PDF to Audiobook

1prathi naveena, 2tankala santosh, 3pathivada pavan kumar, 4pilla srinu, 5boopana likitha

1assistant professor, 2student, 3student, 4student, 5student

Anil neerukonda institute of technology and sciences

Abstract - **Audiobooks are ideal for anyone who, like the majority of us, likes to listen rather than read. It is just not possible to purchase and store them in your home on a bookshelf. Audiobooks are also a good method to relax your eyes and take a break from the continual stimulation of digital devices. Others can take advantage of them to save time. For example, keep up with books while doing different tasks at same time. It has the potential to not only alleviate problems for millennials, but also to be a highly useful tool for visually impaired people. The ability to transform any material into an audiobook is a true gift to civilization. Our technology can be put to use in the development of such tools. Text-to-speech and other read-aloud programmes are widely utilised to assist students in developing their reading comprehension abilities. A PDF to audio system is a screen reader programme that has been designed and developed specifically for the purpose of effective audio communication. International Organization for Standardization (ISO) established PDFs as an open standard document format for the aim of displaying and transmitting information securely (ISO). One of the most convenient formats for electronic communication and information transmission is the document format. It's critical if we want to improve accessibility for screen readers by including audio into our material. Among the features PDF documents provide are text links and buttons as well as audio and video files. Many languages may be supported by the PDF to audio technology, which will allow users to hear text being read aloud (spoken).**

keywords - **Python 2.8 or above, PyPDF2, pyttsx3, Tkinter, Text to Speech Convertor**

I. INTRODUCTION

Audiobooks are recorded versions of a book's text that we listen to rather than read. Audiobooks can be literal word-for-word transcriptions of novels or shortened versions that exclude unnecessary language. For instance, we may convert short stories otherwise novels into audiobooks so that scholars can rapidly obtain a description of the work. Text-to-speech related audio reading tools are widely accepted across the world and they are implemented in an attempt to encourage students listening skills. They help children in listening short stories at bed time as parents can now convert any story stories or light novels into audiobook. PDF to audio system is a screen interpretation application designed for an effective audio communication. PDFs are an open standard document format used internationally, maintained by the International Organization for Standardization (ISO). PDF is one of the most suitable methods for electronic communication. And they are also very easy to share and exchange through electronic information conversation system. PDF documents are intended to contain links, buttons, forms, audios.

Optical-character-recognition is what is meant by the term "OCR." A generally used method for recognising text embedded in images, such as scanned documents and photos, is optical character recognition (OCR). When a picture contains written text, optical character recognition (OCR) technology is used to extract the text from the image. It is conceivably most recognised for its use of optical character recognition to turn printed paper documents into machine-readable text files. Word processors like Microsoft Word and Google Docs may be used to alter the text on a scanned paper document that has been treated using Optical Character Recognition. The following PDF to audiobook converter may benefit users by giving replacements to reading PDFs for the blind, for young people, for lethargic readers, for busy people to listen during their commute and so on and so forth. This is mainly helping to save time of the users because listening can be done more quickly than reading. Also, while travelling people prefers to listen to their PDFs rather than reading because its not relaxing while travelling. The project can help user to convert PDF file to Audiofile while also creating an text file in the process using python libraries such as PDFminer and Gtts (Google text-to-speech) and also using tinker files for GUI.

Background Study:

On July 1, 2008, the ISO issued ISO 32000-1:2008 and disseminated it as an open standard. When it came to defining the term, the responsibility was transferred to an ISO Committee of industry experts. The Adobe Public Patent License, announced in 2008, grants ISO 32000-1 royalty-free use of all Adobe-owned patents necessary for PDF-compliant implementations to be produced, used, sold, and distributed. In 2008, this licence was made available to the public.

As early as 1993, Adobe Systems released the PDF standard open source. Adobe's own PostScript format, as well as other formats like Common Ground Digital Paper and Farallon Replica were all opposing with PDF in the early years.

As a starting point, the user provides a location for a PDF file to be read to them using audio, which is a simple way to association various Python modules.

II. LITERATURE REVIEW

The goal of the study is to see how we can extract text from PDF files using various methods and to investigate these methods. Many Python modules exist, such as pdfminer, which extracts text from PDF files using OCR (Optical Character Recognition). Dumppdf is a python tool for repairing PDF files. Pseudo-XML dumps of the whole system are produced. The Python-based PDF toolkit, PyPDF2, is called PyPDF2. Document info (such as title and author) may be collected, documents can be divided and combined page by page, pages can be clipped, PDF files can be encrypted and decrypted, and much more! This research shows how listening to audiobooks increases readers' attention and keeps them interested in what they're reading. And because the mainstream of web content is available in pdf format. This could be useful if we could convert these pdf files to audio formats like mp3.

Google Text-to-Speech (Gtts) is a Python open source library. The text-to-speech API of Google Translate may be interrelated with through a Python library and command-line programme. Use the mp3 data to write to a file, a file-like object (byte string), or the standard output (stdout). Google Translate TTS request URLs may also be constructed in before and sent to an external programme. Tone, abbreviations, and decimal places are all preserved while reading long passages of text using a speech-specific sentence tokenizer. Text pre-processors that may be customised, for example, to correct articulation, are available.

It is possible to use Google text-to-speech, or gtts, on a wide range of personal digital devices, including PCs, smartphones and tablets. Text files of any form, plus Word and Pages documents, can have their contents read aloud. It is possible to read aloud even web pages that are accessed over the Internet. A computer generates the voice for TTS, and the reading speed may be adjusted to suit the user's preferences. Voice excellence varies, but roughly noises sounded like they were coming from a human. Additionally, computer-generated voices imitate the speaking patterns of children. Optical character recognition is also used by some TTS systems (OCR). A text-to-speech system can read aloud text from a picture using optical character recognition. To give you an idea, your child may take a picture of a street sign and then ask to have the words recorded.

OCR stands for "optical character recognition." Images, such as scanned documents and pictures may be recognised by text recognition software. A text picture may be converted into machine-understandable data using optical character recognition (OCR). The most common use of OCR is to convert printed paper documents into computer-readable text documents. Microsoft Word and Google Docs may be used to modify a scanned paper document's OCR-processed text.

The machine learning method creates a link between phonemes and sounds, permitting precise inflection. A sound wave producer is used in the system to produce a voice sound. The auditory model's frequency attributes are finally sent into the sound wave generator.

III. METHODOLOG

System Design

There isn't enough time in people's lives these days to read a book or use internet services to convert a PDF file to an MP3 player. As a fellow bibliophile, I have a folder where I save all of the books I'd want to read but on no occasion get around to it. As a result, I reasoned, why not turn them into audiobooks so I could listen to them while doing other work? A Python GUI programme is being developed to convert the PDF file to audio format and play it back for the user on this machine. No audio file or MP3 player is required for the application, making it easier for users to get started and utilise. After selecting a PDF file, the user can begin listening.

In contrast to distinctive audiobook convertors, they convert PDF texts (or pictures) to speech and enable volume control for single-voice conversion (either male or female). When it comes to modifying the voice, the user has only one option. They provide the ability to play and pause the video. The pace of the voice is persistent.

IV. PROJECT SCOPE

The translation of PDF text to audio signals is in high demand due to PDF's extensive use as the world's most popular document format. In addition to instructional purposes, they may be used for a variety of other tasks, such as automobile steering, train station announcements, telecommunications response services, and reading e-mail. PDF files cannot be viewed or read by those with vision damages, which is a big disadvantage. This study examines the problems of turning PDF text into spoken word. One of the ways to make PDF-based text seem more accurate when played over an audio system is to make synthetic speech sound more natural.

Problem statement :

Because PDF is the most widely used document format on the planet, there is a demand for PDF text to be converted to audio signals. In addition to instructional objectives, they may be used for a variety of other tasks, such as automobile navigation, train station notifications, telecommunications reply services, and reading e-mail. PDF files are not accessible to those with poor vision, which is a severe drawback. In this research, the difficulties of translating PDF text to voice are examined. In PDF-based writing, one challenge is to make the audio output of artificial speech sound more truthful. You'll notice when you look at our Audiobook converter that it converts PDF text (or pictures) into voice by making slight alterations inside the same PDF document. Voices may be converted. Using a single button press, the user may switch between male and female voices. Using voice speed as an example, it is possible to alter the pace at which one speaks (rapid, normal, and slow). Alternate methods of regulating the volume are available. Since the audiobook converts the PDF to text, we have complete control over what we want to include or exclude from the final product. Adding this additional competence to our audiobook converter is a first. This makes it easier for the user to add or remove lines, for example. For example, the user may use this notepad to combine or delete sections of text as they usually would. While the music is playing, we may also open the relevant PDF or any other PDF.

V. WORKFLOW:

1. When the user selects a pdf file, the pdf to audiobook converter begins the process of converting it into an audiobook.
2. After selecting a PDF file, we must type in the rate of speech per minute using the keyboard. " In this way, we are able to set the amount of words per minute that we like to hear.
3. Now we need to select the voice type, we can select male voice ore female voice
4. We can control the volume of the speech produced.
5. We can also extract .txt file using command prompt which helps in extracting the file in text format.
6. The exit button can be used to exit the application.
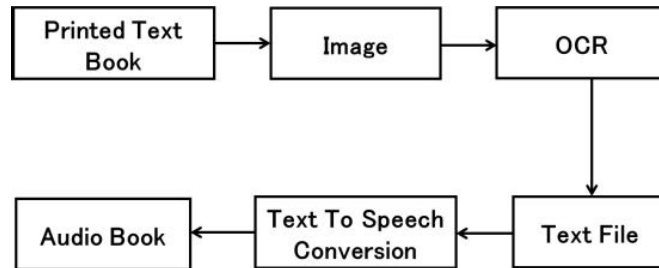
VI. WORKING

In our project, we are converting a.pdf file into a.mp3 file as a core functionality. Two libraries, pyttsx3 and pdfminer, were used to accomplish this.

- One of the most popular programmes for mining data from PDF files is PDFMiner. It's the only PDF application that concentates only on text data mining and analysis. If you use PDFMiner, you'll be able to see where text appears on a page, along with font and line information.
- A Python library for converting text to voice is available in the form of pyttsx3. If you don't have access to the Internet, this library can still be used. Extraction of data, page-by-page document splitting, merging, and more are all possible.
- The PyPDF2, pdfminer, and pyttsx3 modules will be imported.
- Using pdfminer, the text is then mined from the pdf file and saved as a text file.
- It will be transformed to an mp3 file using Google's text-to-speech module, gtts.
- Windows player will be used to play the output.

VII. SYSTEM REQUIREMENTS

For the smooth working of the application,

- Hardware Requirements
  - RAM: 4 GB or higher
  - Processor: Obtain Better Performance in i3 or above and Ryzen 3 or above
  - Hard Disk
- Software Requirements
  - Operating System: Windows Family
  - Python

VIII. ADVANTAGES OF THIS SYSTEM

- This application makes its users listening skills strong.
- The users will have ease in listening audiobooks.
- User-Friendly interface
- No prior environment required to run this application like Java, python.
- The query will be executed only in the command line interface and Graphical user Interface
- Free of cost to make this project.
- User can get text file as well as audiofile.

IX. DISADVANTAGES OF THIS SYSTEM

- The user can select any of the two voices.
- The user has to kill the program in order to stop the audio file while running in GUI.
- Works fine only in Windows family, application is not compatible with macOS and Linux.

X. ACKNOWLEDGEMENT

XI. CONCLUSION

When it comes to reading basic PDF text files, this strategy performs remarkably well. When a user selects a PDF, they should be able to convert it to audio, and then read the text to see what was read. Contribution by pupils who have trouble reading should be permitted. That this study was a success is no surprise, given how widely audiobooks are used in literateness and library programmes around the country. Students who struggle with reading might benefit from using audiobooks as a reading strategy, according to these findings. Research on the use of audiobooks with reading understanding younger and older than those who participated in the study would be of interest to us, as would discussion of the use of audiobooks by English Language Learners. Code does not currently have a stop option, but I intend to design one and include more advanced machine learning techniques into the audiobook. The ability to add features that recognise a user's voice and perform the function exactly how the user wants may be attained with machine learning. Those with incapacities, such as the blind and the disabled, will be the primary beneficiaries of this new feature.

XII. REFERENCES

[1]. R. M. K. Sinha, "A Syntactic pattern analysis system and its application to Devanagari script recognition," Ph.D. Thesis, Dept. Elect. Eng., Indian Institute of Technology, Kanpur, India, 1973.

[2]. R. M. K. Sinha and H. Mahabala, "Machine recognition of Devanagari script" IEEE Trans. Syst. ManCybern., vol. 9, no. 8, pp. 435–441, Aug.1979.

[3]. Bernard Conan Pobiak - ADJUSTABLE ACCESS ELECTRONIC BOOKS, The Reader Project ofWashington, D.C., 1992 IEEE

[4]. S. Palit and B. B. Chaudhuri, "A feature-based scheme for the machine recognition of printed Devanagari script" in Pattern Recognition, Image Processing and Computer Vision, P. P. Das and B. N. Chatterjee, Eds. New Delhi, India: Narosa Publishing House, 1995, pp. 163–168.

[5]. P. V. S. Rao and R. B. Thosar, "A programming system for studies in speech synthesis," IEEE Trans. Acoust., Speech and Signal Processing, vol. 22, no. 3, pp. 217–225, 1974.

[6]. P. Bhaskararao and S. Mathew, "Phonemic transcription rules for text-to-speech synthesis of Hindi," in In: Sinha, R.M.K. (ed) Proc. Computer Processing of Asian Languages-2. New Delhi: Tata McGraw Hill., 1992.

[7]. S. S. Agrawal, "Synthesizing Hindi speech using Klsyn and Hlsyn for natural sounding," in Workshop on Spoken Language Processing, TIFR, Mumbai, 2003.

Command Prompt
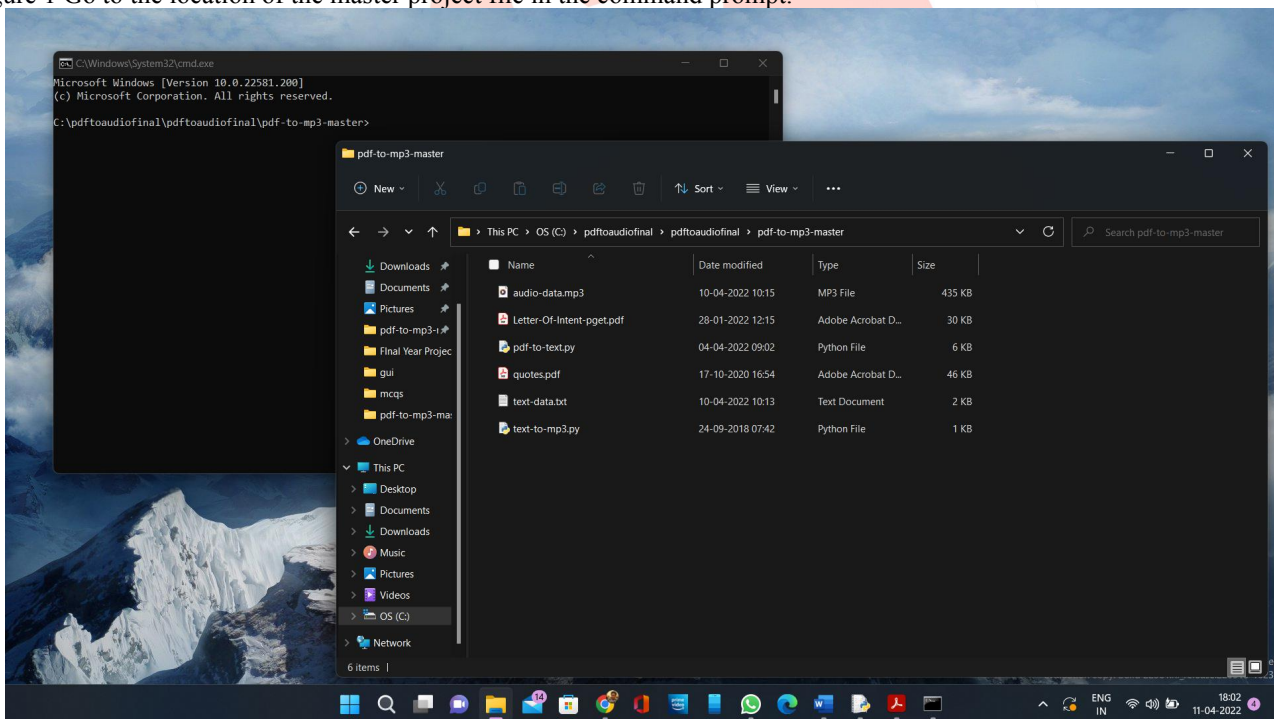Figure 1 Go to the location of the master project file in the command prompt.

Figure 2 Execute pdf-to-text.py with the preferred pdf file to extract the text from the pdf.
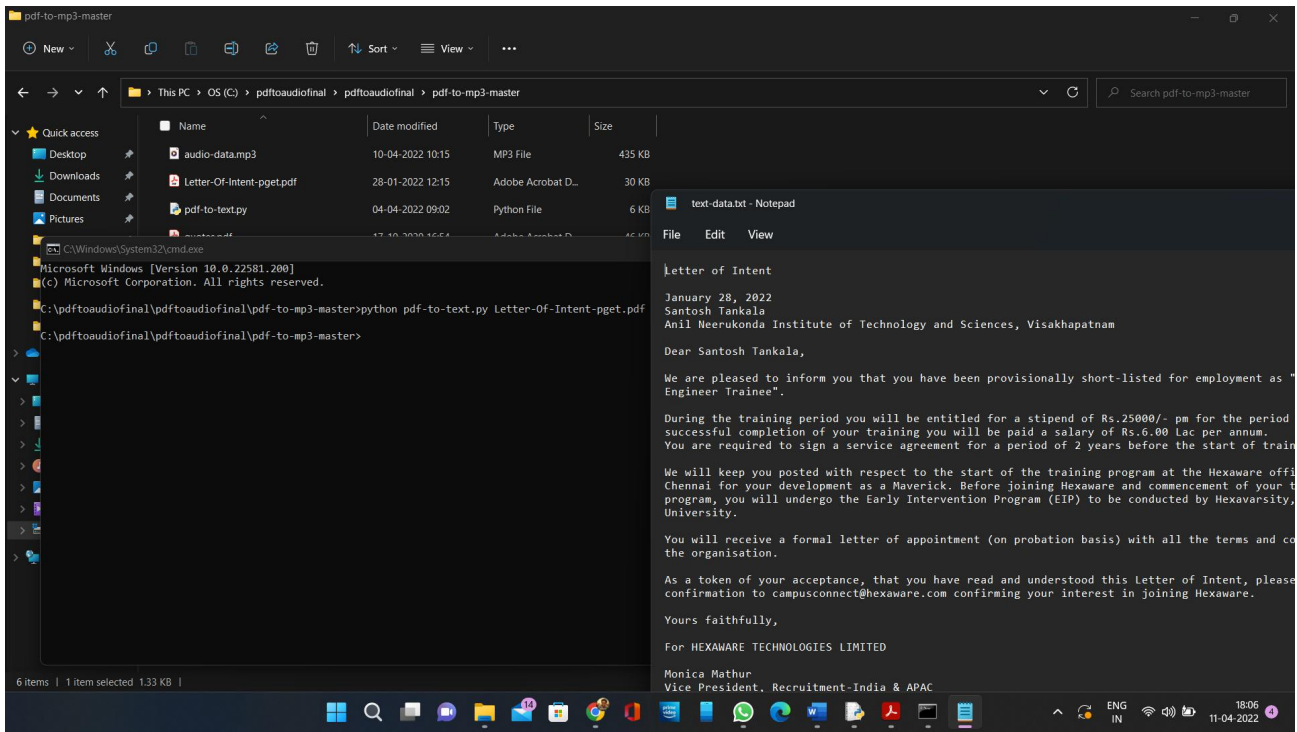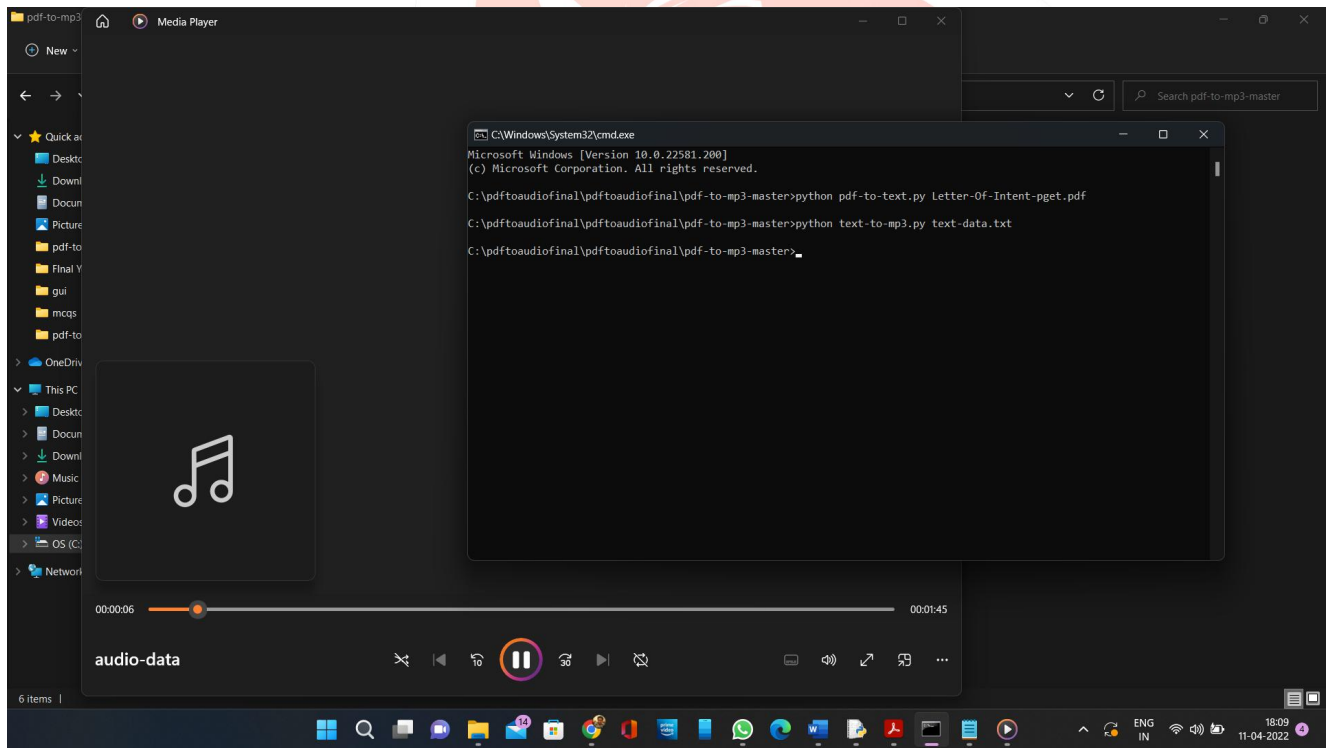Command: python pdf-to-text.py pdffilename.pdf

Figure 3
Execute text-to-mp3.py with text-data.txt to convert text to speech using google text-to-speech module.
Command: python text-to-mp3.py text-data.txt
This command will convert text to speech and windows player to play the sound.

GUI method
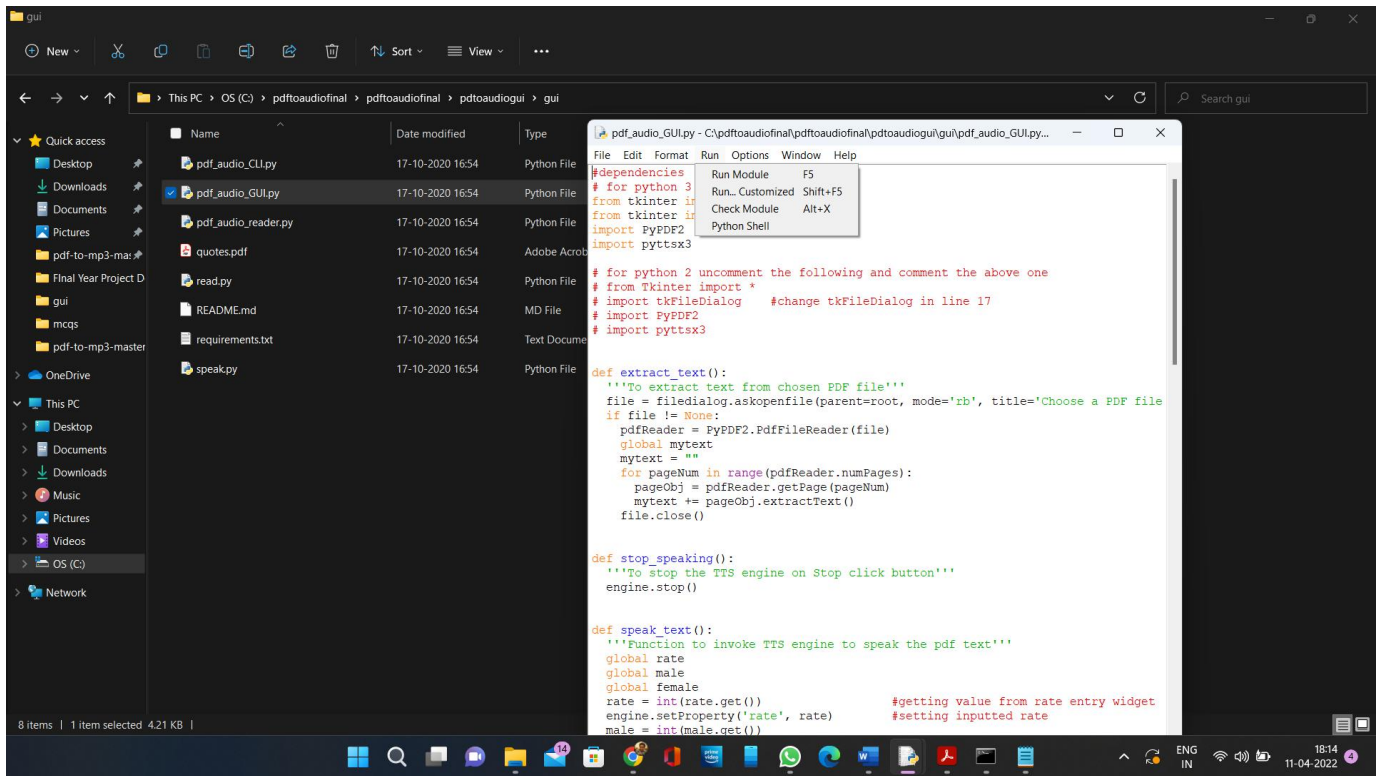Figure 4
Open pdf_audio_GUI.py using IDLE version 3.10

Figure 5
Run the module and then an Graphical User Interface will open with buttons Select PDF File, Enter rate of Speech and voice and play PDF file.
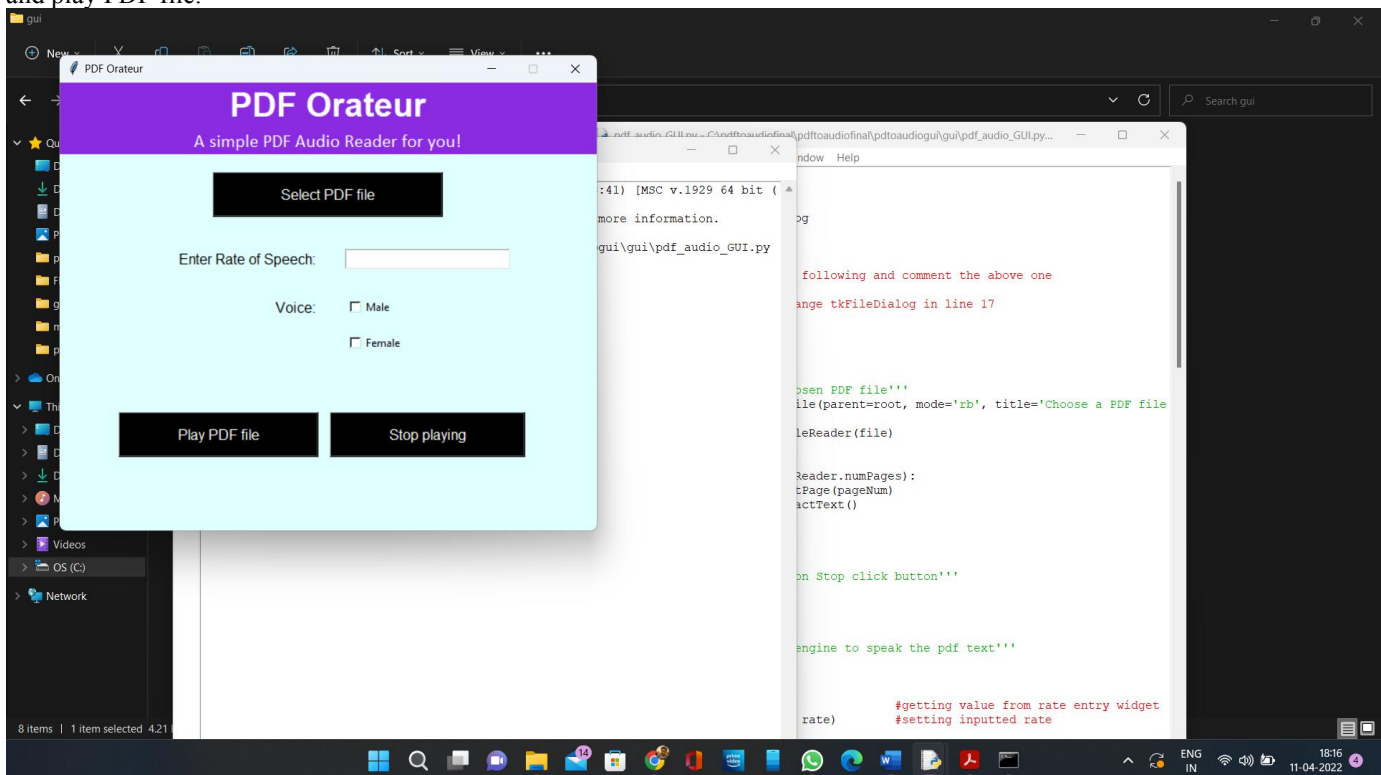

Figure 6
Select the pdf file, enter the rate of speech and select the type of voice.