Annexure – 1

MINI PROJECT NAME

A Mini Project-I Report submitted in partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE & ENGINEERING

By

1.V.Harshita Sai [19B01A05A2]

2.M.Sarvani [19B01A0590]

3.K.Jaya Sravani [19B01A0571]

4.N.Rishika [19B01A05C0]

5.M.Harshitha [19B01A05A9]

Under the esteemed guidance of

Mr. A.Seenu M.Tech(Ph.D.)

Assoc.Professor

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

SHRI VISHNU ENGINEERING COLLEGE FOR WOMEN(A)

(Approved by AICTE, Accredited by NBA & NAAC, Affiliated to JNTU Kakinada)

BHIMAVARAM – 534 202

2020 – 2021

Annexure – 2

SHRI VISHNU ENGINEERING COLLEGE FOR WOMEN (A)

(Approved by AICTE, Accredited by NBA & NAAC, Affiliated to JNTU Kakinada)

BHIMAVARAM – 534 202

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that the Mini Project-1 entitled "**Youtube Transcript Summarizer**", is being submitted by **V.Harshita Sai** bearing the **Regd. No. 19B01A05A2** in partial fulfillment of the requirements for the award of the degree of "**Bachelor of Technology** in **Computer Science & Engineering**" is a record of bonafide work carried out by them under my guidance and supervision during the academic year 2020–2021 and it has been found worthy of acceptance according to the requirements of the university.

Internal Guide **Head of the Department**

Acknowledgements

The satisfaction that accompanies the successful completion of any task would be incomplete without the mention of the people who made it possible and whose constant encouragement and guidance has been a source of inspiration throughout the course of this project. I take this opportunity to express our gratitude to all those who have helped me in this project.

I wish to place my deep sense of gratitude to ***Sri. K. V. Vishnu Raju, chairman of SVES***, for his constant support on each and every progressive work of mine.

My deep sense of gratitude and sincere thanks to ***Dr. G. Srinivasa Rao, Principal of SVECW*** for being a source of an inspirational constant encouragement.

My deep sense of gratitude and sincere thanks to ***Dr. P. Srinivasa Raju, Vice-Principal of SVECW*** for being a source of an inspirational constant encouragement.

I deeply indebted and sincere thanks to ***Dr. P. Kiran Sree, Head of the Department***, for his valuable advices in completing this project successfully.

My deep sense of gratitude and sincere thanks to for his unflinching devotion and valuable suggestion throughout my project.

Project Associates

1.V.HARSHITA SAI (19B01A05A2)

2.M.SARVANI (19B01A0590)

3.N.RISHIKA (19B01A05C0)

4.M.HARSHITHA (19B01A05A9)

5.K.JAYA SRAVANI (19B01A0571)

CONTENTS

S.NO	UML DIAGRAMS	PAGE NO
1	Use Case Diagram	13-14
2	Class Diagram	15
3	Sequence Diagram	16
4	Activity Diagram	17-18
5	State Chart Diagram	19

INTRODUCTION

Introduction

We all interact with applications which uses text summarization. Many of those applications are for the platform which publishes articles on daily news, entertainment, sports. With our busy schedule, we prefer to read the summary of those article before we decide to jump in for reading entire article. Reading a summary help us to identify the interest area, gives a brief context of the story.

The main objective of a text summarization system is to identify the most important information from the given text and present it to the end users.

Think of the internet, comprised of web pages, news articles, status updates, blogs and so much more. The data is unstructured and the best that we can do to navigate it is to use search and skim the results.There is a great need to reduce much of this text data to shorter, focused summaries that capture the salient details, both so we can navigate itmore effectively as well as check whether the larger documents contain the information that we are looking for.

Enormous number of video recordings are being created and shared on the Internet through out 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.

We spend a noticeable amount of our weekly time watching YouTube videos, be it for entertainment, education, or exploring our interests. In most cases, the overall intent is to obtain some form of information from the video. We were seeking a solution to increase the efficiency of this "information extraction" process as YouTube's speed adjustment option is the only relevant tool. And so we decided to develop YouTube Transcript Summarizer!

Text summarisation is done in two ways

1.abstractive summarisation and

2.extractive summarisation

our project is based on extractive summarisation

Extractive Text Summarization :

It is the traditional method developed first. The main objective is to identify the significant sentences of the text and add them to the summary. You need to note that the summary obtained contains exact sentences from the original text.

As part of the summarization we can change the default parameters of the summarize function according to our requirements. one of the parameters is the ratio It can take values between 0 to 1. It represents the proportion of the summary compared to the original text.

SYSTEM ANALYSIS

System Analysis

2.1 Existing System

The summarizer sumarize the captions or transcript of any given video, it is able to pull the information and condense it into a paragraph. The exisiting system can take time to analyse the video and summarize it and can involve all the matter which is even irrelevant to the user.

Disadvantages

Time taking process.

It gives some irrelevant information.

2.2 Proposed System

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. So In the proposed system we used advanced online services to provide a summary of the description more accurately analyse and summarize the video more briefly.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. 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.

Advantages

It reduces reading time.

It improves the effectiveness of indexing.

It algorithms are less biased than human summarizers.

2.3 Feasibility Study

An important outcome of preliminary investigation is the determination that the system request is feasible. This is possible only if it is feasible within limited resource and time. The different feasibilities that have to be analyzed are

- Operational Feasibility
- Economic Feasibility
- Technical Feasibility

Operational Feasibility:

Operational feasibility deals with the study of prospects of the system to be developed. This system provides the summarizer of the youtube video using Youtube Transcript API. Based on the study, the system is proved to be operationally feasible.

Economic Feasibility:

Economic feasibility or Cost-benefit is an assessment of the economic justification for the computer based project. As hardware was installed from the beginning and for lots of purposes thus the cost on hardware is low. So, the project is economically feasible.

Technical Feasibility:

Technical Feasibility is the process of validating the technology assumptions, architecture and design of a product or project.. We used colab notebook for running the code. Thus, the project is technically feasible.

SYSTEM REQUIREMENTS SPECIFICATION

System Requirements Specification

3.1 Software Requirements

- IDE : Colab Notebook
- Language : Python
- Browser : Any type

3.2 Hardware Requirements

- Processor : intel I3
- RAM : 8GB
- Space on Hard Disk : 500GB

3.3 Functional Requirements

In Software Engineering, a functional requirement defines a function of a software system or its component. A function is described as a set of inputs, behavior, outputs. Functional requirements may be calculations, technical details, data manipulation and supposed to accomplish. The plan for implementing functional requirements is detailed in the system design. In requirements engineering, functional requirement specify particular results of a system. Functional requirements drive the application architecture of a system.

The following are the functional requirements of our system

- User should be able to process the video and generate the summarizer.
- The summarizer should contain main points and should have brief content according to the video.
- The summarizer should be easy to understand.

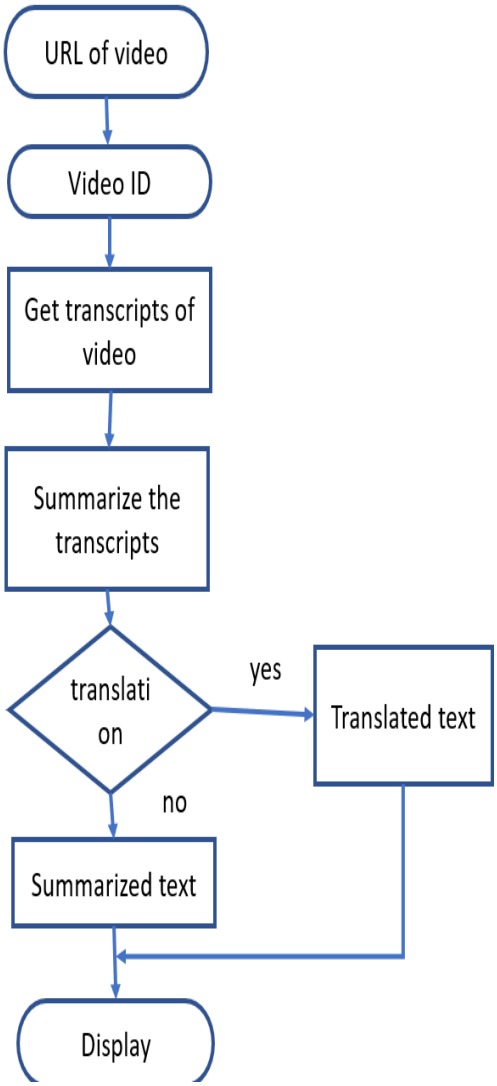
SYSTEM DESIGN

System Design

4.1 Introduction

Design is the first step in the development phase of an engineering product or system. Design is the place where quality is considered in the software development. Design is the only way that we can accurately translate user requirements into finished software product or system. Software design serves as the foundation for all the software engineers and software maintenance that steps follow. Without design we risk building an unstable design one that will fail when small changes are made, one that may be difficult to test and one whose quantity cannot be assessed until late in software engineering process.

The software of the project is mainly based on Hugging face transformers and NLP technique of abstractive text summarization . The video Id is taken from the URL given using split() . The transcripts for a given Youtube video Id is abstracted using a youtube_transcript_api API. The response from the Transcript API will return a list of dictionaries. A function Parse the JSON objects into strings utilize as a feed input for the NLP processor in the pipeline. Using genism package perform text summarization on obtained transcripts strings using Hugging Face Transformers pre_trained BART model to summarize the transcripts with given ratio by the user. Using the translator imported from GoogleTrans the summarized text is translated to userfriendly language which is convenient to user. On overall by using an url of an youtube video the summarized text is generated in user friendly language.


URL of video
Video ID
Get transcripts of video
Summarize the transcripts
translati on
yes
Translated text
no
Summarized text
Display

4.2 Data flow diagrams (UML Diagrams)

Introduction to UML

A model is an abstract representation of system, constructed to understand the system priority to building or modifying it. A model is a simplified representation of reality and it provides a means for conceptualization and communication of ideas in a precise and ambiguous form. We build models so that we can better understand the system we are developing. The elements are like components which can be associated in different ways to make a complete UML picture, which is known as diagram. Thus, it is very important to understand the different diagrams to implement the knowledge in real life systems.

UML (Unified Modeling Language) is a standard language for specifying, visualizing, constructing, and documenting the artifacts of software systems. It is a method for describing the system architecture in detail using the blueprint. We use UML diagrams to portray the behavior and structure of a system. This is the step while developing any product after analysis. The goal from this is to produce a model of the entities involved in the project which later need to be built. The representation of the entities that are to be used in the product being developed need to be designed.

There are various kinds of methods in software design:

- Use case Diagram
- Class Diagram
- Sequence Diagram
- Activity Diagram
- State Chart Diagram

4.2.1 Use Case Diagram

Use Case Diagrams are used to depict the functionality of a system or a part of a system. They are widely used to illustrate the functional requirements of the system and its interaction with external agents (actors).

A use case is basically a diagram representing different scenarios where the system can be used. A use case diagram gives us a high level view of what the system or a part of the system does without going into implementation details. When the initial task is complete, use case diagrams are modelled to present the outside view.

In brief, the purposes of use case diagrams can be said to be as follows

- Used to gather the requirements of a system.
- Used to get an outside view of a system.
- Identify the external and internal factors influencing the system.

Use case diagrams commonly contains

- Use cases
- Actors
- Dependency, generalization and association relationships.

Use cases

A use case is a software and system engineering term that describes how a user uses system to accomplish a particular goal.

Actors

An actor is a person, organisation or external system that plays a role in one or more interactions with the system.

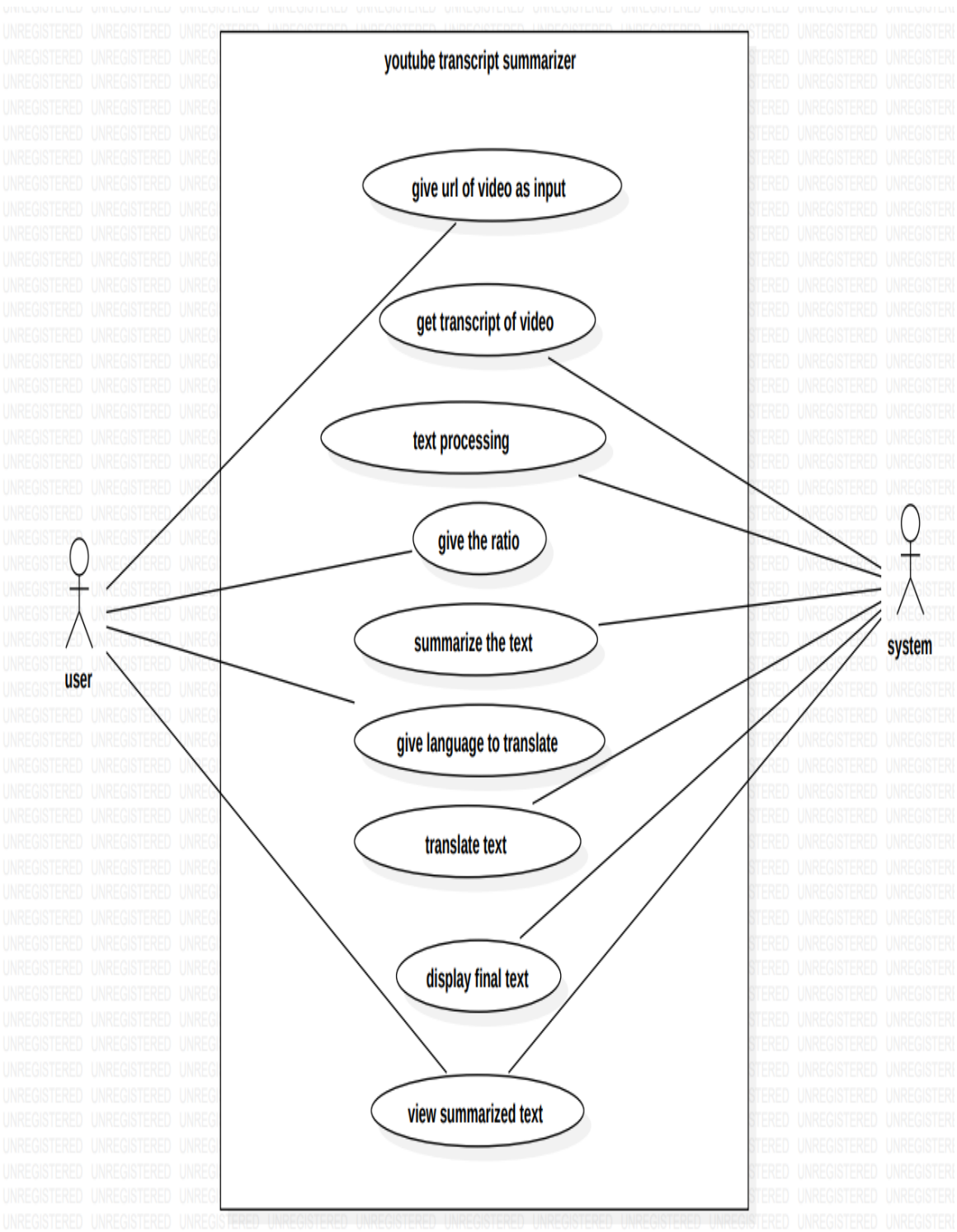


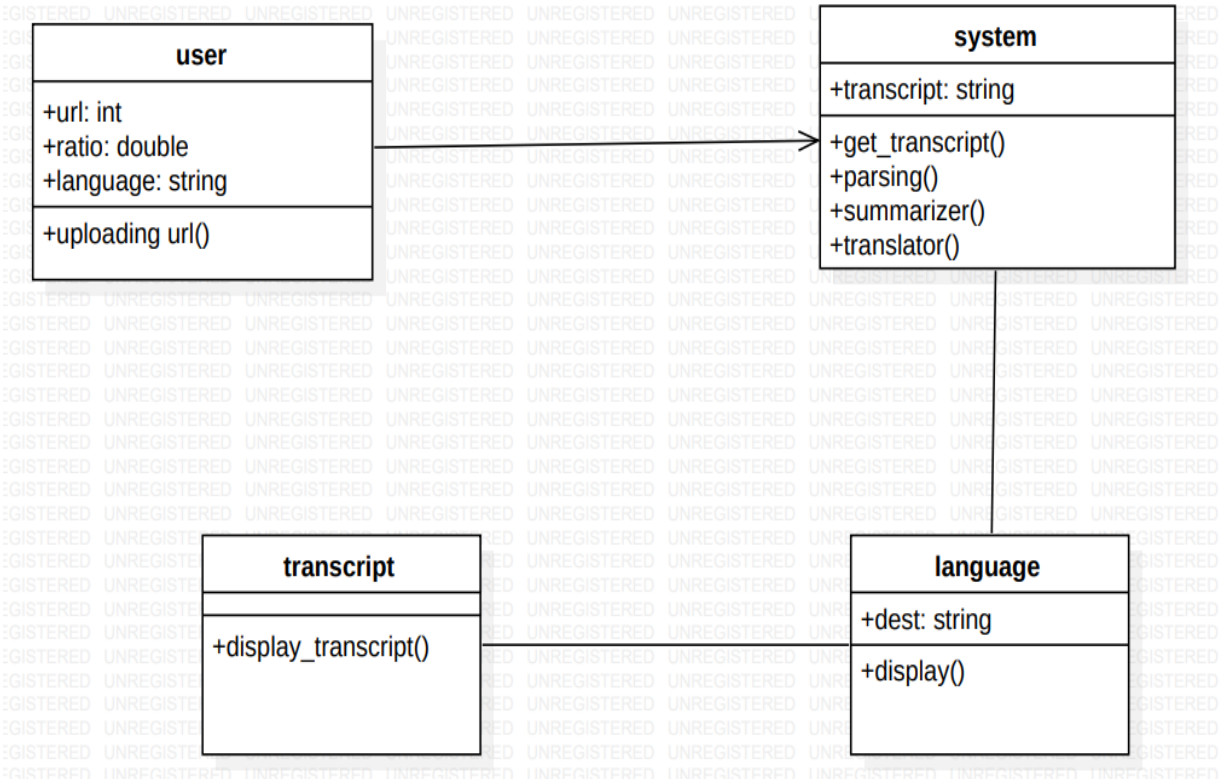
Figure 4.1 Use case Diagram

4.2.2 Class Diagram

Class diagram is a static diagram. It represents the static view of an application. Class diagram is not only used for visualizing, describing, and documenting different aspects of a system but also for constructing executable code of the software application. Class diagram describes the attributes and operations of a class and also the constraints imposed on the system. The class diagrams are widely used in the modelling of object oriented systems because they are the only UML diagrams, which can be mapped directly with object-oriented languages. It is also known as a structural diagram.

Class diagram contains

- Classes
- Interfaces
- Dependency, generalization and association.



4.2.3 Sequence Diagram

A sequence diagram simply depicts interaction between objects in a sequential order i.e. the order in which these interactions take place. We can also use the terms event diagrams or event scenarios to refer to a sequence diagram. Sequence diagrams describe how and in what order the objects in a system function. Sequence diagrams are used to formalize the behavior of the system and to visualize the communication among objects. These are useful for identifying additional objects that participate in the use cases. These diagrams are widely used by businessmen and software developers to document and understand requirements for new and existing systems.

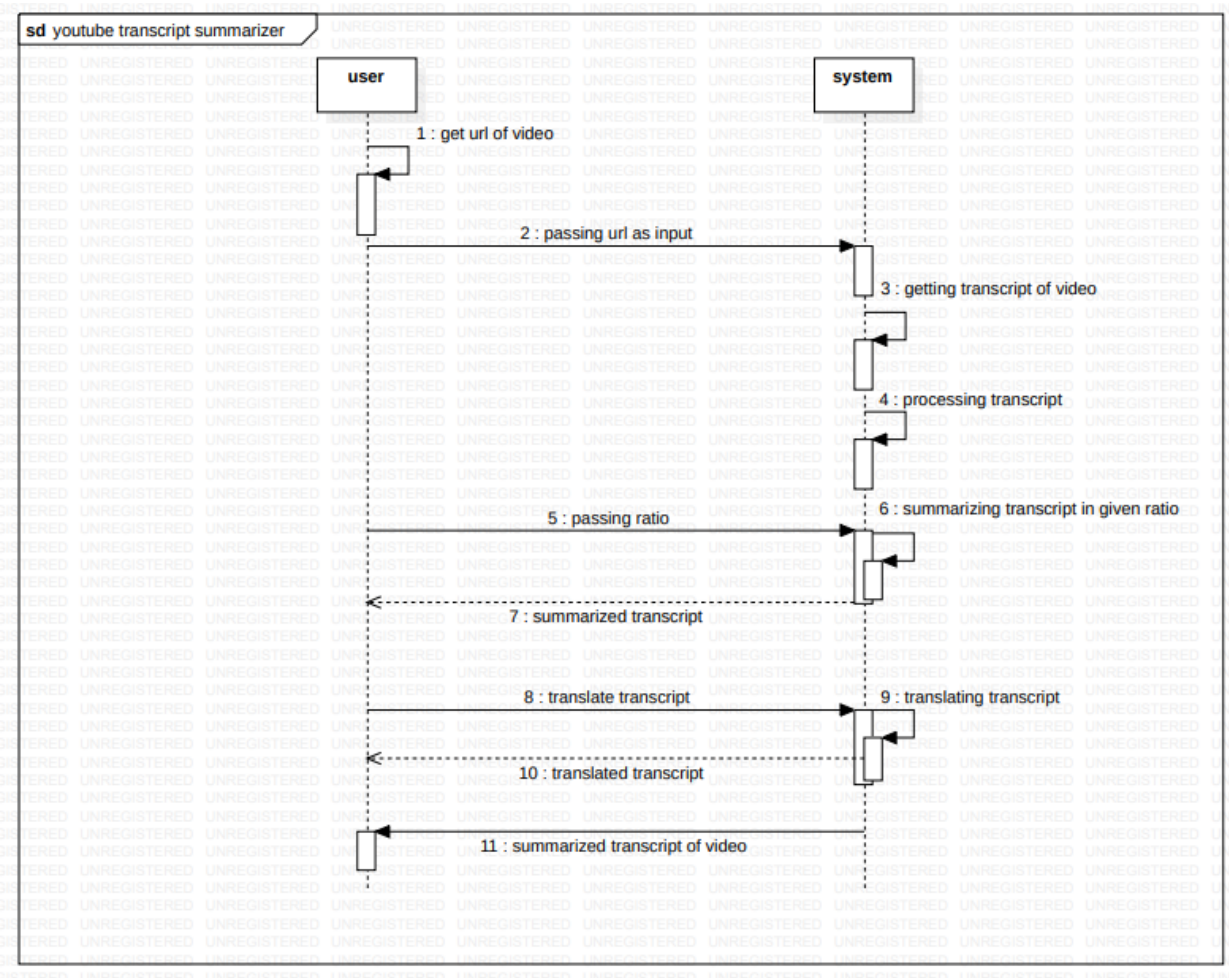


Figure 4.3 Sequence Diagram

4.2.4 Activity Diagram

Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system. The control flow is drawn from one operation to another. This flow can be sequential, branched, or concurrent. Activity diagrams deal with all type of flow control by using different elements such as fork, join, etc. The basic purposes of activity diagram is similar to other four diagrams. It captures the dynamic behaviour of the system. Other four diagrams are used to show the message flow from one object to another but activity diagram is used to show message flow from one activity to another.

Activity is a particular operation of the system. Activity diagrams are not only used for visualizing the dynamic nature of a system, but they are also used to construct the executable system by using forward and reverse engineering techniques. The only missing thing in the activity diagram is the message part. It does not show any message flow from one activity to another. Activity diagram is sometimes considered as the flowchart. Although the diagrams look like a flowchart, they are not. It shows different flows such as parallel, branched, concurrent, and single.

- Describe the sequence from one activity to another.
- Draw the activity flow of a system.
- Describe the parallel, branched and concurrent flow of the system.

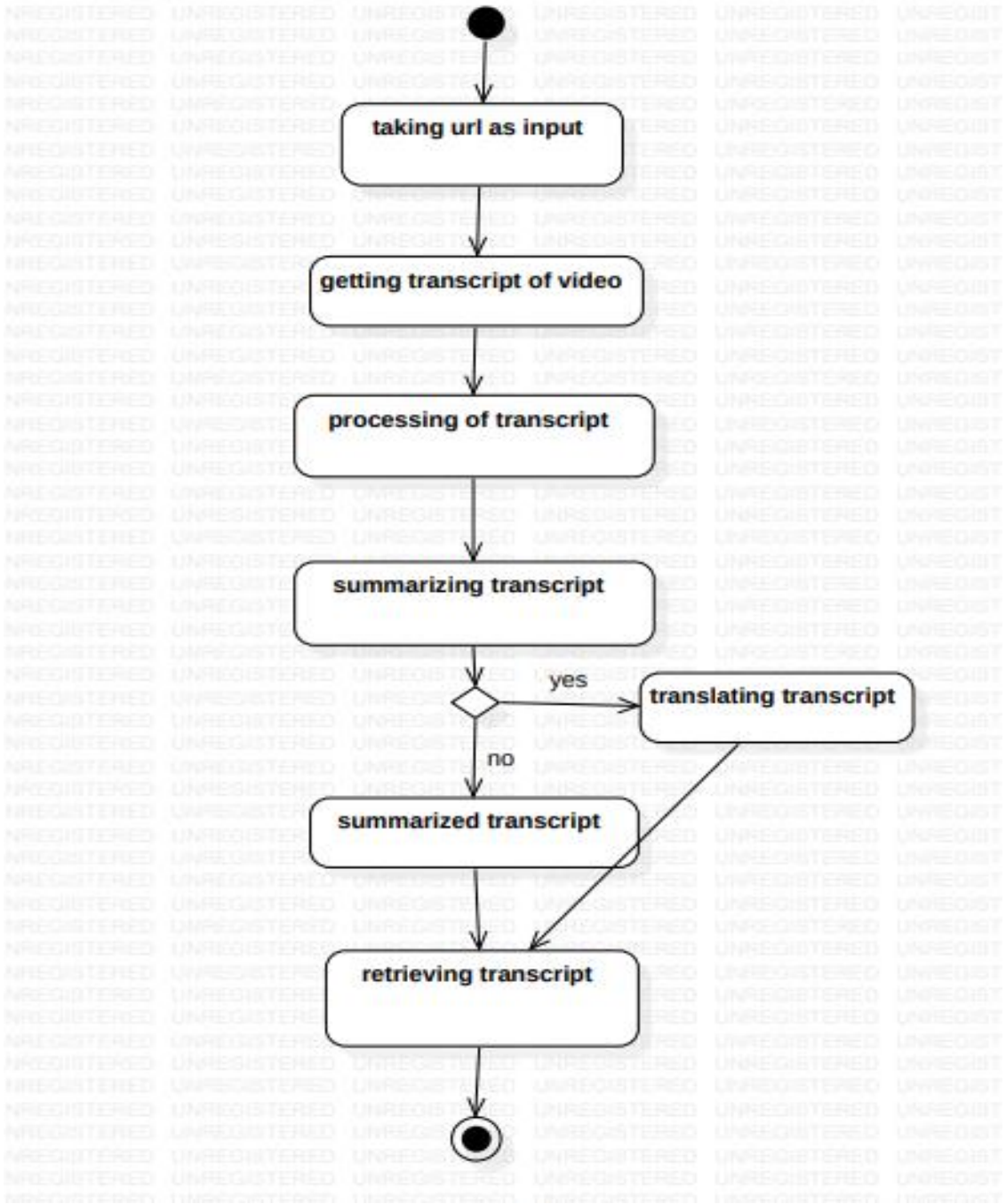
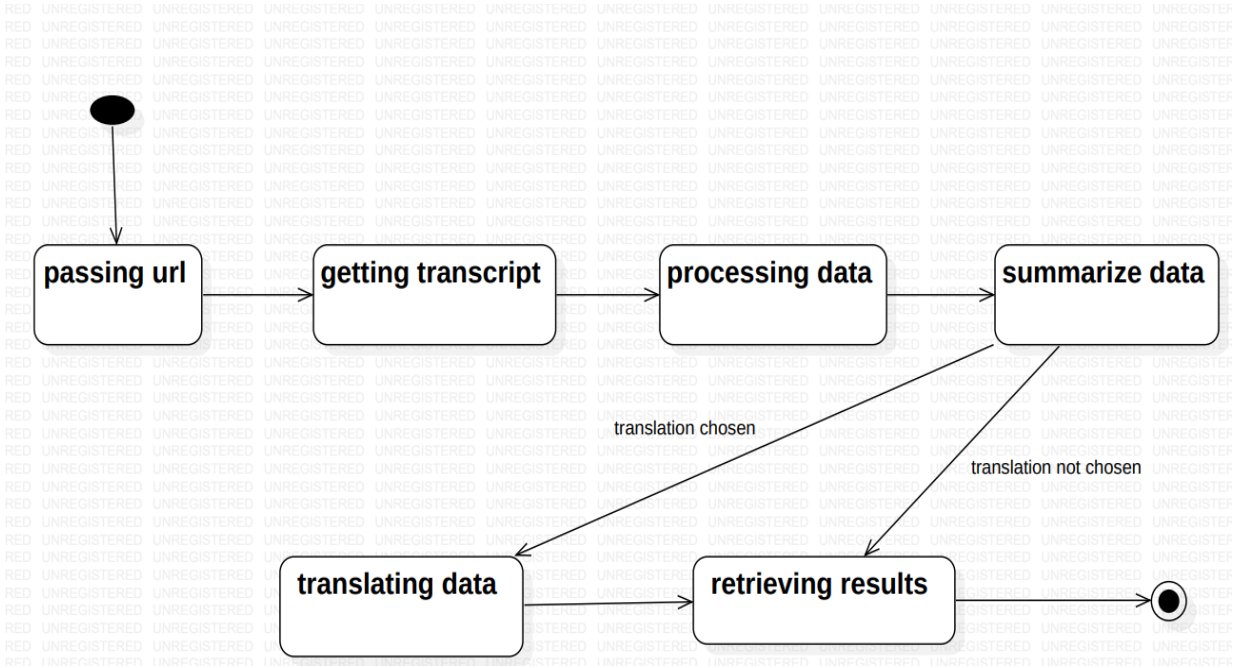


Figure 4.4 Activity Diagram

4.2.5 State Chart Diagram

State chart diagram describes the flow of control from one state to another state. States are defined as a condition in which an object exists and it changes when some event is triggered. The most important purpose of State chart diagram is to model lifetime of an object from creation to termination. State chart diagrams are also used for forward and reverse engineering of a system. However, the main purpose is to model the reactive system.



SYSTEM IMPLEMENTATION

System Implementation

5.1 Introduction

The purpose of system implementation can be summarized as follow making the new system available to the prepared set of users (the deployment), and positioning ongoing support and maintenance of the system within the performing organization (the transaction). At a finer necessary to educate the consumer on the use of system, placing the newly developed system into production, confirming that business functions that interact with the system and functioning properly. Transitioning the system support responsibilities involve changing from a system development to the system and maintenance mode of operation, with ownership of the new system moving from the project team to the performing organization.

A key difference between system implementation and all other phases of lifecycle is that all project activities up to this point have been performed in safe, protected and check your environments. It is through the careful planning, execution and management of system implementation activities that the project team can minimize the likelihood of these occurrences and determine appropriate contingency plans in the event of the problem.

Our project explores the summarization of the youtube video using youtube_transcript_api The input to our algorithm is a url of the youtube video and our model output is the summary of the video. We use hugging face transformers and NLP in the text summarization.

5.2 Project Modules

1. Extract Transcript text from the given YouTube Video
2. Summarizing the extracted Text (Transcript)
3. Summarize the summary
4. Convert the summary into other major languages

Extract Transcript text from the given YouTube Video:

1. Firstly, we would be giving the URL of the video assigning to youtube_video.

```
[ ]  youtube_video = "https://www.youtube.com/watch?v=aOL7wzEIZSc"
```

2. But it is not necessary of having the complete link we need the video id from that. So, we need to split it as the index value is 1 as assign it to the video_id . From this it would be extracting the video id that is needed.

```
[ ]  video_id = youtube_video.split("=")[1]
```

```
[ ]  video_id
```

```
'aOL7wzEIZSc'
```

3. from IPython.display import YouTubeVideo

Here ,basically we are showing what video we were trying to translate.

```
[ ]  from IPython.display import YouTubeVideo
     YouTubeVideo(video_id)
```

4.

```
[ ]  YouTubeTranscriptApi.get_transcript(video_id)
     transcript = YouTubeTranscriptApi.get_transcript(video_id)
```

This is going to extract the transcript as it is assigning to transcript.From this we would be obtaining a dictionary for list at every point where ever there is a transcript that means for every duration there is a speech (caption) it is going to be recorded and shown.

Now converting this into a pure text. We would be iterating through this object and everytime we are encountered text that is going to be taken and add it to the result object.

It takes the text at every point and it to the result with the space separator**.**

Summarizing the extracted Text (Transcript)

1.We have to summarize the text that we have got (large amount of text) in a meaning full way.We will be using the transformers summary pipeline that is imported from transformers.

```
[ ] summarizer = pipeline('summarization')
```

```
Downloading: 100% 1.80k/1.80k [00:32<00:00, 54.8B/s]
Downloading: 100% 1.22G/1.22G [00:32<00:00, 37.7MB/s]
Downloading: 100% 899k/899k [00:00<00:00, 3.03MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 606kB/s]
Downloading: 100% 26.0/26.0 [00:00<00:00, 229B/s]
```

Here we will be initializing a pipeline for summarization by default it is going to take …… model.Once we have this model we can simply summarize by using summerizer that may not work as we have got the large text by default would support this large text and tokens. So, we are creating chunks of the text from the input text that we got.

2.

```
min_length = 10
max_length = 50
num_iters = int(len(result)/1000)
summarized_text = []
for i in range(0, num_iters + 1):
  start = 0
  start = i * 1000
  end = (i + 1) * 1000
  out = summarizer(result[start:end], min_length, max_length)
  out = out[0]
  out = out['summary_text']
  summarized_text.append(out)

text1 = summarized_text

print(summarized_text)
```

We were going to declare the amount of iteration .At every iteration we will going to handle thousands of linesof text.

num_iters = int(len(result)/1000)

Now, initiating an empty list where we are going to have the summarized text.

summarized_text = []

And then iterate through the number we calculated for the total we got

```
for i in range(0, num_iters + 1):
```

We would be having start and end value

```
start = 0
start = i * 1000
end = (i + 1) * 1000
```

Now,from this start to end take the chunk of the text and add it to summarize it.

```
out = summarizer(result[start:end]
```

After we summarized it ,the output is taken and appended to the object that we initiated at the start

```
out = out[0]
out = out['summary_text']
summarized_text.append(out)
```

Lastly, print summarized_text

```
print(summarized_text)
```

Summarize the summary:

This is nothing but getting the key points from the text .

This is done in a process ……

1. We will be having raw text .This raw text is cleaned (text pre-processing).Next it would be dividing the whole text into sentences so that the important sentences will be identified .
2. After that it tokenize the sentences(into words).Then converts these into vectors(numeric values) and then append the word vectors to get a long sentence vector.
3. After that creating a similarity matrix(having n rows and n columns where n is no of sentences) and then create a graph out of similarity matrix (for text summarization)we need to extract the sentences.So,nodes will be sentences ,edges

are similarity.Next rank the nodes(using page ranking algorithm).Next output the top key sentences.

4.In this way the summary will be summarized.

For this process it is needed to install summa

```
[ ] !pip install summa
```

```
Collecting summa
  Downloading https://files.pythonhosted.org/packages/45/3b/1c7dc435d05aef474c4137328400f1e11787b9bffab1f87a3f160c1fef54/summa-1.2.0.tar.gz (54kB)
     |████████████████████████████████| 61kB 3.8MB/s
Requirement already satisfied: scipy>=0.19 in /usr/local/lib/python3.7/dist-packages (from summa) (1.4.1)
Requirement already satisfied: numpy>=1.13.3 in /usr/local/lib/python3.7/dist-packages (from scipy>=0.19->summa) (1.19.5)
Building wheels for collected packages: summa
  Building wheel for summa (setup.py) ... done
  Created wheel for summa: filename=summa-1.2.0-cp37-none-any.whl size=54411 sha256=ec584d41ead87c729a8534084266907ad00436a60254db1766d8ca52ef54869d
  Stored in directory: /root/.cache/pip/wheels/6a/09/68/e2f2861c01d86407c3fa5220826ed7eed2abaa56b001be5970
Successfully built summa
Installing collected packages: summa
Successfully installed summa-1.2.0
```

5. Next we will be converting the text1 which is list into string and store in text.And the we will be given ratio so that according to that ratio the summary would get summarized

```
[ ] from summa.summarizer import summarize

[ ] text = str(text1)

[ ] print(text)

[ ] summarize(text, ratio = 0.5)
```

Convert the summary into other major languages:

As, there are different languages , all the languages are not known to everyone .One would be familiar with one language like English, telugu, hindi etc … So , converting the summary into other languages is important so that every one would be able to understand.

We can convert one/more statements or passages from one language to another language.We have to specify which language we have to covert.

For this , we have to install google trans module for doing this translation.

```
[ ] from googletrans import Translator
    translator = Translator()
    out = translator.translate(text, dest='te')
    print(out.text)
```

```
!pip install googletrans==3.1.0a0
```

```
Collecting googletrans==3.1.0a0
  Downloading https://files.pythonhosted.org/packages/19/3d/4e3a1609bf52f2f7b00436cc751eb977e27040665dde2bd57e7152989672/googletrans-3.1.0a0.tar.gz
Collecting httpx==0.13.3
  Downloading https://files.pythonhosted.org/packages/54/b4/698b284c6aed4d7c2b4fe3ba5df1fcf6093612423797e76fbb24890dd22f/httpx-0.13.3-py3-none-any.whl (55kB)
     |████████████████████████████████| 61kB 3.5MB/s
Requirement already satisfied: chardet==3.* in /usr/local/lib/python3.7/dist-packages (from httpx==0.13.3->googletrans==3.1.0a0) (3.0.4)
Collecting hstspreload
  Downloading https://files.pythonhosted.org/packages/dd/50/606213e12fb49c5eb667df0936223dcaf461f94e215ea60244b2b1e9b039/hstspreload-2020.12.22-py3-none-any.whl (994kB)
     |████████████████████████████████| 1.0MB 10.2MB/s
Requirement already satisfied: certifi in /usr/local/lib/python3.7/dist-packages (from httpx==0.13.3->googletrans==3.1.0a0) (2021.5.30)
Collecting rfc3986<2,>=1.3
  Downloading https://files.pythonhosted.org/packages/c4/e5/63ca2c4edf4e00657584608bee1001302bbf8c5f569340b78304f2f446cb/rfc3986-1.5.0-py2.py3-none-any.whl
Collecting sniffio
  Downloading https://files.pythonhosted.org/packages/52/b0/7b2e028b63d092804b6794595871f936aafa5e9322dcaaad50ebf67445b3/sniffio-1.2.0-py3-none-any.whl
Requirement already satisfied: idna==2.* in /usr/local/lib/python3.7/dist-packages (from httpx==0.13.3->googletrans==3.1.0a0) (2.10)
Collecting httpcore==0.9.*
  Downloading https://files.pythonhosted.org/packages/dd/d5/e4ff9318693ac6101a2095e580908b591838c6f33df8d3ee8dd953ba96a8/httpcore-0.9.1-py3-none-any.whl (42kB)
     |████████████████████████████████| 51kB 6.8MB/s
Collecting h2==3.*
  Downloading https://files.pythonhosted.org/packages/25/de/da019bcc539eeab02f6d45836f23858ac467f584bfec7a526ef200242afe/h2-3.2.0-py2.py3-none-any.whl (65kB)
     |████████████████████████████████| 71kB 8.4MB/s
Collecting h11<0.10,>=0.8
  Downloading https://files.pythonhosted.org/packages/5a/fd/3dad730b0f95e78aeeb742f96fa7bbecbdd56a58e405d3da440d5bfb90c6/h11-0.9.0-py2.py3-none-any.whl (53kB)
     |████████████████████████████████| 61kB 7.2MB/s
Collecting hyperframe<6,>=5.2.0
  Downloading https://files.pythonhosted.org/packages/19/0c/bf88182bcb5dce3094e2f3e4fe20db28a9928cb7bd5b08024030e4b140db/hyperframe-5.2.0-py2.py3-none-any.whl
Collecting hpack<4,>=3.0
  Downloading https://files.pythonhosted.org/packages/8a/cc/e53517f4a1e13f74776ca93271caef378dadec14d71c61c949d759d3db69/hpack-3.0.0-py2.py3-none-any.whl
Building wheels for collected packages: googletrans
  Building wheel for googletrans (setup.py) ... done
  Created wheel for googletrans: filename=googletrans-3.1.0a0-cp37-none-any.whl size=16368 sha256=68db45e25d064ff2aa75bb0fc9d0471ec815f36a61483e3e370bb9cc1d8552e7
  Stored in directory: /root/.cache/pip/wheels/27/7a/a0/aff3babbb775549ce6813cb8fa7ff3c0848c4dc62c20f8fdac
Successfully built googletrans
Installing collected packages: hstspreload, rfc3986, sniffio, hyperframe, hpack, h2, h11, httpcore, httpx, googletrans
Successfully installed googletrans-3.1.0a0 h11-0.9.0 h2-3.2.0 hpack-3.0.0 hstspreload-2020.12.22 httpcore-0.9.1 httpx-0.13.3 hyperframe-5.2.0 rfc3986-1.5.0 sniffio-1.2.0
```

Screens

1..Youtube Video

2. Summarized text

```
[' I want to thank every American who participated in this election whether '
 'you voted for the very first time or weighted it in wine for a very long '
 'time . We are an American family and we rise or fall together as one nation '
 'and as one people tonight . We know in our hearts but for the United States '
 'of America the best is yet to come .',
 ' President Obama congratulates Mitt Romney and Paul Ryan on a hard-fought '
 'campaign . Michelle Obama says she has never been prouder to Michelle Obama '
 '. Biden says he has never loved Michelle Obama more than he has loved her . '
 'Biden: "I wouldn\'t be the man I am today without the woman who agreed to '
 'marry me 20 years ago"',
 ' The president says he is proud of his campaign team and volunteers in the '
 'history of politics the best the best ever . He says he will always be '
 "grateful for everything that you've done and all the incredible work that "
 "you put in . 'I know that political campaigns can sometimes seem small even "
 "silly and that provides plenty of fodder for the cynics who say they're just "
 "silly'",
 ' L. ll us that politics nothing more than a contest of eagles or the domain '
 'of special interests but if you ever get the chance to talk to folks who '
 "turned out at our rally you'll hear the determination in the voice of a "
 "young field organizer who's working his way through college and wants to "
 'make sure every child has that same opportunity . The deep patriotism and '
 "the deep patriotism of a military spouse who's going door-to-door because "
 'her brother was finally hired when the local auto plant added another shift '
 '.',
 ' Democracy in a nation of 300 million can be noisy and messy and complicated '
 'we have our own opinions each of us has deeply held beliefs . Despite all '
 "our differences most of us share certain hopes for America's future . We "
 'want our kids to grow up in a country where they have access to the best '
 'schools and the best teachers .',
 " We want to pass on a country that's safe and respected and admired around "
 'the world a nation that is defended by the strongest military on earth and '
 'the best troops this this world has ever known . We believe in a generous '
 'America in a compassionate America in compassionate America  in a tolerant '
 'America open to the dreams of an immigrants daughter who studies in our '
 'schools .',
 ' President Barack Obama returns to the White House more determined and more '
 'inspired than ever about the work there is to do and the future of the '
 'blazing home tonight . Obama: "I have listened to you I have learned from '
```

```
"you and you've made me a better president and with your stories and your "
'struggles I return to the . White House .',
" I'm looking forward to reaching out and working with leaders of both "
'parties to meet the challenges we can only solve together reducing our '
'deficit reforming our tax code reforming our immigration system and freeing '
'ourselves from foreign oil . The role of citizen in our democracy does not '
"end with your vote but that doesn't mean your work is done . America has "
"never been about what can be done for us it's about what we can do by us "
'together too .',
" I've seen it in the family business whose owners would rather cut their own "
"pay than lay out their neighbors . I've also seen it on the shores of New "
'Jersey in New York where leaders from every party and level of government '
'have swept aside their differences to help a community rebuild from the '
'wreckage of a terrible storm .',
' President Obama spoke to a crowd in Mentor Ohio where a father told the '
'story of his eight-year-old daughter whose long battle with leukemia nearly '
'cost their family everything . Obama said he has never been that more '
'hopeful about America and he asks you to sustain that hope . Obama: "I\'m '
'not talking about blind optimism the kind of hope that just ignores the '
'enormity of t',
" I believe we can build on the progress we've made and continue to fight for "
'new jobs and new opportunity and new security for the middle class . I have '
'always believed that hope is that stubborn thing inside us that insists '
'despite all the evidence to the contrary that something better awaits us .',
' The U.S. is not as cynical as the pundits believe we are greater than the '
'sum of our individual ambitions . We are and forever will be the United '
'States of America and remind the world just why it is that we live in the '
'greatest nation on earth Thank You America you .']
```

3. `summarize(text, ratio = 0.5)`

'We know in our hearts but for the United States of America the best is yet to come .\', \' President Obama congratulates Mitt Romney and Paul Ryan on a hard-fought campaign .\nMichelle Obama says she has never been prouder to Michelle Obama .\nBiden says he has never loved Michelle Obama more than he has loved her .\nBiden: "I wouldn\\\'t be the man I am today without the woman who agreed to marry me 20 years ago"\', " The president says he is proud of his campaign team and volunteers in the history of politics the best the best ever .\nHe says he will always be grateful for everything that you\'ve done and all the incredible work that you put in .\nWe believe in a generous America in a compassionate America in compassionate America in a tolerant America open to the dreams of an immigrants daughter who studies in our schools .", \' President Barack Obama returns to the White House more determined and more inspired than ever about the work there is to do and the future of the blazing home tonight .\nObama: "I have listened to you I have learned from you and you\\\'ve made me a better president and with your stories and your struggles I return to the .\nI\'ve also seen it on the shores of New Jersey in New York where leaders from every party and level of government have swept aside their differences to help a community rebuild from the wreckage of a terrible storm .", \' President Obama spoke to a crowd in Mentor Ohio where a father told the story of his eight-year-old daughter whose long battle with leukemia nearly cost their family everything .\nObama said he has never been that more hopeful about America and he asks you to sustain that hope .\nObama: "I\\\'m not talking about blind optimism the kind of hope that just ignores the enormity of t\', " I believe we can build on the progress we\'ve made and continue to fight for new jobs and new opportunity and new security for the middle class .\nWe are and forever will be the United States of America and remind the world just why it is that we live in the greatest nation on earth Thank You America you .\']<'

4. Translate to telugu

మన హృదయాలలో మనకు తెలుసు, కాని యునైటెడ్ స్టేట్స్ ఆఫ్ అమెరికాకు ఇంకా ఉత్తమమైనది రాదు. \ ', President'
అధ్యక్షుడు ఒబామా మిట్ రోమ్నీ మరియు పాల్ ర్యాన్‌లను గట్టిగా పోరాడినందుకు అభినందించారు. \ N
మిచెల్ ఒబామా మాట్లాడుతూ మిచెల్‌కు అమె ఎప్పుడూ కలవరపడలేదు ఒబామా.
\ N మిచెల్ ఒబామాను తాను ప్రేమించిన దానికంటే ఎక్కువగా ప్రేమించలేదని బిడెన్ చెప్పారు.
Id n బిడెన్: "20 సంవత్సరాల క్రితం నన్ను వివాహం చేసుకోవడానికి అంగీకరించిన మహిళ లేకుండా నేను ఈ రోజు ఉన్న వ్యక్తిని కాను" \
', "రాజకీయ చరిత్రలో తన ప్రచార బృందం మరియు స్వచ్ఛంద సేవకుల గురించి తాను గర్విస్తున్నానని అధ్యక్షుడు చెప్పారు. \ N
మీరు చేసిన ప్రతిదానికీ మరియు మీరు పెట్టిన అద్భుతమైన పనికి ఆయన ఎల్లప్పుడూ కృతజ్ఞతలు తెలుపుతారని ఆయన అన్నారు.
మా పాఠశాలల్లో చదువుకునే పేద కూతురు కలల కోసం తెరిచిన సహనంతో ఉన్న అమెరికాలో కారుణ్య అమెరికాలో దయగల
అమెరికాలో ఉదారమైన అమెరికాను మేము నమ్ముతున్నాము. ",
'అధ్యక్షుడు బరాక్ ఒబామా వైట్ హౌస్కు తిరిగి వస్తారు. ఎప్పుడైనా చేయవలసిన పని గురించి మరియు మండుతున్న భవిష్యత్తు గురించి h ome ఒబామా:
"నేను మీ మాటలు విన్నాను, నేను మీ నుండి నేర్చుకున్నాను మరియు మీరు
నన్ను మంచి అధ్యక్షునిగా చేసారు మరియు మీ కథలు మరియు మీ పోరాటాలతో నేను తిరిగి వస్తాను. \ n
నేను కూడా దీనిని చూశాను న్యూయార్క్‌లోని న్యూజెర్సీ తీరం, ప్రతి పార్టీ మరియు ప్రభుత్వ స్థాయి నాయకులు తమ విభేదాలను పక్కనబెట్టి,
ఒక సమాజం భయంకరమైన తుఫాను శిధిలాల నుండి పునర్నిర్మించడంలో సహాయపడుతుంది.
తండ్రి తన ఎనిమిదేళ్ల కుమార్తె కథను చెప్పాడు, లుకేమియాతో సుదీర్ఘ పోరాటం వారి కుటుంబానికి దాదాపు ఖర్చవుతుంది. \ n
ఒబామా మాట్లాడుతూ, తాను అమెరికా గురించి ఇంత ఆశాజనకంగా ఎప్పుడూ లేనని, ఆ ఆశను నిలబెట్టుకోమని అతను మిమ్మల్ని అడుగుతున్నాడు. \ n
ఒబామా: "నేను \ బ్లైండ్ ఆశావాదం గురించి నేను మాట్లాడటం లేదు, ఇది కేవలం of 'యొక్క అపారతను విస్మరిస్తుంది,
"మేము సాధించిన పురోగతిని మనం నిర్మించగలమని మరియు కొత్త ఉద్యోగాలు మరియు కొత్త అవకాశం మరియు కొత్త భద్రత
కోసం పోరాటం కొనసాగించగలమని నేను నమ్ముతున్నాను. మధ్యతరగతి కోసం. \ n
మేము ఎప్పటికీ మరియు ఎప్పటికీ యునైటెడ్ స్టేట్స్ ఆఫ్ అమెరికాగా ఉంటాము. ధన్యవాదాలు ...

TESTING

Testing

6.1 Introduction

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, subassemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the

Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

6.2 TESTING METHODOLOGIES

The following are the Testing Methodologies:

- Unit Testing.
- Integration Testing.
- User Acceptance Testing.
- Output Testing.
- Validation Testing.

Unit Testing

Unit testing focuses verification effort on the smallest unit of Software design that is the module. Unit testing exercises specific paths in a module's control structure to ensure complete coverage and maximum error detection. This test focuses on each module individually, ensuring that it functions properly as a unit. Hence, the naming is Unit Testing.

During this testing, each module is tested individually and the module interfaces are verified for the consistency with design specification. All important processing path are tested for the expected results. All error handling paths are also tested.

Integration Testing

Integration testing addresses the issues associated with the dual problems of verification and program construction. After the software has been integrated a set of high order tests are conducted. The main objective in this testing process is to take unit tested modules and builds a program structure that has been dictated by design.

The following are the types of Integration Testing:

1.Top Down Integration

This method is an incremental approach to the construction of program structure. Modules are integrated by moving downward through the control hierarchy, beginning with the main program module. The module subordinates to the main program module are incorporated into the structure in either a depth first or breadth first manner. In this method, the software is tested from main module and individual stubs are replaced when the test proceeds downwards.

2. Bottom-up Integration

This method begins the construction and testing with the modules at the lowest level in the program structure. Since the modules are integrated from the bottom up, processing required for modules subordinate to a given level is always available and the need for stubs is eliminated. The bottom up integration strategy may be implemented with the following steps:

- The low-level modules are combined into clusters into clusters that perform a specific Software sub-function.
- A driver (i.e.) the control program for testing is written to coordinate test case input and output.
- The cluster is tested.
- Drivers are removed and clusters are combined moving upward in the program structure

The bottom up approaches test each module individually and then each module is module is integrated with a main module and tested for functionality.

OTHER TESTING METHODOLOGIES

User Acceptance Testing

User Acceptance of a system is the key factor for the success of any system. The system under consideration is tested for user acceptance by constantly keeping in touch with the prospective system users at the time of developing and making changes wherever required. The system developed provides a friendly user interface that can easily be understood even by a person who is new to the system.

Output Testing

After performing the validation testing, the next step is output testing of the proposed system, since no system could be useful if it does not produce the required output in the specified format. Asking the users about the format required by them tests the outputs generated or displayed by the system under consideration. Hence the output format is considered in 2 ways – one is on screen and another in printed format.

Validation Checking

Validation checks are performed on the following fields.

Text Field:

The text field can contain only the number of characters lesser than or equal to its size. The text fields are alphanumeric in some tables and alphabetic in other tables. Incorrect entry always flashes and error message.

Numeric Field:

The numeric field can contain only numbers from 0 to 9. An entry of any character flashes an error messages.

The individual modules are checked for accuracy and what it has to perform. Each module is subjected to test run along with sample data. The individually tested modules are

integrated into a single system. Testing involves executing the real data information is used in the program the existence of any program defect is inferred from the output. The testing should be planned so that all the requirements are individually tested. A successful test is one that gives out the defects for the inappropriate data and produces and output revealing the errors in the system.

Preparation of Test Data

Taking various kinds of test data does the above testing. Preparation of test data plays a vital role in the system testing. After preparing the test data the system under study is tested using that test data. While testing the system by using test data errors are again uncovered and corrected by using above testing steps and corrections are also noted for future use.

Using Live Test Data:

Live test data are those that are actually extracted from organization files. After a system is partially constructed, programmers or analysts often ask users to key in a set of data from their normal activities. Then, the systems person uses this data as a way to partially test the system. In other instances, programmers or analysts extract a set of live data from the files and have them entered themselves.

It is difficult to obtain live data in sufficient amounts to conduct extensive testing. And, although it is realistic data that will show how the system will perform for the typical processing requirement, assuming that the live data entered are in fact typical, such data generally will not test all combinations or formats that can enter the system. This bias toward typical values then does not provide a true systems test and in fact ignores the cases most likely to cause system failure.

Using Artificial Test Data:

Artificial test data are created solely for test purposes, since they can be generated to test all combinations of formats and values. In other words, the artificial data, which can quickly be prepared by a data generating utility program in the information systems department, make possible the testing of all login and control paths through the program. The most effective test programs use artificial test data generated by persons other than those who wrote the programs. Often, an independent team of testers formulates a testing plan, using the systems specifications. The package "Virtual Private Network" has satisfied all the requirements specified as per software requirement specification and was accepted.

USER TRAINING

Whenever a new system is developed, user training is required to educate them about the working of the system so that it can be put to efficient use by those for whom the system has been primarily designed. For this purpose the normal working of the project was demonstrated to the prospective users. Its working is easily understandable and since the expected users are people who have good knowledge of computers, the use of this system is very easy.

MAINTAINENCE

This covers a wide range of activities including correcting code and design errors. To reduce the need for maintenance in the long run, we have more accurately defined the user's requirements during the process of system development. Depending on the requirements, this system has been developed to satisfy the needs to the largest possible extent. With development in technology, it may be possible to add many more features based on the requirements in future. The coding and designing is simple and easy to understand which will make maintenance easier.

6.3 TESTING STRATEGY :

A strategy for system testing integrates system test cases and design techniques into a well planned series of steps that results in the successful construction of software. The testing strategy must co-operate test planning, test case design, test execution, and the resultant data collection and evaluation .A strategy for software testing must accommodate low-level tests that are necessary to verify that a small source code segment has been correctly implemented as well as high level tests that validate major system functions against user requirements.

Software testing is a critical element of software quality assurance and represents the ultimate review of specification design and coding. Testing represents an interesting anomaly for the software. Thus, a series of testing are performed for the proposed system before the system is ready for user acceptance testing.

SYSTEM TESTING:

Software once validated must be combined with other system elements (e.g. Hardware, people, database). System testing verifies that all the elements are proper and that overall system function performance is achieved. It also tests to find discrepancies between the system and its original objective, current specifications and system documentation.

UNIT TESTING:

In unit testing different are modules are tested against the specifications produced during the design for the modules. Unit testing is essential for verification of the code produced during the coding phase, and hence the goals to test the internal logic of the modules. Using the detailed design description as a guide, important Conrail paths are tested to uncover errors within the boundary of the modules. This testing is carried out during the programming stage itself. In this type of testing step, each module was found to be working satisfactorily as regards to the expected output from the module.

In Due Course, latest technology advancements will be taken into consideration. As part of technical build-up many components of the networking system will be generic in nature so that future projects can either use or interact with this. The future holds a lot to offer to the development and refinement of this project.

CONCLUSION

CONCLUSION

This project is mainly focused on the shortening of the valuable time of the people by summarising a huge amount of information that people tend to get on daily bases via youtube. In our project along with the text summarization we performed text translation that is the generated text can be translated into different languages

.The main objective of translation is to transfer the intent of a message and original tone, taking into consideration regional and cultural differences between target and source languages.

Nowadays, translation is a field that is seeing a lot of activity these days, it’s all because of the increasing globalization of businesses. There is a large number of people in a country that do not speak their local language well.

so even if any person encountering the problem would rely on this tool for the information needed in one's daily life

BIBILOGRAPHY

Bibilography

1. https://youtu.be/XO97Uon83Os
2. https://youtu.be/aOL7wzEIZSc
3. https://pypi.org/project/youtube-transcript-api/
4. https://www.kdnuggets.com/2021/02/hugging-face-transformer-basics.html

APPENDIX

APPENDIX

9.1 Introduction to python

Python is an open source , high-level programming language developed by Guido van Rossum in the late 1980s and presently administered by Python Software Foundation. It came from the ABC language that he helped create early on in his career. Python is a powerful language that you can use to create games, write GUIs and develop web applications.

It is a high-level language. Reading and writing codes in Python is much like reading and writing regular English statements. Because they are not written in machinereadable language, Python programs need to be processed before machines can run them. Python is an interpreted language. This means that every time a program is run, its interpreter runs through the code and translates it into machine readable byte code.

Python is an object-oriented language that allows users to manage and control data structures or objects to create and run programs. Everything in Python is latest version of Python in fact, first class. All objects, data types, functions, methods, and classes take equal position in Python. Programming languages are created to satisfy the needs of programmers and users for an effective tool to develop applications that impact lives, lifestyles, economy, and society. They help make lives better by increasing productivity, enhancing communication, and improving efficiency. Languages die and become obsolete when they fail to live up to expectations and are replaced and superseded by languages that are more powerful.

Python is a programming language that has stood the test of time and has remained relevant across industries and businesses and among programmers, and individual users. It is a living, thriving, and highly useful language that is highly recommended as a first programming language for those who want to dive
In to and experience programming. Advantages of Using Python Here are reasons why you would prefer to learn and use Python over other high-level languages

Readability

Python programs use clear, simple, and concise instructions that are easy to read even by those who have no substantial programming background. Programs written in Python are, therefore, easier to maintain, debug, or enhance.

Higher productivity

Codes used in Python are considerably shorter, simpler, and less verbose than other high level programming languages such as Java and C++. In addition, it has well-designed built-in features and standard library as well as access to third party modules and source libraries. These features make programming in Python more efficient.

Less learning time

Python is relatively easy to learn. Many find Python a good first language for learning programming because it uses simple syntax and shorter codes. Python works on Windows, Linux/UNIX, Mac OS X, other operating systems and small form devices. It also runs on microcontrollers used in appliances, toys, remote controls, embedded devices, and other similar devices.

Installing Python in Windows

To install Python, you must first download the installation package of your preferred version from this link: https://www.python.org/downloads/ On this page, you will be asked to choose between the two latest versions for Python 2 and 3: Python 3.5.1 and Python 2.7.11. Alternatively, if you are looking for a specific release, you can scroll down the page to find download links for earlier versions. You would normally opt to download the latest version, which is Python 3.5.1. This was released on December 7, 2015. However, you may opt for the latest version of Python 2, 2.7.11. Your preferences will usually depend on which version will be most usable for your project. While Python 3 is the present and future of the language, issues such as third-party utility or compatibility may require you to download Python 2.

Colab Notebook:

Colab notebooks allows you to combine executable code and rich text in a single document, along with images, HTML, LaTeX and more. When you create your own Colab notebooks, they are stored in your Google Drive account. You can easily share your Colab notebooks with co-workers or friends, allowing them to comment on your notebooks or even edit them.

Getting Started:

There are two ways that python is commonly used. The first is an interactive command environment, each as Python or IDLE, which are commonly bundled with the Python interpreter. Starting Python with one of these (using Start/Python in Windows, or by typing python at a command prompt (which will be shown as >>>). Unlike with C or Java, you can type commands at this prompt and the interpreter will run the commands and display the results, if any, on the screen. You can write functions in a text editor and run them from the command prompt by calling them by name.