A MINI PROJECT REPORT

On

CLOUD BASED ONLINE TEXT SUMMARIZATION SYSTEM

Submitted in partial fulfillment of the requirement of
University of Mumbai for the Course

Bachelor of Engineering
In
Computer Engineering

Submitted By
P V Athira
Anoop Pillai
Nikhil Chandran
Mayuresh Pitale

Supervisor
Prof. Sheetal Gawande

Department Of Computer Engineering
PILLAI COLLEGE OF ENGINEERING
New Panvel – 410 206
UNIVERSITY OF MUMBAI

Academic Year 2019 – 20

Department of Computer Engineering
Pillai College of Engineering
New Panvel – 410 206

CERTIFICATE

This is to certify that the requirements for the project report entitled ' **Cloud Based Online Text Summarization System** ' have been successfully completed by the following students:

Name	**Roll No.**
P V Athira	B853
Anoop Pillai	B847
Nikhil Chandran	B850
Mayuresh Pitale	B851

in partial fulfillment of Bachelor of Engineering of Mumbai University in the Department of Computer Engineering,Pillai College of Engineering, New Panvel – 410 206 during the Academic Year 2019 – 20.

_____

Supervisor

Subject Incharge

Department of Computer Engineering
Pillai College of Engineering
New Panvel – 410 206

PROJECT APPROVAL

This project entitled "**Cloud Based Online Text Summarization System**" by **P V Athira**, **Anoop Pillai**, **Nikhil Chandran** and **Mayuresh Pitale** are approved for the course of Cloud Computing in Computer Engineering (VIII sem) of Mumbai University in the Department of **Computer Engineering**.

Examiners:

1._____

2._____

Supervisors:

1._____

2._____

Subject Incharge:

1._____

Date:

Place: New Panvel

DECLARATION

We declare that this written submission for the Cloud Computing mini project entitled "**Cloud Based Online Text Summarization System**" 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 ideas / data / fact / source in our submission. We understand that any violation of the above will cause disciplinary action by the institute and also evoke penal action from the sources which have not been properly cited or from whom prior permission has not been taken when needed

Project Group Members :

P V Athira : _____

Anoop Pillai : _____

Nikhil Chandran : _____

Mayuresh Pitale : _____

Date:

Place: New Panvel

Table of Contents

Abstract

A person uses the internet to access information available globally. While doing so, he uses many cloud services, with or without his knowledge, while using various websites. A person may be looking for precise and summarized data, rather than huge documents of information. Due to an exponential growth in the generation of textual data, the need for tools and mechanisms for automatic summarization of documents has become very critical. Text documents are vital to any organization's day-to-day working and as such, long documents often hamper trivial work. Most of the text resources such as books, news articles, blog posts, research papers, emails, and tweets present on the internet have repetitive content which is vestigial to the day-to-day user. Tremendous amounts of information is available on the internet, it is important to develop a mechanism to extract concise and meaningful summary of text from multiple text resources quickly and most efficiently from document files and websites. Therefore, an automatic cloud based text summarizer is vital towards reducing human effort. A cloud based summarizer helps to provide a user flexibility and portability due to the ability of being able to use it anywhere, anytime.

List of Figures

List of Tables

Chapter 1
Introduction

Text summarization simply means the technique of shortening the long pieces of text. The intention is to create a systematic and fluent summary having only the main points outlined in the document. With such a big amount of data circulating in the digital space, there is a need to develop machine learning algorithms that can automatically shorten longer texts and deliver accurate summaries that can fluently pass the intended messages in very less time. Text summarization methods are greatly needed to address the ever-growing amount of text data available online to both better help discover relevant information and to consume relevant information faster. Now in order for the application to have maximum reach and be easily available to everyone, it has to be online. Hence, in this project, we have created a cloud based summarization website that provides a summarized report of the input text and which is cloud hosted to attain the services of cloud.

Some of the services provided by cloud computing organizations are :

1.Scalable : A cloud service allows quick scaling up and down of computing resources to accommodate your changing needs.

2.Affordable : You pay less for a cloud service, as it eliminates unnecessary costs involved in hardware upgrades and maintenance.

3.Secure : By signing up for a cloud service, you are essentially making your data more secure using their industry-grade security protocols.Cloud computing offers great security when any sensitive data has been lost. As the data is stored in the system, it can be easily accessed even if something happens to your computer. You can even remotely wipe out data from the lost machines for avoiding it getting in the wrong hands.

4.Mobility : Cloud computing allows mobile access to corporate data via smartphones and devices, which is a great way to ensure that no one is ever left out of the loop. Resources in the cloud can be easily stored, retrieved, recovered, or processed with just a couple of clicks. Users can get access to their works on-the-go, 24/7, via any devices of their choice, in any corner of the world as long as you stay connected to the internet.

5.Speed : Most cloud computing services are provided self service and on demand, typically with just a few mouse clicks, giving businesses a lot of flexibility.

6.Reliability : Cloud computing makes data backup, disaster recovery and business continuity easier and less expensive because data can be mirrored at multiple redundant sites on the cloud provider's network.

1.1 Fundamentals :

Cloud computing simply is the delivery of computing services that include servers, storage, databases, networking, software, analytics, and intelligence over the Internet ("the cloud") to offer faster innovation, flexible resources, and economies of scale.Cloud computing acts as a virtual server and helps in processing for the systems that aren't capable of computing the data due to hardware limitations.Cloud hosting is the acquisition of computing resources from a cloud computing provider or facility to host data, services and/or solutions.

In cloud computing, cloud models can be distinguished into three types based on the business requirement.Each of these cloud models has their own set of benefits that could serve the needs of various businesses.

- IaaS (Infrastructure as a Service)
- PaaS (Platform as a Service).
- SaaS (Software as a Service)

IaaS : Infrastructure as a service (IaaS) is a cloud computing offering in which a vendor provides users access to computing resources such as servers, storage and networking. Organizations use

their own platforms and applications within a service provider's infrastructure.

Key features :

- Instead of purchasing hardware outright, users pay for IaaS on demand.
- Infrastructure is scalable depending on processing and storage needs.
- Saves enterprises the costs of buying and maintaining their own hardware.
- Because data is on the cloud, there can be no single point of failure.
- Enables the virtualization of administrative tasks, freeing up time for other work.

PaaS : Platform as a service (PaaS) is a cloud computing offering that provides users with a cloud environment in which they can develop, manage and deliver applications. In addition to storage and other computing resources, users are able to use a suite of prebuilt tools to develop, customize and test their own applications.

Key features :

- PaaS provides a platform with tools to test, develop and host applications in the same environment.
- Enables organizations to focus on development without having to worry about underlying infrastructure.
- Providers manage security, operating systems, server software and backups.
- Facilitates collaborative work even if teams work remotely.

SaaS : Software as a service (SaaS) is a cloud computing offering that provides users with access to a vendor's cloud-based software. Users do not install applications on their local devices. Instead, the applications reside on a remote cloud network accessed through the web or an API. Through the application, users can store and analyze data and collaborate on projects.

Key features

- SaaS vendors provide users with software and applications via a subscription model.

- Users do not have to manage, install or upgrade software; SaaS providers manage this.
- Data is secure in the cloud; equipment failure does not result in loss of data.
- Applications are accessible from almost any internet-connected device, from virtually anywhere in the world.

1.2 Objectives

The objective of this work is as follows:

1. Reduction of a given text to a smaller number of sentences
2. To summarize the document/text on the cloud without leaving out the main ideas of the original text.
3. To identify the most important information from the given text and present it to the end users.
4. Provide the summarized text quickly and efficiently.
5. Easy availability and mobility on the cloud

1.3 Scope :

One of the plans is to apply the topic-focused summarization framework to news articles or blogs and to extend the work in the machine learning approaches. Topic focused summaries of news articles would be a lot more accurate and valuable to users. It would be more interesting to work on topic modeling and summarization in the domain of social media in future. The rate at which the information is growing is tremendous. Hence it is very important to build a multilingual summarization system that would be accessible to anyone and works with a single click. The work presented by the thesis can also be developed into a dedicated cloud based application that a user can use on his device. The implemented system in this thesis can work as a framework for the research community to understand and extend the applicability of cognitive and symbolic approach in various domains of business needs. Research in summarization continues to enhance the diversity and information richness and strive to produce coherent and focused answers to users' information needs.

1.4 Organization of the Report

The report is organized as follows:

Chapter 1 The introduction is given in Chapter 1. It describes the fundamental terms used in this project. It motivates to study and understand the different techniques used in this work. This chapter also presents the outline of the objective of the report.

Chapter 2 It describes the review of the relevant various techniques in the literature systems. It describes the pros and cons of each technique.

Chapter 3 This chapter presents the theory and proposed work. It describes the major approaches used in this project work.

Chapter 4 The societal and technical applications are mentioned in Chapter 4.

Chapter 5 The summary of the report is presented in Chapter 5.

Chapter 2

Literature Survey

2.1 Introduction :

Literature surveys provide brief overviews or a summary of the current research on topics.Literature surveys are used in ensuring that the used experiments, methodologies and experiments offer reliability and validity in the research being conducted.They are useful in validating or providing proof and also provides a base of moving a research idea forward on what researchers have done and exciting avenues that it opens for investigation during future work in the field.

In this literature survey research papers and articles on automatic text summarization were referred through for the project.

2.2 Literature Survey :

The following research papers and reports were used for the topic of Text Summarization System for English Language :

1. Automatic Text Summarization Using Natural Language Processing :

Authors Pratibha Devihosur, Naseer R implemented an automatic text summarization mechanism based on an unsupervised learning system.The significance of the generated summary was assessed with the assistance of Simplified Lesk calculation along with an online semantic lexicon WordNet.Based on their evaluation the algorithm provides best summarized outcome ranging from 25-50 percent with respect to the source data. In this project they also focussed on ambiguous words because a specific word may have distinctive significance in various setting.Hence they tried to inculcate the principle of word sense disambiguation to decide the right feeling of a word utilized as part of a specific setting.

2. Text Summarization using Natural Language Processing :

In this paper authors Ankit Kumar, Zixin Luo and Ming Xu created an end to end web application which can take an article as input and generate a summary.The model was trained using deep learning approach and trained on Juniper's datasets.Juniper is a corporate organization that develops and markets networking devices.In order to provide a better customer experience, Juniper Networks maintains large datasets of articles wherein each of these articles can be long and verbose.Hence these datasets were used to train the text summarization model.The model built used abstractive summarization technique and significantly generated excellent human readable sentences from given inputs. However, it did not always generate summaries capturing all the important information in the input documents.

3. Text Summarization Techniques: A Brief Survey :

In this survey Mehdi Allahyari, Elizabeth D. Trippe, Saeid Safaei and others study the main approaches to automatic text summarization and also review the different processes for summarization and describe the effectiveness and shortcomings of different methods.Topics like the impact of context in summarization and semantic analysis are also mentioned.

4. NLP Based Text Summarization Using Semantic Analysis:

In this project paper authors Harsh Desai, Dhairya Pawar, Geet Agrawal reviewed the different methods for text summarization and provided a novel technique generating the summarization of domain specific text by using Semantic Analysis for text summarization.

5. Text Summarization:An Overview :

In this research paper author Samrat Babar provides an analysis about the meaning of text summarization in natural language processing and their types along with the technical and mathematical analysis of text summarization in detail. This paper basically is a documentation for all the information required to study as well refer through on the topic of automatic text summarization.

2.3 Literature Survey Summary :

The detailed summary of the above referred paper are mentioned in the table below **:**

SN	Techniques	Author & Year of Publication	Characteristics
1.	Automatic Text Summarization Using Natural Language Processing	Pratibha Devihosur , Naseer R.2017	1.Text summarization model based on unsupervised learning model 2.Includes word sense disambiguation to determine the sentiment of the word in that particular setting 3. Generated summary ranges from 25-50 percent with respect to the source data.

2.	Text Summarization using Natural Language Processing	Ankit Kumar,Zixin Luo and Ming Xu.2018	1.Creates an end to end text summarization web application using the abstractive summarization technique. 2.Training datasets used contain long and verbose articles which increases the accuracy of the trained model. 3.The trained model usually generated significantly excellent human readable sentences.
3.	Text Summarization Techniques: A Brief Survey	Mehdi Allahyari, Elizabeth D. Trippe, Saeid Safaei and others.2017	1.Provides detailed description about the main approaches on automatic text summarization. 2.Reviews the different processes for summarization and also describes the effectiveness and shortcomings of different methods 3.Also includes topics like the impact of context in summarization and semantic analysis.
4.	NLP Based Text Summarization Using Semantic Analysis	Harsh Desai, Dhairya Pawar, Geet Agrawal.2016	1.Mentions the different methods used in text summarization. 2. Creates a .novel technique generating the summarization of domain specific text by using Semantic Analysis.
5.	Text Summarization: An Overview	Samrat Babar.2013	1.Comparatively provides a very detailed report about the process of text summarization in natural language processing. 2. The paper includes details about the meaning of summarization , their types and also includes the technical and mathematical analysis of the different approaches used.

Table 2.3 Literature survey summary

Chapter 3

Proposed System

3.1 Overview :

With the dramatic growth of the internet in the modern times people are able to access a tremendous amount of information on any topic. This information can be overwhelming at times to go through since people require these information quickly and with ease. This expanding availability of documents has demanded an exhaustive research in the field of text summarization. Now a summary is a text that is produced from one or more texts, that conveys the important information in the original text but is shorter in format.So therefore automatic text summarization is the process of producing a concise and fluent summary while preserving the key information content and overall meaning of the text.

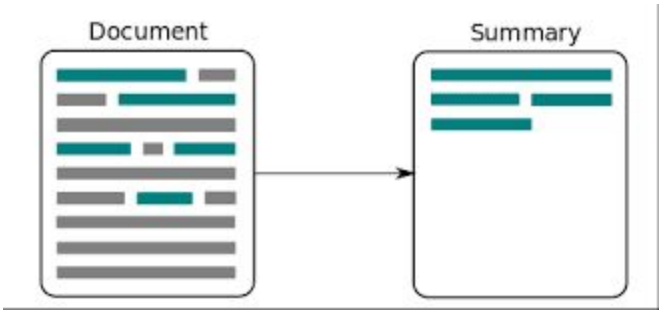

Fig : 3.1.1 Diagrammatic representation of text summarization

3.1.1 Existing System Architecture :

There are a few ways of going about classifying automated text summarization techniques such as extractive and abstractive text summarization.While abstractive text summarization methods employ more powerful natural language processing techniques to interpret text and generate new summary text , it is a far more difficult method to approach. The majority of summarization processes today are extraction-based since much work is needed to produce a strong foothold in the abstractive technique.

Whereas extractive text summarization methods function by identifying the important sentences or excerpts from the text and reproducing them almost precisely as part of the summary. No new text

is generated; only existing text is used in the summarization process. Now extractive text summarization techniques vary, yet they all share the same basic tasks:

1. Construct an intermediate representation of the input text (text to be summarized)

2. Score the sentences based on the constructed intermediate representation

3. Select a summary consisting of the top k most important sentence

Tasks 2 and 3 are straightforward enough; in sentence scoring, it is determined how well each sentence relays important aspects of the text being summarized, while sentence selection is performed using some specific optimization approach. Algorithms for each of these 2 steps can vary, but they are conceptually quite simple i.e assign a score to each sentence using some metric, and then select from the best-scored sentences via some well-defined sentence selection method.

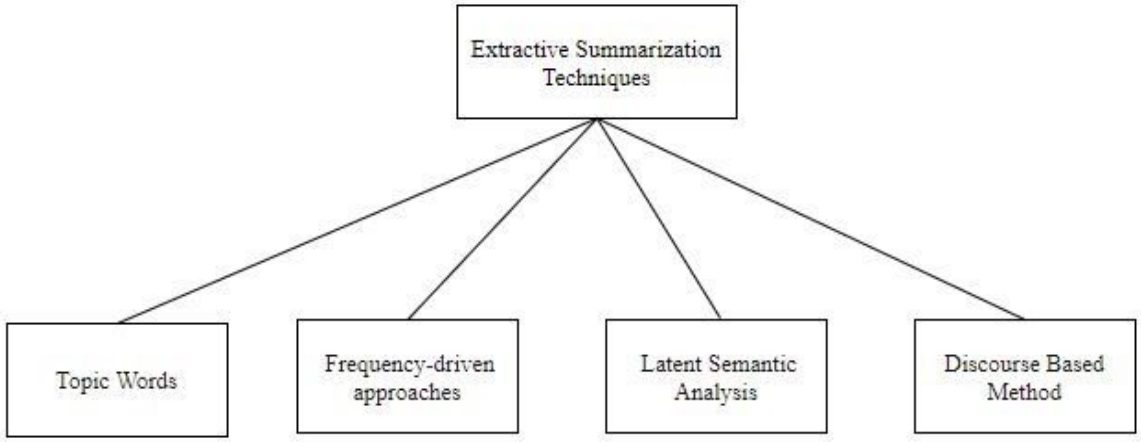

Fig : 3.1.1_1 Classification of extractive summarization techniques

Existing summarization systems :

(1) Topic Words : This common technique aims to identify words that describe the topic of the input document. Generally speaking, there are two ways to compute the importance of a sentence as a function of the number of topic signatures it contains, or as the proportion of the topic signatures in the sentence. While the first method gives higher scores to longer sentences with more words, the second one measures the density of the topic words.

(2) Frequency-driven approaches : This approach uses frequency of words as indicators of importance. The two most common techniques in this category are: word probability and TFIDF (Term Frequency Inverse Document Frequency). The probability of a word w is determined as the number of occurrences of the word, f (w), divided by the number of all words in the input. Words with highest probability are assumed to represent the topic of the document and are included in the summary. TFIDF, a more sophisticated technique, assesses the importance of words and identifies very common words (that should be omitted from consideration) in the document by giving low weights to words appearing in most documents. After creation of TFIDF vector representations of documents, the documents that describe the same topic are clustered together and centroids are computed — pseudo-documents that consist of the words whose TF IDF scores are higher than a certain threshold and form the cluster. Afterwards, the centroids are used to identify sentences in each cluster that are central to the topic.

(3) **Latent Semantic Analysis :** Latent semantic analysis (LSA) is an unsupervised method for extracting a representation of text semantics based on observed words. The first step is to build a term-sentence matrix, where each row corresponds to a word from the input (n words) and each column corresponds to a sentence. Each entry of the matrix is the weight of the word i in sentence j computed by TFIDF technique. Then singular value decomposition (SVD) is used on the matrix that transforms the initial matrix into three matrices: a term-topic matrix having weights of words, a diagonal matrix where each row corresponds to the weight of a topic, and a topic-sentence matrix. If you multiply the diagonal matrix with weights with the topic-sentence matrix, the result will describe how much a sentence represents a topic, in other words, the weight of the topic i in sentence j.

(4) Discourse Based Method : A logical development of analyzing semantics is performing discourse analysis, finding the semantic relations between textual units, to form a summary. The study on cross-document relations was initiated by Radev, who came up with the Cross-Document Structure Theory (CST) model. In his model, words, phrases or sentences can be linked with each other if they are semantically connected. CST was indeed useful for document summarization to determine sentence relevance as well as to treat repetition, complementarity and inconsistency among the diverse data sources. Nonetheless, the significant limitation of this method is that the CST relations should be explicitly determined by humans.

3.1.2 Proposed System Architecture :

In our proposed system, we will be hosting the Text Summarizer on a cloud, accessible through a website for ease of use to the user. Firstly, we have prepared the project online using Python programming language and various other machine learning libraries. After that, the required files were imported to 'pythonanywhere.com', a website which provides a hosting platform for cloud based applications. It is one of the simplest ways to deploy a web2py project online with cloud integration.

The cloud based automatic text summarizer architecture consist of two parts:

1. Implementing a successfully running automatic text summarization website
2. To host the website over cloud.

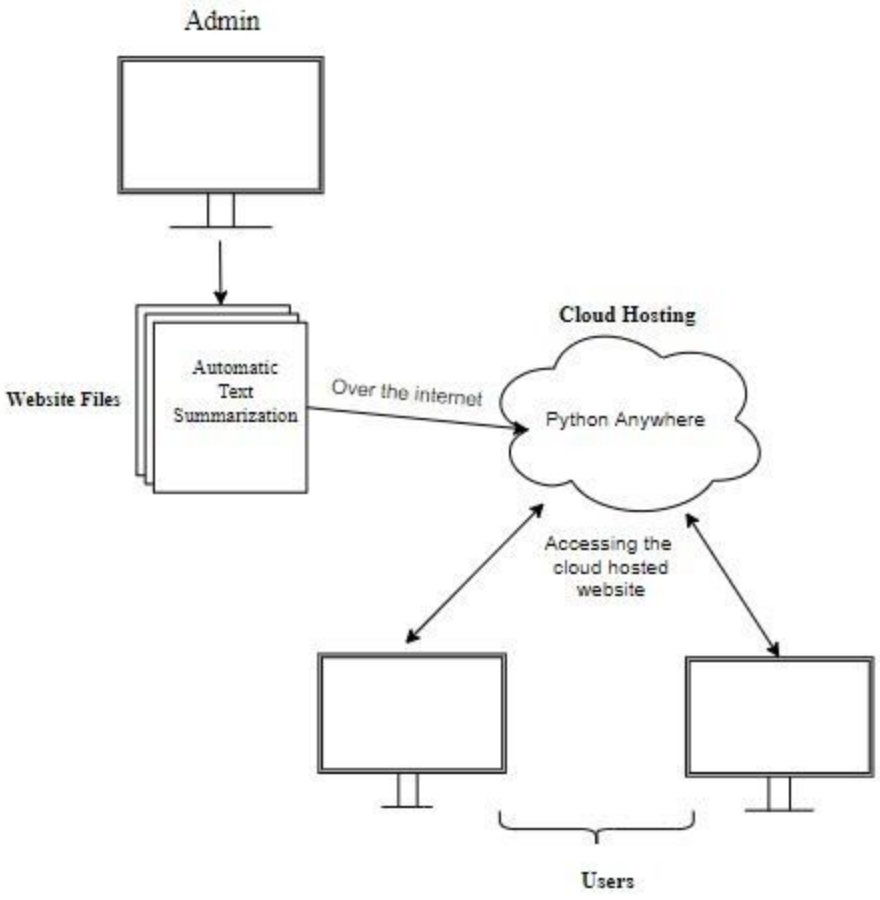

Fig : 3.1.2_1 Proposed System Architecture

3.2 Requirements for Implementation :

3.2.1 Tools

Html:- Hypertext Markup Language is the standard markup language for documents designed to be displayed in a web browser. It can be assisted by technologies such as Cascading Style Sheets and scripting languages such as JavaScript.

Python:- Python is a high-level programming language designed to be easy to read and simple to implement. Python is considered a scripting language, like Ruby or Perl and is often used for creating Web applications and dynamic Web content.

Flask:-

Flask is a micro web framework written in Python. It is classified as a microframework because it does not require particular tools or libraries. It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions.

PythonAnywhere:- PythonAnywhere is an online integrated development environment (IDE) and web hosting service (Platform as a service) based on the Python programming language. It provides in-browser access to server-based Python and Bash command-line interfaces, along with a code editor with syntax highlighting. Program files can be transferred to and from the service using the user's browser. Web applications hosted by the service can be written using any WSGI-based application framework.

3.2.2 Implementation steps :

Step 1: Go to pythonanywhere website https://www.pythonanywhere.com/

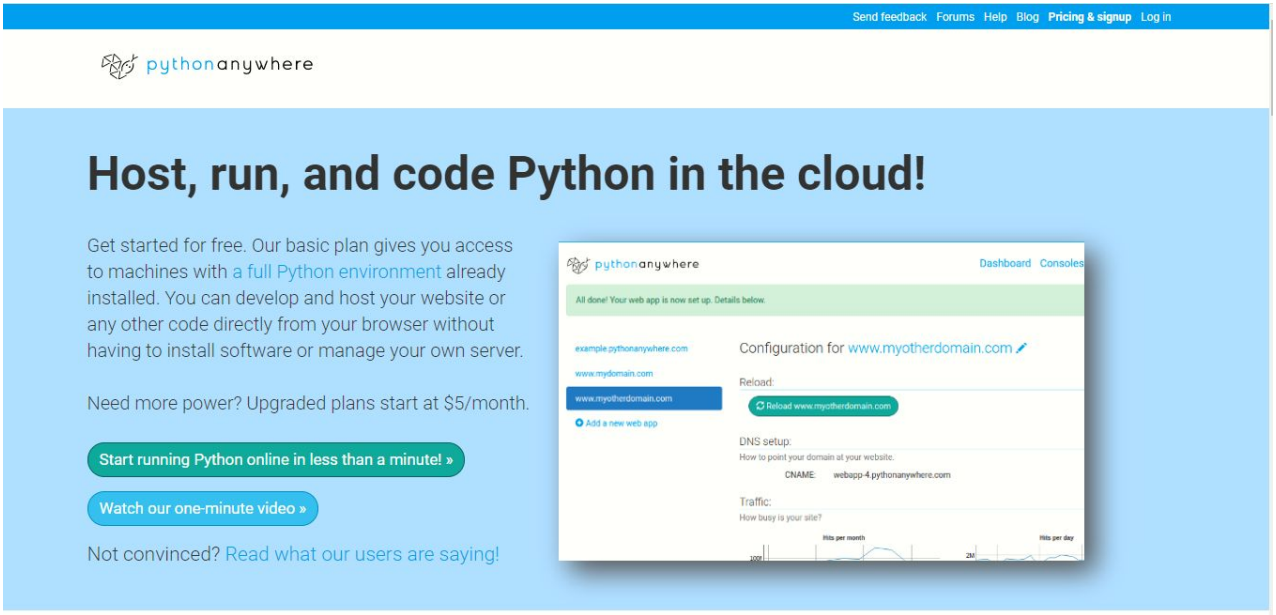

Fig.3.2.2_1 Pythonanywhere homepage

Step 2: Click on pricing & signup tab to create a new account.

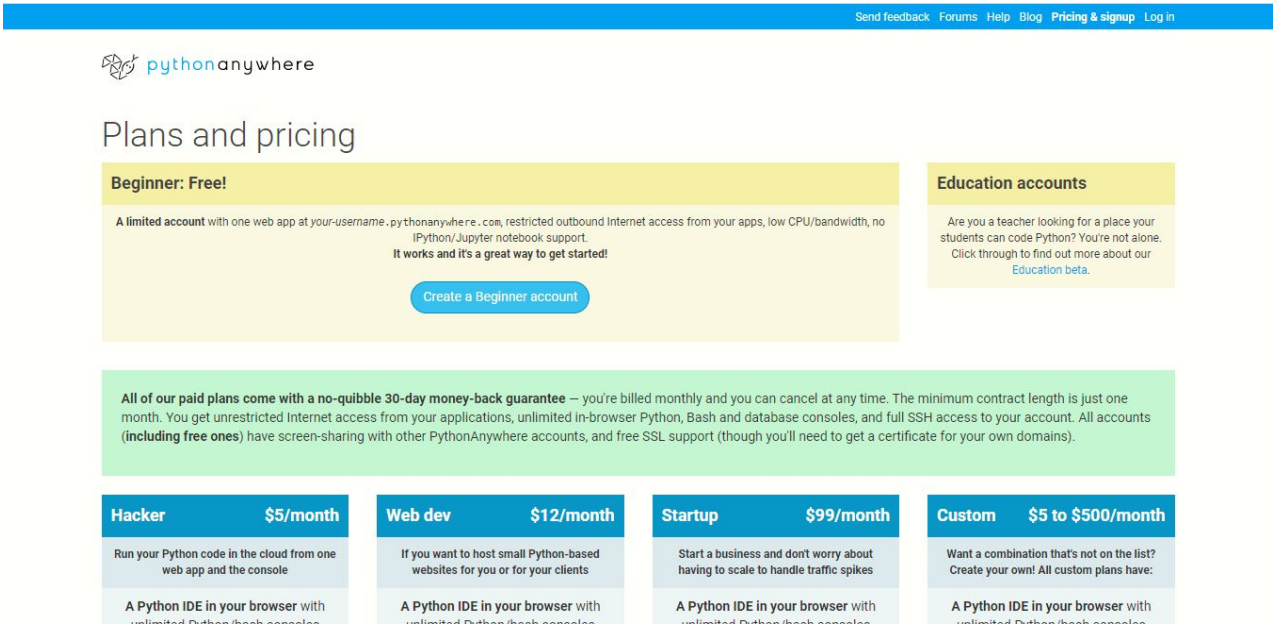

Fig.3.2.2_2 Pythonanywhere pricing & signup page

Step 3: Login to account and add upload files to be hosted

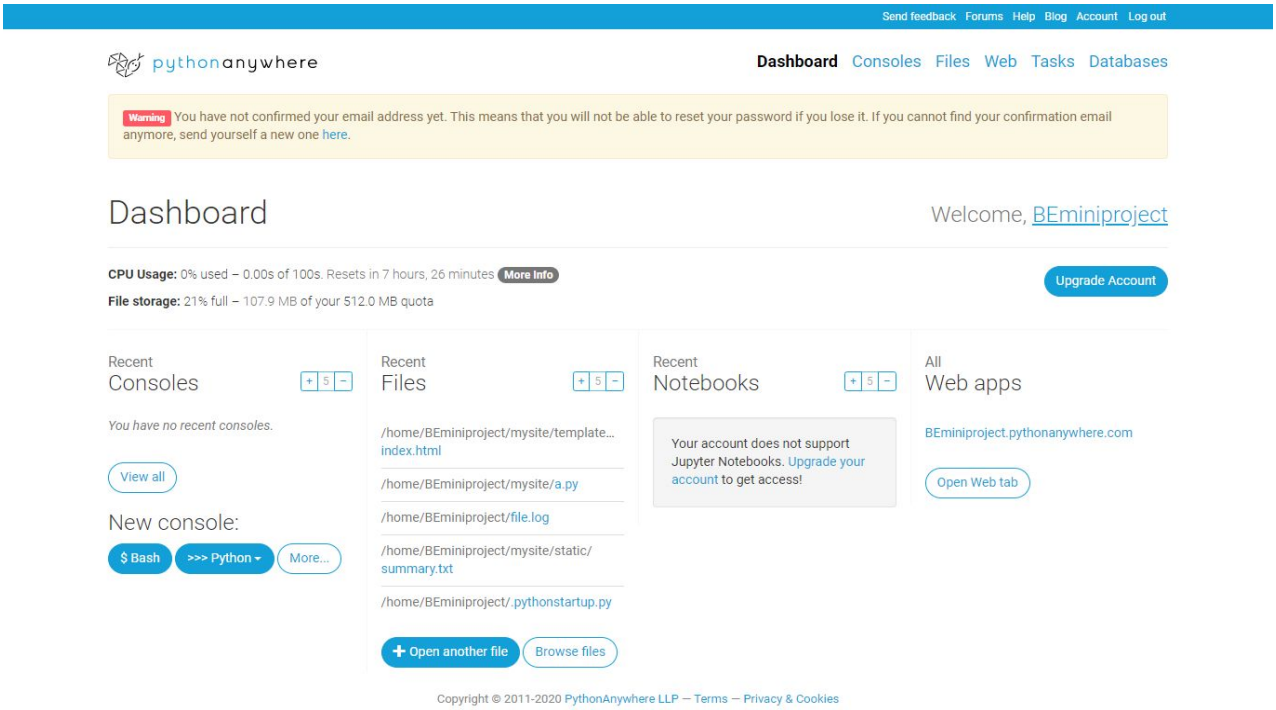

Fig.3.2.2_3 Dashboard

Step 4: Upload all code files in /home/BEminiproject/mysite/

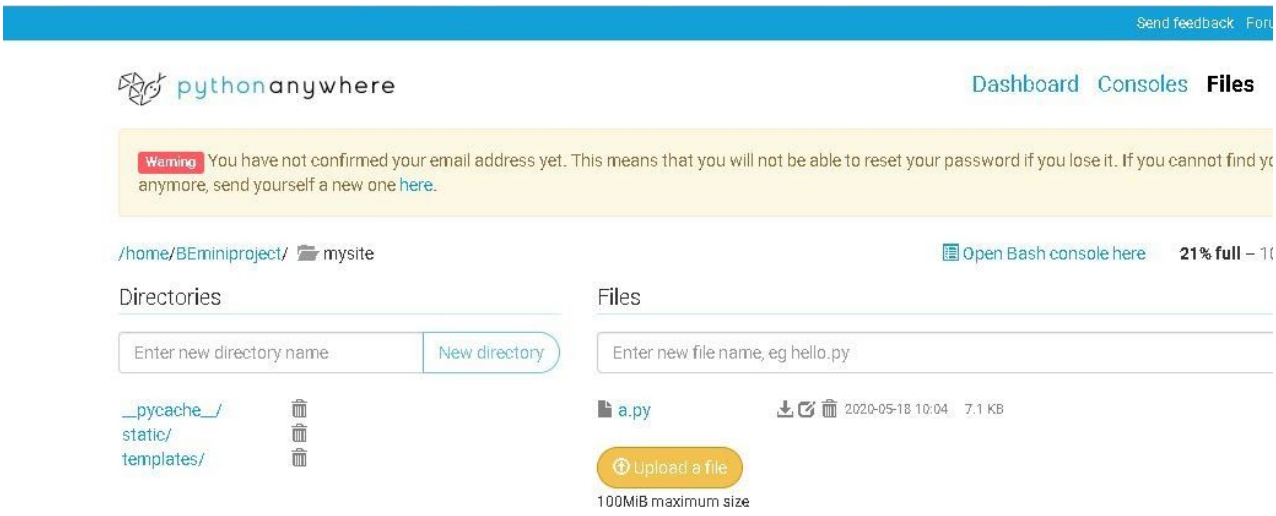

Fig.3.2.2_4 Folder ”mysite” which stores the html templates and python code of our project

Step 4: Setting up virtualenv

Open up a new Bash console from your Dashboard and run

```
mkvirtualenv --python=/usr/bin/python3.8 my-virtualenv
pip install flask
```

Step 5: Setting up the Web app using Manual configuration

Go to the Web Tab and hit Add a new web app. Choose Manual Configuration, and then choose the Python version -- make sure it's the same version as the one you used in your virtualenv

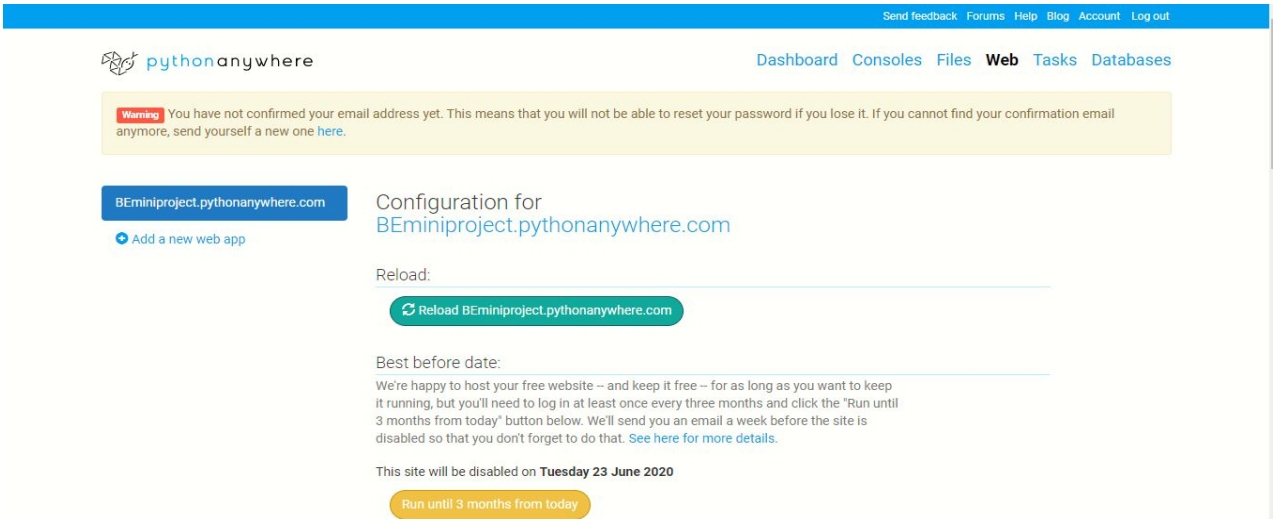

Fig. 3.2.2_5 Setting up theWeb app using manual configuration

Step 6: Configuring the WSGI file

/var/www/beminiproject_pythonanywhere_com_wsgi.py

```
1  # This file contains the WSGI configuration required to serve up your
2  # web application at http://<your-username>.pythonanywhere.com/
3  # It works by setting the variable 'application' to a WSGI handler of some
4  # description.
5  #
6  # The below has been auto-generated for your Flask project
7
8  import sys
9
10 # add your project directory to the sys.path
11 project_home = '/home/BEminiproject/mysite'
12 if project_home not in sys.path:
13     sys.path = [project_home] + sys.path
14
15 # import flask app but need to call it "application" for WSGI to work
16 from a import app as application  # noqa
17
```

Fig. 3.2.2_6 Configuring the WSGI file

Step 7: Our website is now successfully hosted on http://beminiproject.pythonanywhere.com/

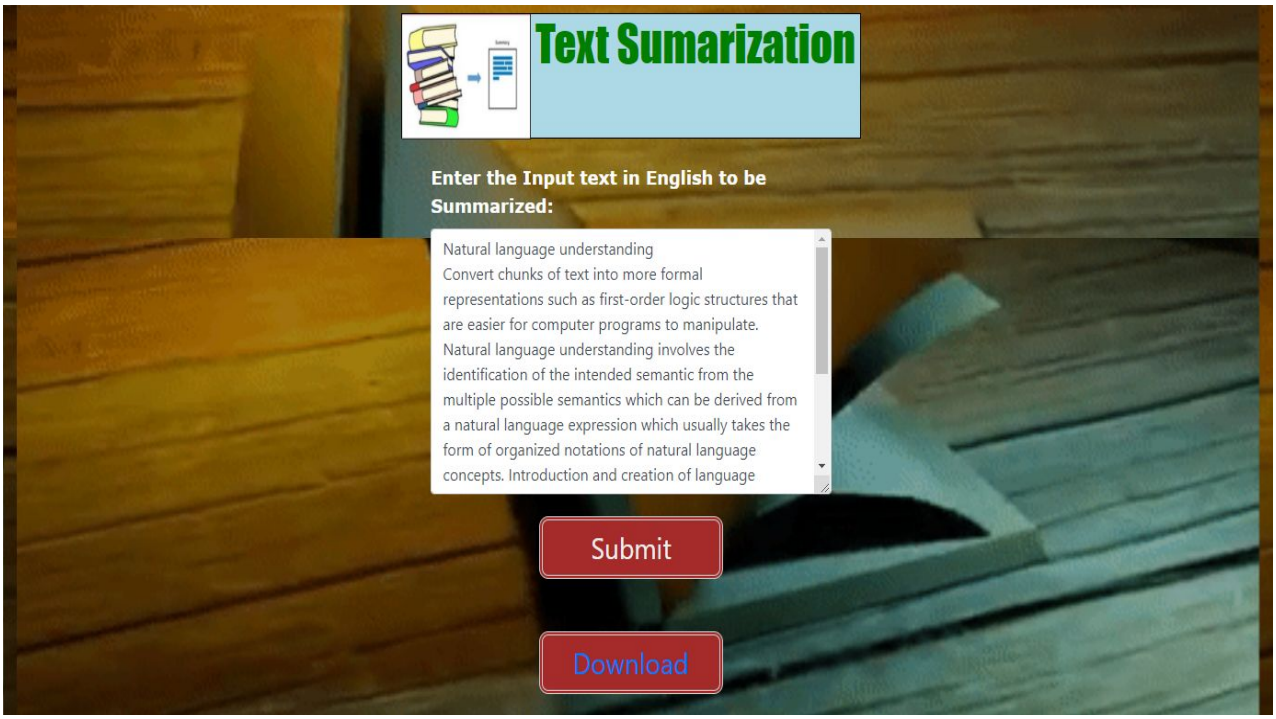

Fig. 3.2.2_7 Home page

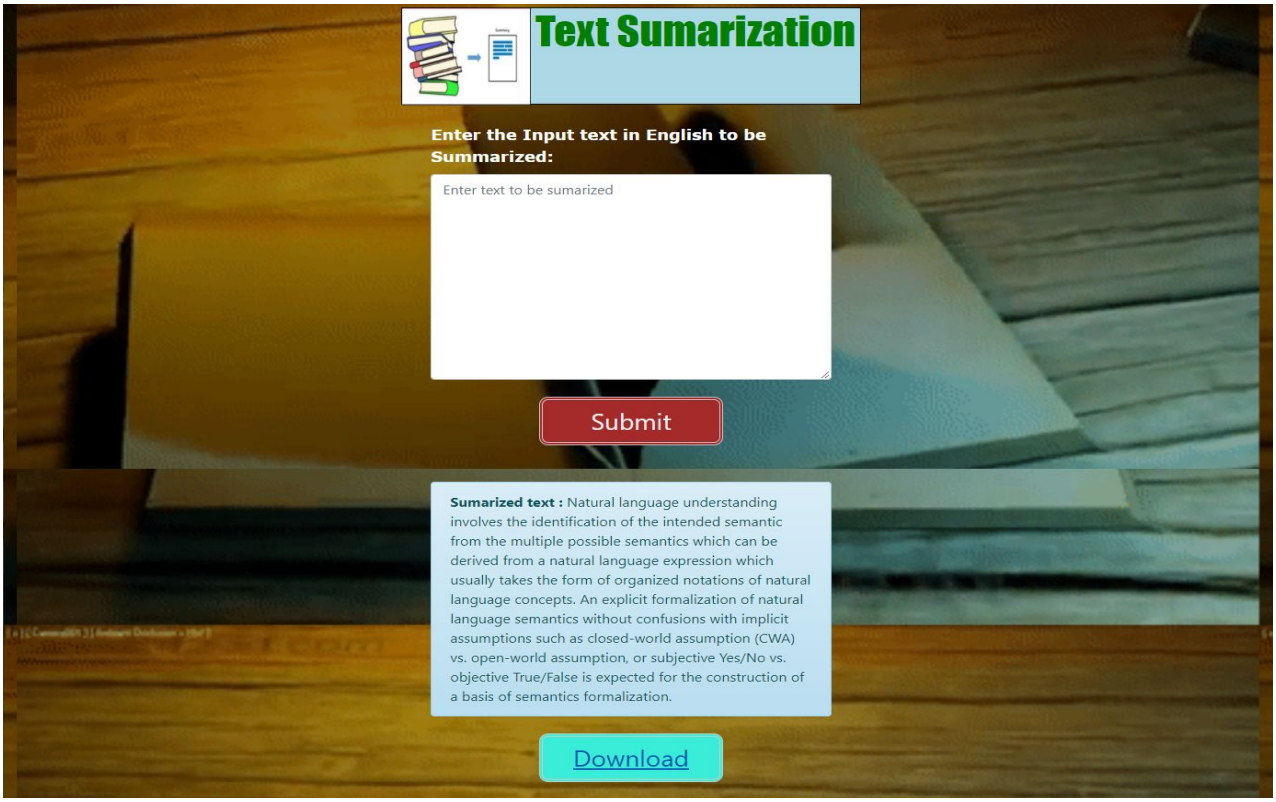

Fig. 3.2.2_8 : Adding in the text to be summarized

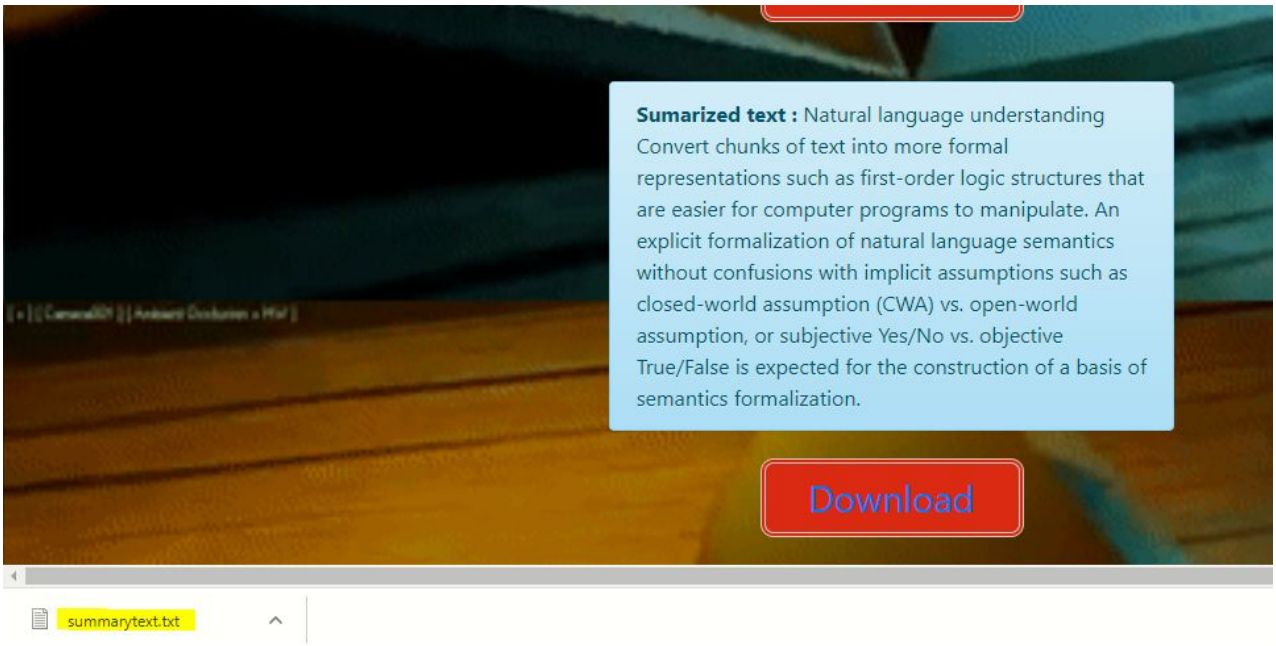

Fig.3.2.2_9 Downloading the summarized text in a file

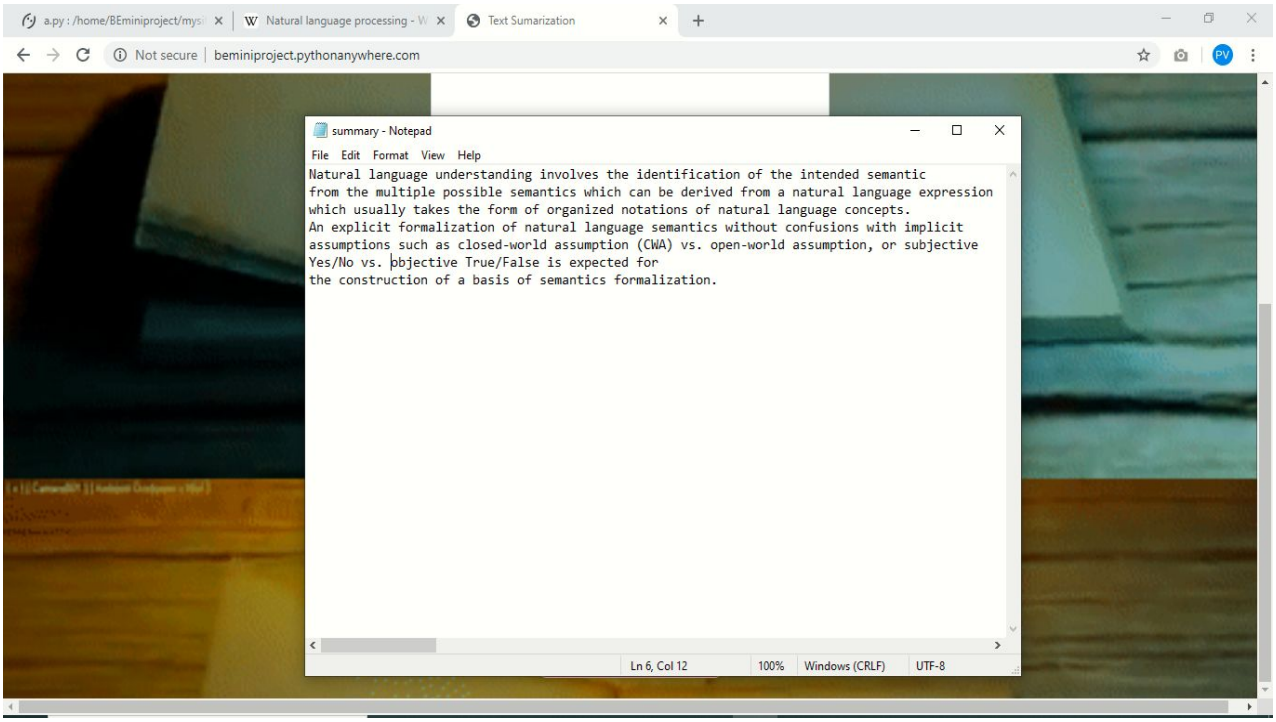

Fig.3.2.2_10 Viewing the downloaded summarized text file

3.2.3 Use Case Diagram

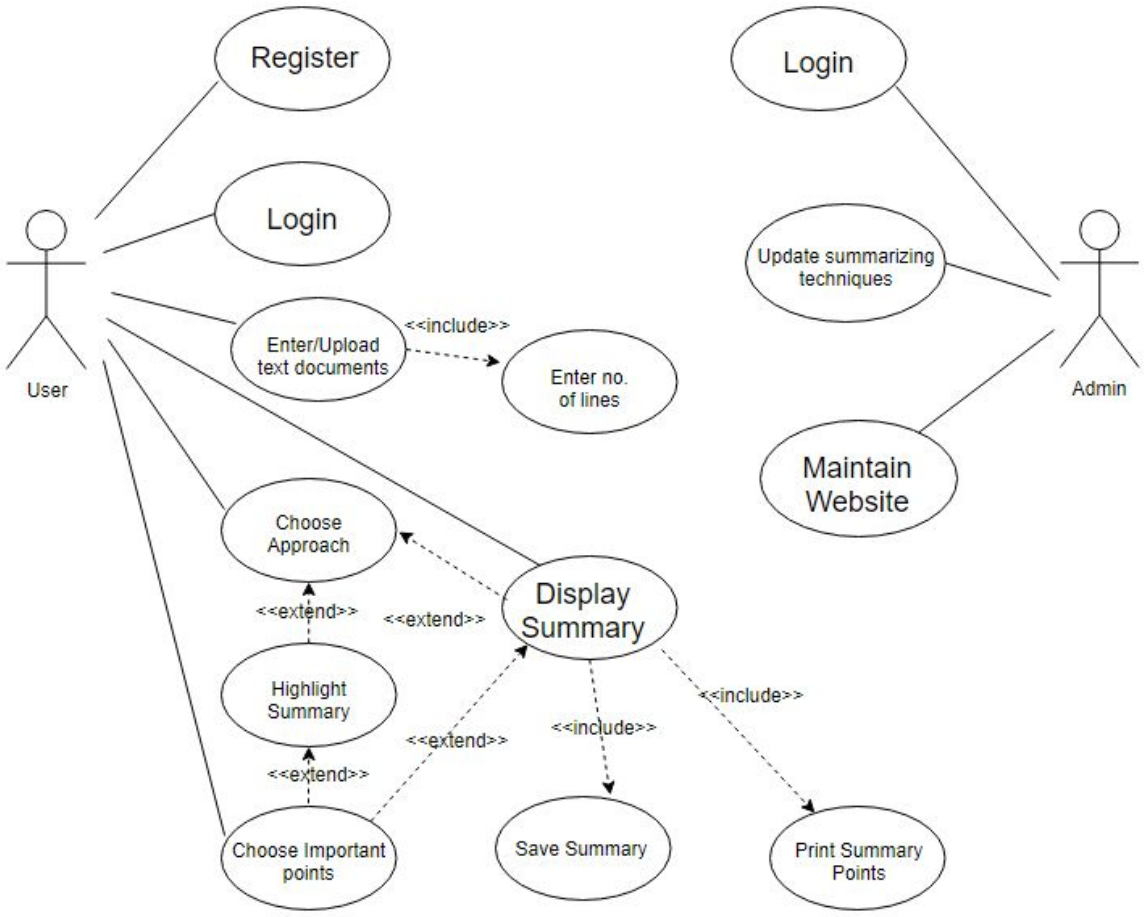

Fig 3.2.3 Use case diagram for cloud based online text summarization system

3.2.4 Hardware and Software Specifications :

The experiment setup is carried out on a computer system which has a different hardware and software specifications. This cloud based automatic text summarization can be executed on any basic system with minimal specifications. Though the hardware and software specifications of the system it was processed and executed are given in Table 3.1 and Table 3.2 respectively.

Hardware Specifications :

Processor	Intel Core i3
HDD	500 GB
RAM	4 GB

Table : 3.2.4_1 Hardware Details

Software Specifications :

Operating System	Windows 10
Technologies used	Python, HTML,CSS,Flask,Pythonanywhere

Table : 3.2.4_2 Software Details

Chapter 4

Applications

There are various applications of this domain system. The applications are listed below :

4.1 Media monitoring

The problem of information overload and “content shock” has been widely discussed. Automatic text summarization presents an opportunity to condense the continuous torrent of information into smaller pieces of information.

4.2 Newsletters

Many weekly newsletters take the form of an introduction followed by a curated selection of relevant articles. Summarization would allow organizations to further enrich newsletters with a stream of summaries (versus a list of links), which can be a particularly convenient format in mobile.

4.3 Search marketing and Search engine optimization

When evaluating search queries for Search engine optimization , it is critical to have a well-rounded understanding of what your competitors are talking about in their content. This has become particularly important since Google updated its algorithm and shifted focus towards topical authority (versus keywords). Multi-document summarization can be a powerful tool to quickly analyze dozens of search results, understand shared themes and skim the most important points.

4.4 Programming languages

There have been multiple attempts to build AI technology that could write code and build websites by itself. It is a possibility that custom “code summarizers” will emerge to help developers get the big picture out of a new project.

4.5 Patent research

Researching patents can be a tedious process. Whether you are doing market intelligence research or looking to file a new patent, a summarizer to extract the most salient claims across patents could be a time saver.

4.6 Email overload

Companies like Slack were born to keep us away from constant emailing. Summarization could surface the most important content within email and let us skim emails faster.

4.7 E-learning and class assignments

Many teachers utilize case studies and news to frame their lectures. Summarization can help teachers more quickly update their content by producing summarized reports on their subject of interest.

4.8 Meetings and video-conferencing

With the growth of tele-working, the ability to capture key ideas and content from conversations is increasingly needed. A system that could turn voice to text and generate summaries from your team meetings would be fantastic.

4.9 Help desk and customer support

Knowledge bases have been around for a while, and they are critical for SAAS platforms to provide customer support at scale. Still, users can sometimes feel overwhelmed when browsing help documents.

4.10 Social media marketing

Companies producing long-form content, like whitepapers, e-books and blogs, might be able to leverage summarization to break down this content and make it shareable on social media sites like Twitter or Facebook. This would allow companies to further re-use existing content.

Chapter 5

Summary

5.1 Summary :

Thus, we have observed the project uses various libraries to provide a summary of the given text while preserving the original meaning of the input document. For the automatic text summarizer to be more accessible the summarizer project is hosted over the cloud, so a user can easily access the website with ease and mobility by typing in just the URL. Thus the user can access the website anytime and anywhere. Once the website is accessed the required text can be copy pasted onto the site and it summarizes and provides it back to the user. The summarized report is stored in the cloud as a text file and can be downloaded from the website onto the user's personal device and be saved as a file. The analysis of the summarized report revealed that the project acquires a 50 percent retention rate over the input text. Though there is some distortion in the continuity of the information the summarizer preserves the overall idea of the input text. After hosting the website was easily available and very responsive without facing any connectivity issues over the internet.

References :

[1] Pratibha Devihosur, Naseer R "Automatic Text Summarization Using Natural Language Processing" , International Research Journal of Engineering and Technology (IRJET) , Aug 2017

[2] Mehdi Allahyari, Elizabeth D. Trippe, Saeid Safaei and others "Text Summarization Techniques: A Brief Survey " , 2017

[3] Harsh Desai, Dhairya Pawar, Geet Agrawal, "NLP Based Text Summarization Using Semantic Analysis" , International Journal of Advanced Engineering, Management and Science (IJAEMS), Oct 2016.

[4] Ankit Kumar Zixin Luo Ming Xu, "Text Summarization using Natural Language Processing", Worcester Polytechnic Institute (WPI) , March 2018.

[5] Samrat Babar, "Text Summarization:An Overview", .researchgate.net , October 2013.

[6] Towards Automatic Text Summarization: Extractive Methods, medium.com, January 2019.

[7] A Gentle Introduction to Text Summarization in Machine Learning , Floydhub, April 2019.

[8] Richa Bathija, "Data Scientist's Guide to Summarization", towards data science, March 2019

Acknowledgement

The success and final outcome of this project required a lot of guidance and assistance from many people and we are extremely privileged to have got this all along the completion of this project. All that we have done is only due to such supervision and assistance and we would not forget to thank them.

We respect and thank **Prof. Sheetal Gawande and Prof. Suresh Babu**, for providing us an opportunity to do the project and giving us all the support and guidance which made us complete the project duly.We are extremely thankful to them.

We would like to express a wealth of gratitude to **Dr. Sharvari Govilkar**, our supervisor and Head Of Department for Computer Engineering, for her constant encouragement and guidance till the completion of our project work.

We heartily thank our principal **Dr. Sandeep Joshi**, Pillai's College of Engineering for providing us with a chance to polish our skills and work on a project which would improve our skills and add more value to our resume.

We are also thankful and fortunate enough to get constant encouragement, support and guidance from all teaching staff of the Computer Engineering department which helped us in successfully completing our project work.

PV Athira Chandran
Anoop Pillai
Nikhil Chandran
Mayuresh Pitale