant leap forward in the field of video summarization, with implications that extend to content discovery and accessibility on a broader scale.

Chapter 2

OBJECTIVES OF THE PROJECT

Our project has following objectives:

- Develop an automated video summarization web application for efficient content consumption.
- Prioritize user-friendliness and customization, allowing users to tailor their summaries.
- Ensure versatility by accommodating diverse video genres .
- Enhance accessibility through features like textual descriptions and subtitles

Chapter 3

SCOPE OF THE PROJECT

Following points summarize the scope of our project :

1. Develop a user-friendly web application for summarizing diverse video content.
2. Provide customization options for users to tailor their summaries.
3. Support multiple languages and include accessibility features like subtitles.
4. Integrate with external APIs to enhance text conversion and language processing.
5. Ensure security, scalability of our system.

Chapter 4

REVIEW OF LITERATURE

Sr. No.	Paper Title	Citation	Database	Summary
1	Review of automatic text summarization techniques & methods	133	Science Direct	The text discusses text summarization, highlighting extractive and abstractive approaches. It notes the shift towards abstractive summarization and the lack of comprehensive research despite achievements in the field. The paper presents a systematic review of 85 publications from 2008 to 2019, providing insights into research trends, methods, and challenges in text summarization.
2.	Exploring the Limits of ChatGPT for Query or Aspect-based Text Summarization	53	Arxiv	Text summarization is a key challenge in natural language processing. ChatGPT's potential for practical applications beyond news summarization is underexplored. Our evaluation found ChatGPT's performance comparable to traditional methods across diverse datasets, revealing unique characteristics in its summaries compared to human references. Further research is planned.
3.	Deep reinforcement and transfer learning for abstractive text summarization: A review	56	Elsevier	Automatic Text Summarization (ATS) in Natural Language Processing (NLP) condenses lengthy texts by emphasizing key points for readability. State-of-the-Art (SotA) techniques like deep neural models, Reinforcement Learning (RL), and Transfer Learning (TL) with Pre-Trained Language Models (PTLMs) have significantly improved abstractive ATS. This study offers a comprehensive six-year review of abstractive text summarization, covering challenges, solutions, datasets, and future research directions.
4.	Recent Challenges and Opportunities in Video Summarization with Machine Learning Algorithms	0	ieee explore	Video summarization is essential for efficiently managing the vast amount of digital content, allowing users to access key highlights from various domains. It offers both static and dynamic summaries, with the former using keyframes and the latter incorporating relevant video segments. This comprehensive process enhances the accessibility of information and streamlines video browsing and retrieval.

Chapter 5

PROPOSED SYSTEM

5.1 DRAWBACKS OF EXISTING SYSTEM

Following points summarizes drawbacks of existing system:

- Limited context understanding.
- Overreliance on extractive methods
- Inadequate evaluation metrics.
- Accuracy and coherence challenges.
- Limited accessibility and inclusivity.

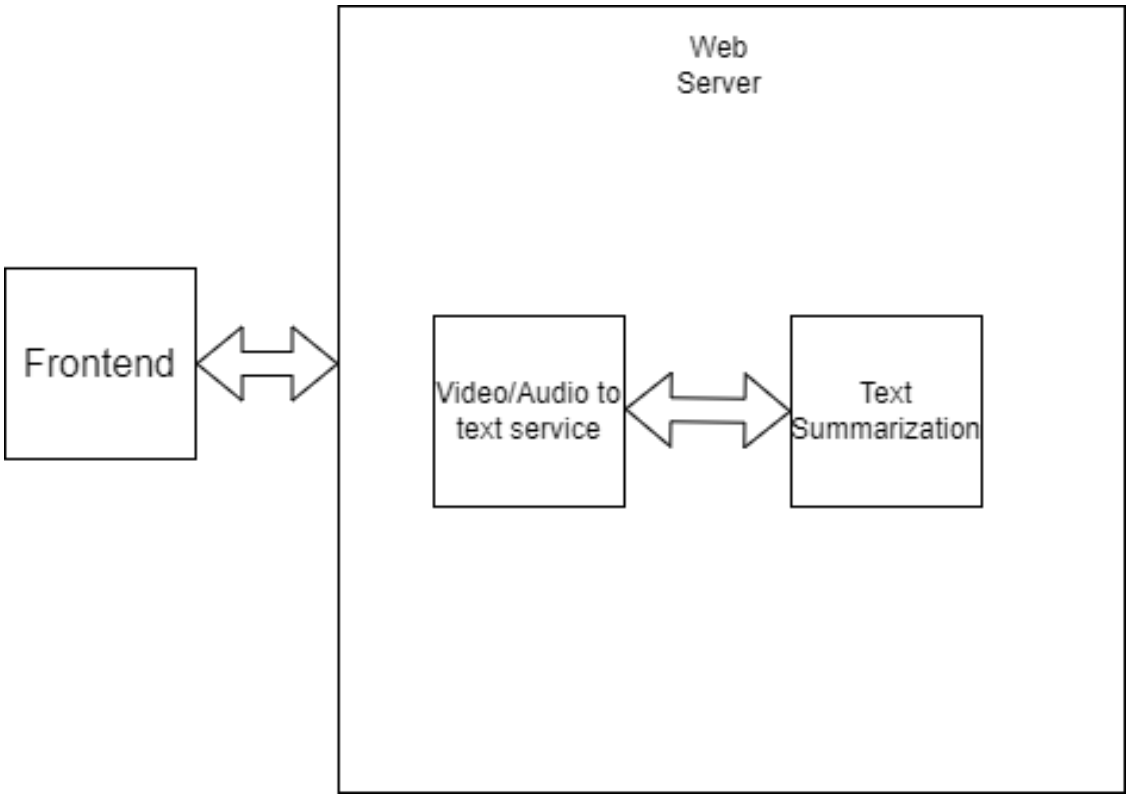
5.2 PROBLEM STATEMENT

Existing video summarization systems suffer from limitations in context understanding, linguistic versatility, customization, scalability, and accessibility. These shortcomings impede efficient access and comprehension of the vast digital video content available across diverse domains and languages. A need exists for an improved video summarization system to enhance the user experience.

Chapter 6

SYSTEM DESIGN

6.1 MODULE DESIGN



6.2 MODULE DESCRIPTION

In Video Synopsi system, we have following modules:

Frontend:

Frontend is main entry point of our application. The user can enter the video link(youtube) or upload video easily through user friendly interface

Web Server:

The request gets forwarded to our web service which has following pillars :

Video to Text conversion API:

This module contain API calls for converting video/audio to text document

This module is at the heart of the system, as most of the data processing happens here. The cloud has services deployed to facilitate respective functionalities. Following services are there in the system namely IoT service, Crop recommendation service, Crop Disease Detection Service and E-commerce service.

The IoT service facilitates the data flow from hardware to related services and data storage. Each service contains API endpoints which shall be used at client side to render relevant information. The crop-recommendation services and crop disease detection use trained models for prediction results and these results can be used by client-side via API endpoints created.

6.3 SOFTWARE AND HARDWARE USED

Software:

Operating System	Microsoft Windows 7 or higher/ Similar Linux System
Tool	Appropriate IDE like VS Code, Android Studio etc
Coding language	JavaScript , HTML, CSS, React, Next,Python
Storage	8 GB RAM and minimum 10 GB free space
Processor	Intel core I5 or higher
Other Requirements	Good Internet connectivity

Table 6.3 Software Requirements

Chapter 7

IMPLEMENTATION

Framework Choice: The project utilized the Django web framework for the entire development cycle. Django is well-suited for web application development and offers features for rapid development, security, and scalability.

Large Language Models (LLMs): The project leveraged Large Language Models (LLMs), specifically LLama 2, for various text-related tasks. These models are pre-trained on massive text corpora and provide a powerful foundation for natural language understanding and generation.

Backend Development: The backend of the web application was implemented using Django, encompassing data storage, APIs, and business logic. Django's robust ORM (Object-Relational Mapping) facilitated efficient data handling.

User Interface: The frontend was designed using Django's templating system, which allows for the creation of dynamic and user-friendly web interfaces. User interactions were implemented using HTML, CSS, and JavaScript.

Data Storage: Django's built-in support for relational databases was employed to store user data, video information, and summarization results. This ensured data consistency and reliability.

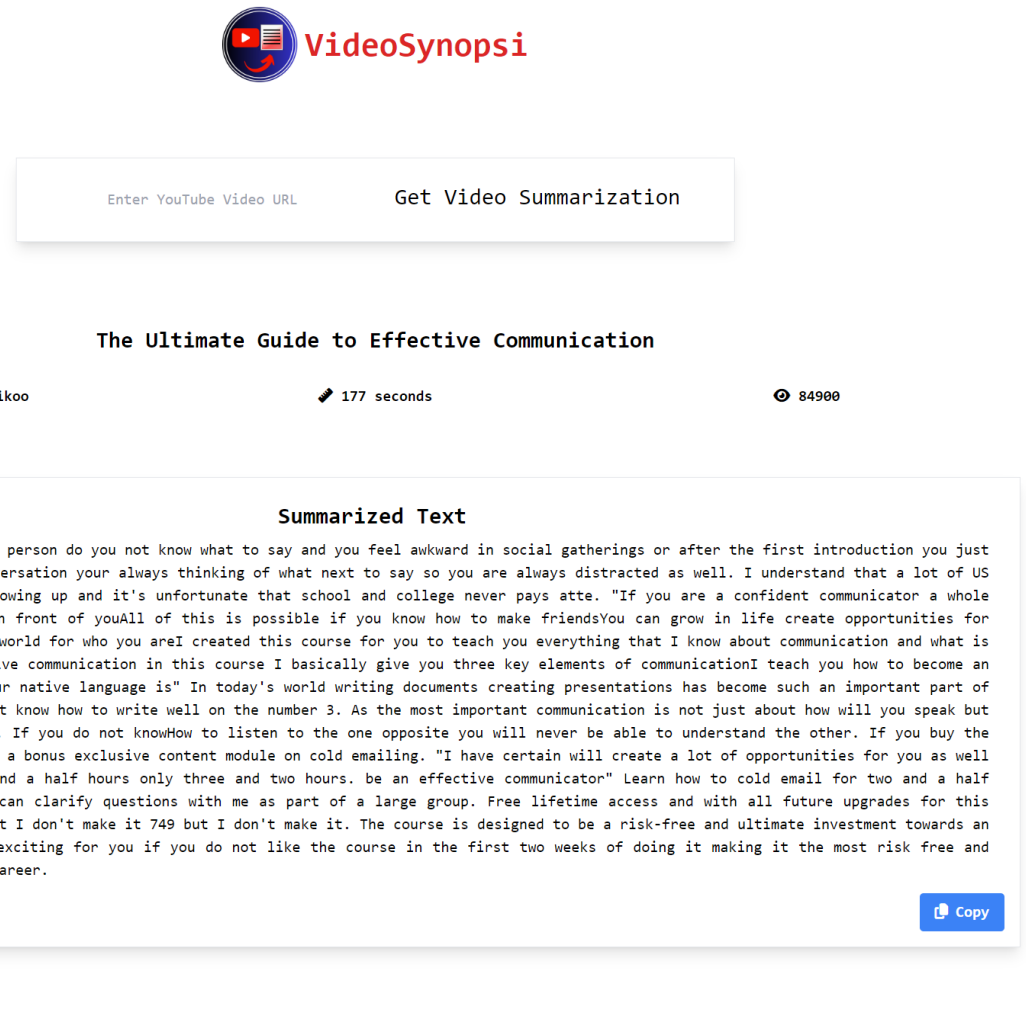
Summarization Engine: LLama 2, a Large Language Model, was integrated into the system to perform the video summarization. It processed video content and generated textual summaries using its powerful natural language processing capabilities.

Chapter 8

RESULTS

The project successfully implemented a web application using the Django framework, providing users with a platform for video summarization. The integration of the LLama 2 Large Language Model facilitated the generation of textual video summaries. The application offered features are user-friendly interface, enhancing accessibility to summarized video content.

8.1 WebApp



Show Raw Video Text

Raw Text:

are you an introvert or a shy person do you not know what to say and you feel awkward in social gatherings or after the first introduction you just completely words or in a conversation your always thinking of what next to say so you are always distracted as well I get it I understand that a lot of US I want to that when I was growing up and it's unfortunate that school and college never pays attention in teaching how to become an effective communicator because if you effective communicator if you are a confident communicator a whole host of opportunities open in front of you you can know how to make friends you can know how to you can know how to get growth you can grow in life create opportunities for yourself get respected in the world for who you are all of this is possible if you know how to become a good communicator and that's why I created this course for you to teach you everything that I know about communication and what is the ultimate guide to effective communication in this course I basically give you three key elements of communication I teach you how to become an effective speaker of what your native language is I also didn't teach you how to become an effective writer in today's world writing documents creating presentations has become such an important part of our growth that you cannot not know how to write well on the number 3 but as the most important communication is not just about how will you speak it's not just about how will you right it's a lot more about how well do you listen and if you do not know how to listen to the one opposite you will never be an effective communicator so when you combine these three things of speaking writing and listening you get the ultimate guide to active communication skill if you buy the 749 premium plan you also get a bonus exclusive content module on cold emailing Industries my life and have certain will create a lot of opportunities for you as well so what we have with three and a half hours only three and a half hours to become an effective communicator we have a bonus content on cold emailing we have PDF notes to revise everything with exercises to own your skills and whatever you learnt and then you have live class where you can clarify questions with me as part of a large group all of this with lifetime access and with all future upgrades for this course absolutely free 749 but I don't make it even more exciting for you if you do not like the course in the first two weeks of doing it making it the most risk free and ultimate investment towards an effective

GET Api

API:

API Endpoint
Endpoint:
Method: GET

Request Parameters
Required api_key (string):
-- An API key for authentication. Use api123 for testing purposes.
Required url (string):
-- The URL of the video you want to process.

Example Request
GET /apiVideo2Text?api_key=api123&url=https://www.youtube.com/watch?v=ankpGxGh8cA

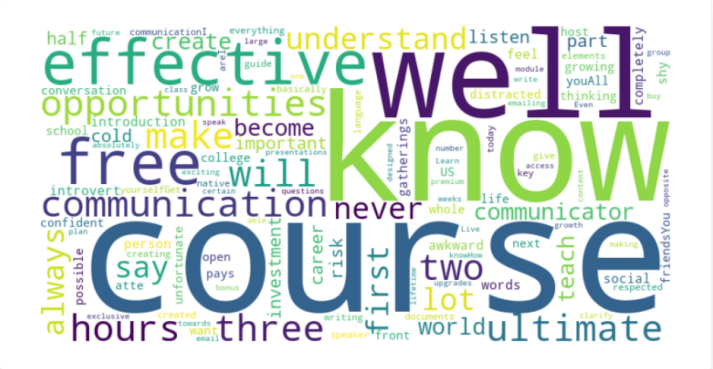
Response
JSON Object:


```
{
"title": "Unit 1- Lesson 7 - How To Pronounce the S in Plural Nouns - Pronunciation - Beginners Level",
"url": "https://www.youtube.com/watch?v=ankpGxGh8cA",
"author": "TEFLship",
"views": 334,
"length": 72,

"Raw_text": "hello this time we got pronunciation yes let's pronounce these words together we got three sounds to pronounce the letter S at the end of the plural nouns now let's listen to them and repeat the words tape script 1.14 pronunciation books students cause computers hamburgers Cambridge televisions bags phones some images houses buses wait for more units more lessons for the beginners English level and more levels coming soon thank you ",

"summarized_text": "14 pronunciation books students cause computers hamburgers Cambridge televisions bags phones some images houses buses wait for more units more lessons for the beginners English level an an English for adults level 1 and 2. 1. Pronunciation books: 1. The sound of the letter S at the end of the plural nouns now let's listen to them and repeat the words. 2. The sounds of the letters A, E, and F at the beginning of the nouns.\nMore levels coming soon thank you. d more levels comingSoon. d more level coming soon. Thank you for your support and support. D more levels coming soon thanks for support. d more levelsComing soon. More levels coming Soon."
}
```


Word Cloud



GITHUB Google Colab

8.2 GitHub Repo:

VideoSynopsi

VideoSynopsi is a NLP subject mini project for Fr CRCE - BE Sem - 7

Team

```
- Hitest Sharma (Back-END :  Video to Text Conversion)
- Aditya Vyas   (Back-END :  Research & LLM for summarization)
- Atharva Pawar (Front-END:  Flask Api & Web-App)
```

This project is designed for text summarization of YouTube videos. It leverages the capabilities of the SpeechRecognition, moviepy, pydub, and pytube libraries for efficient text analysis and processing. The summarization algorithm employed in this project effectively extracts key sentences from the input text, resulting in a concise and informative summary of the video's content.

Installation

```
Project Folder : https://github.com/capstone-project-SECURIX/VideoSynopsi_.git
```

Installation

```
Project Folder : https://github.com/capstone-project-SECURIX/VideoSynopsi_.git
```

1. Clone the repository:

```
git clone https://github.com/capstone-project-SECURIX/VideoSynopsi_.git
cd VideoSynopsi_
```

2. Install Dependencies:

```
pip install -r requirements.txt

#OR

# pip install cmd
pip install SpeechRecognition moviepy pydub pytube
```

Usage

Video to Text Summarization:

The VideoSynopsi_ takes an youtube video link (URL) as input and returns a summary of that Video.

VideoSynopsi_ (Video to Text) API

This API allows you to extract text data from a video using a provided API key and URL. It's designed to provide video-to-text summarization capabilities.

API Endpoint

- **Endpoint:** `/apiVideo2Text`
- **Method:** GET

Request Parameters

- `api_key` (string, required): An API key for authentication. Use `api123` for testing purposes.
- `url` (string, required): The URL of the video you want to process.

Usage:

```
1. Run the Flask app:
    python app.py

2. Send a GET request to http://127.0.0.1:5000/apiVideo2Text?api_key=api123&url=https://www.youtube.com/watch

3. The API will return the summarized text and other relevant information.
```

Google Colab

`https://colab.research.google.com/drive/1_aTfavawuo8sE8waewwT-18BvTmel-Dc#scrollTo=SIB_bYdhF_qK`

Example Request:

- GET /apiVideo2Text?api_key=api123&url=https://www.youtube.com/watch?v=ankpGxGh8cA

Response

- If the API key is valid and the URL is accessible, the API will respond with a JSON object containing the following information:
  - `title` (string): The title of the video.
  - `url` (string): The URL of the video that was processed.
  - `author` (string): The author of the video.
  - `views` (integer): The view count of the video.
  - `length` (integer): The length of the video in seconds.
  - `raw_text` (string): A Raw text extracted from the video content.
  - `summarized_text` (string): A summarized text extracted from the raw_text.

Example Response:


```
{
    "title": "Unit 1- Lesson 7 - How To Pronounce the S in Plural Nouns - Pronunciation - Beginners Level",
    "url": "https://www.youtube.com/watch?v=ankpGxGh8cA",
    "author": "TEFLship",
    "views": 334,
    "length": 72,
    "Raw_text": "hello this time we got pronunciation yes let's pronounce these words together we got three s
    "summarized_text": "14 pronunciation books students cause computers hamburgers Cambridge televisions bags
}
```


Status Codes:

- `200 OK`: The request was successful, and the API has responded with the expected data.
- `401 Unauthorized`: The provided API key is invalid or missing. You need a valid API key for access.
- `500 Internal Server Error`: An error occurred during processing. This could be due to various reasons, such as video download failures or other unexpected errors.

Responses Time

- for `72 Second` long video it took `2 mins`.

Contributing

Contributions are welcome! If you have suggestions, improvements, or bug fixes, please create an issue or a pull request.

License

REFERENCES

[1] *Widyassari, A. P., Rustad, S., Shidik, G. F., Noersasongko, E., Syukur, A., & Affandy, A. (2022). Review of automatic text summarization techniques & methods. Journal of King Saud University-Computer and Information Sciences, 34(4), 1029-1046.*

[2] *Yang, X., Li, Y., Zhang, X., Chen, H., & Cheng, W. (2023). Exploring the Limits of ChatGPT for Query or Aspect-based Text Summarization. ArXiv, abs/2302.08081.*

[3] *Ayham Alomari, Norisma Idris, Aznul Qalid Md Sabri, Izzat Alsmadi,*
Deep reinforcement and transfer learning for abstractive text summarization: A review,Computer Speech & Language,
Volume 71,
2022,
101276,
ISSN 0885-2308,
https://doi.org/10.1016/j.csl.2021.101276.
(https://www.sciencedirect.com/science/article/pii/S0885230821000796)

[4] *P. Kadam et al., "Recent Challenges and Opportunities in Video Summarization With Machine Learning Algorithms," in IEEE Access, vol. 10, pp. 122762-122785, 2022, doi: 10.1109/ACCESS.2022.3223379.*

APPENDIX-II

<<Plagiarism Report>>