TEXT SUMMARIZER

-Surya Tejaswi

-1602-20-737-117

-Rishika

-1602-20-737-097

-Vishwa Vardhan

-1602-20-737-120

Problem Statement

- The abstractive based NLP react based application that we are going to develop helps facilitate our daily lives by summarising the painfully long sentences, articles, PDF files, newspapers, magazines, formal texts, academic papers or even their emails and a lot more in short, crisp and concise sentences for our users to understand. It helps almost everyone in summarising their work and making it easier for them to read instead of having them to read long paragraphs and texts. You can also summarise PDF files using this project, wherein, the text from the PDF is extracted and the desired page number can be summarized. We can also recognise texts from images using optimal character resolution, we can obtain text from an article with the URL using web scraping and also recognise voice and convert it into text and further summarise it.

Abstract

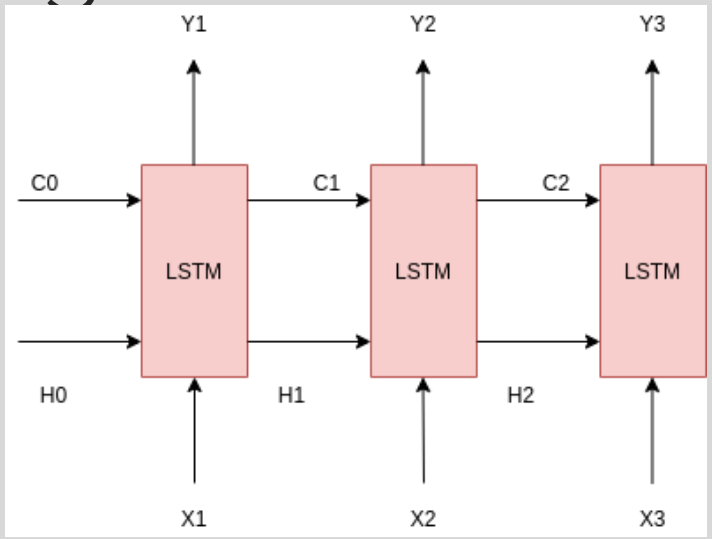
- The project basically involves summarising a block of text which is either directly provided by the user, extracted from a PDF, extracted from an image or recognised from voice. This makes it very easy for the readers who are in a hurry to read a long block of text. We use the concepts of natural language processing and NLTK library to summarise the text. We would develop an application using react native. Voice recognition and Image recognition can be done using deep learning algorithms such as Optimal character resolution etc. Users who want their emails, texts concise and to the point can also use the application to shorten their texts and the interface provided will also be very user friendly and aesthetic.

Algorithm

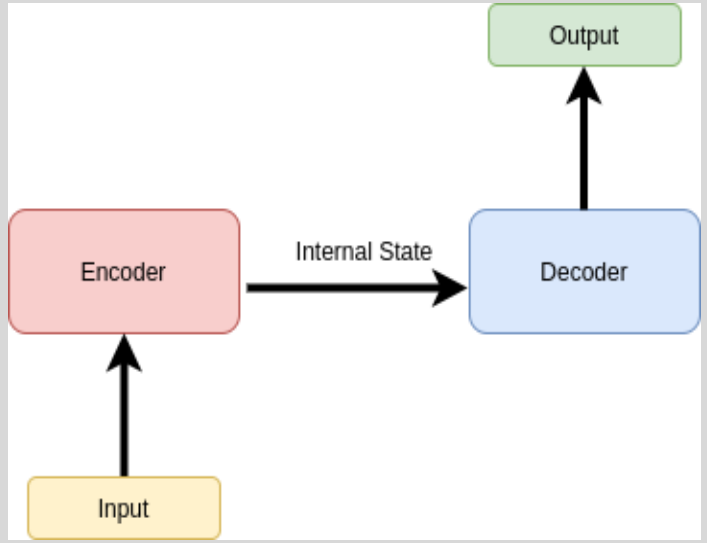
- **Step 1: Install the Transformers Library**
- **Step 2: Import Library**
- **Step 3: Setting GPU (Optional)**
- **Step 4: Setting the Model to Use, which is BART or t5 which are pre-trained CNN models**
- **Step 5: Input the Text to Summarize**
- **Step 6: Summarize using one of the above model**

Internal Architecture of abstractive algorithm

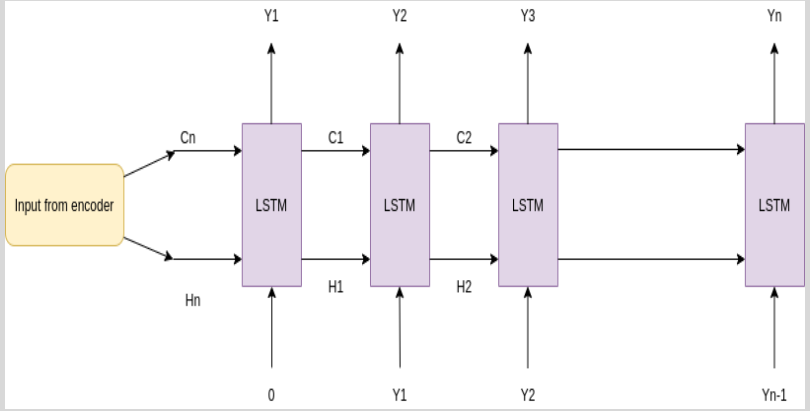
1)



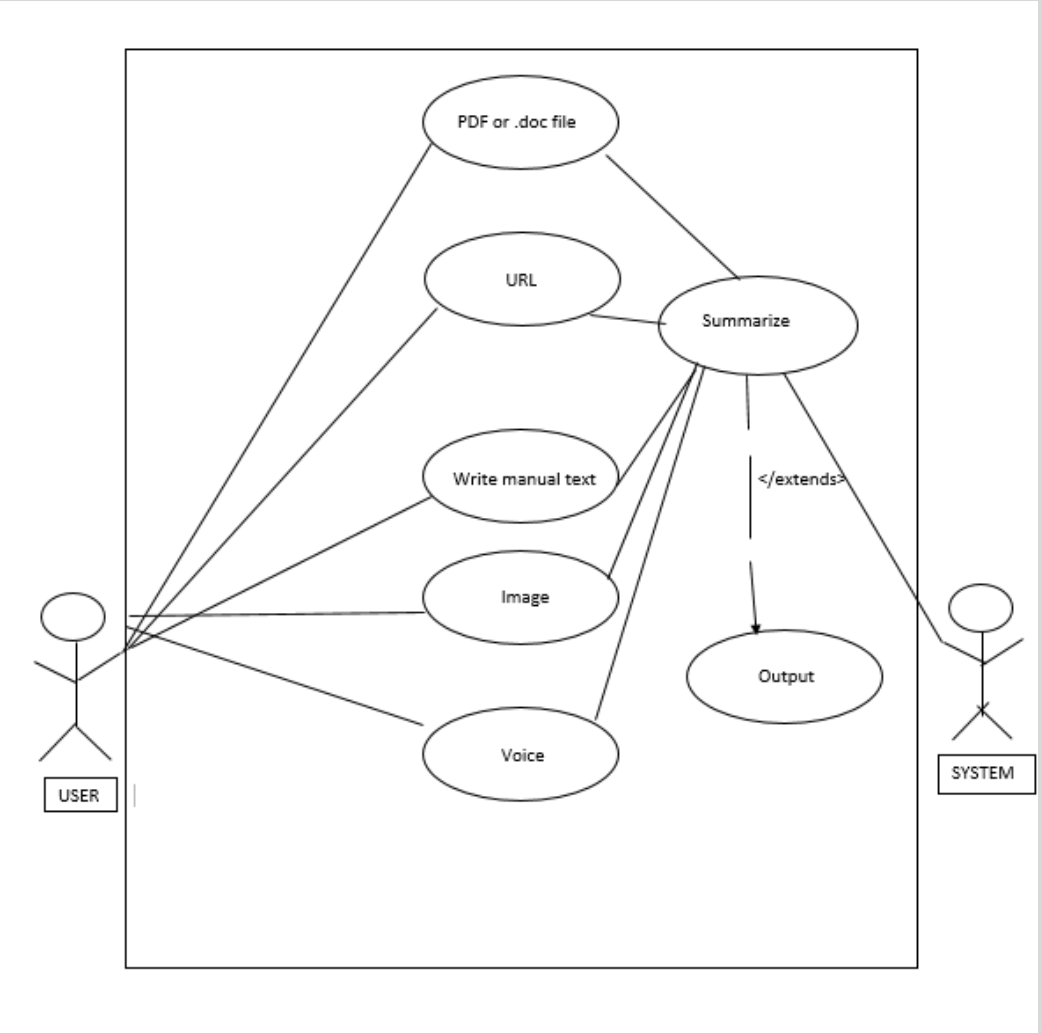
2)



3)



Use Case Diagram



Use Cases

- **>USER:**
- Gives a normal text as input
- Gives a PDF file as input
- Gives a URL as input (Extended use case)
- Gives an image as input (Extended use case)
- Gives Speech as input (Extended use case)
- **>SYSTEM:**
- Uses OCR to recognize text from image (Extended use case)
- Uses Deep learning approaches to convert speech to text (Extended use case)
- Uses the algorithm discussed above to summarize the input text

Technologies Used

- PYTHON
- REACT NATIVE
- NLP
- NTLK
- ML
- OPTIMAL CHARACTER RESOLUTION
- TRANSFORMERS LIBRARY
- CNN

THANK YOU