

PDF to AudioBook Converter

¹Sakshi Bhargava, ² Sravya Tummala, ³Shravani S, ⁴Shivam Koul, ⁵ Nilima Kulkarni
^{1,2,3,4,5} Department of Computer Science and Engineering, MIT School of Engineering,
MIT Arts Design and Technology University, Pune,412201, India

Abstract - In this paper, the PDF to Audio Converter is proposed. It provides an alternative way to access the books and any pdf file for lazy, readers, and others. Using this PDF to Audio Converter the user will be able to listen to his/her favorite PDF and can do their daily routine. The application can be used to read any PDF which has[1].

The following application can be used to convert text from PDF to audio using Tkinter and python files, functions, and definitions. The main packages used in this audiobook converter are pyttsx3 and PyPDF2 libraries. Pyttsx3 is a python library used for text to speech conversions. This is the reason which helps the machine to speak to us. This also contains a notepad. Pyttsx3 extracts the text from the PDF. This is a python library built as a PDF toolkit. To overcome the(issues) the project PDF to Audio Converter has been developed to extract data from the pdf selected by the user, and to extract the data from the pdf, convert it to audio format to read out loud.

PDF to Audio Converter is a GUI application containing buttons for voice conversions, Volume controlling, speed adjustment and also has a feature to open a pdf, it will display the pdf while reading. The GUI is also provided with the feature of having a notepad which makes our project different from others. It will allow the user to note down important notes and it also has the feature of save, cut, copy, and paste.

Index Terms - Python, pyttsx3 and PyPDF2, Text to Speech, Converter, Tkinter.

1.INTRODUCTION

Text-to-speech and related read audio tools are being widely implemented in an attempt to assist students' reading comprehension skills.PDF to the audio system is a screen reader application designed and constructed for an effective audio communication system. PDFs were designed to present and exchange documents reliably, PDFs are an open standard document format used globally, maintained by the International Organization for Standardization (ISO). The

document format is one of the most convenient methods for electronic communication, and also for the exchange of information. Hence, there is a need to make it more accessible to readers on-screen through audio. PDF documents are designed and structured to contain links and buttons, form fields, audio or sounds, video, and business logic. The PDF to the audio system will power text on screens to read aloud (speak) with support for many languages [2]. The PDF to Audio Converter project provides an alternative to access the PDF books for the blind, lazy, readers, and others. Using this PDF to Audio Converter the user will be able to listen to his/her favorite PDF and can do their daily routine. The following application can be used to convert text from PDF to audio using Python predefined libraries [1] like pyttsx3 and PyPDF2 and also using some Tinker files.

In the current busy routine people do not have time to take a book and spend time reading it, instead, everyone needs alternative access to read the content. If a person is traveling, he/she cannot read a book, instead of reading, they can listen to it. Reading stories or essays or any text can be arduous however an audiobook would make the task easy, by reading the text. However, an audio reading of the text is convenient and does not require much concentration as reading requires. When a person tends to read a book, it requires to invest his/her time in reading. Whereas the audiobook makes the task easy, and the user can perform their own task as well as listening to the audio. In this project, we have implemented a simple PDF to audio converter using python.

When we compare with the current features present in a normal audiobook converter, they convert PDF texts (or images) into speech, and they have volume controls with single voice conversion (either male or female). Only a single choice is given to the user in case of voice modification. They provide the play and pause options. The speed of voice is always fixed.

While seeing the Audiobook converter which we have built, it converts the PDF text (or images) into speech with differences in the same PDF document. Voice conversions are possible. It contains both the male and female voice modifying features, which helps the user to change with a single click on the button. The rate of speed of voices can also be changed (fast, normal, and slow) for better clarifications. Volume controlling features are also present. The main distinguishing feature of this audiobook is that it contains a notepad. This is the very new feature added to this audiobook converter. This helps the user to take important notes and adds extra lines. This notepad facility allows the user to edit, like splitting, merging, deleting like how we normally use other notepads. We can also open the required PDF or any other PDF, while the audio is playing in the background.

The user can also save, and modify the PDF document, while the audio of the converted PDF is running in the background. This exclusive feature reduces the time consumption of writing notes and reduces paper consumption.

The objective of our project is:

In the current generation students, researchers, authors don't find time to read a book on an electronic device as that might strain their eyes and might face other issues (headache, itchiness in the eye). So, to overcome those problems we have designed an application that extracts the text from the selected PDF and reads it out to the user.

Background Study of the project is:

Adobe Systems made the PDF specification available free of charge in 1993. In the early years PDF was popular mainly in desktop publishing workflows and competed with a variety of formats such as Common Ground Digital Paper, Farallon Replica, and even Adobe's own PostScript format.

PDF was a proprietary format controlled by Adobe until it was released as an open standard on July 1, 2008 and published by the International Organization for Standardization as ISO 32000-1:2008, at which time control of the specification passed to an ISO Committee of volunteer industry experts. In 2008, Adobe published a Public Patent License to ISO 32000-1 granting royalty-free rights for all patents owned by Adobe that are necessary to make, use, sell, and distribute PDF-compliant implementations [3].

Audiobooks allow students to hear explicit sounds of letters and letter patterns that form words. Audiobooks also help students engage in text and gain exposure to more words, ultimately improving vocabulary, comprehension, and critical thinking skills [4].

In this project, we present a simple way to combine different Python libraries for creating an Audiobook that takes a PDF file path as input and reads the text in the PDF file to the user via audio. Python programming language was used to create this project.

2.LITERATURE SURVEY

Sr. No.	Name	Method	Observations
1.	Dorian Miller., 2003	Microsoft's TTS engine. Microsoft Speech Application Programming Interface (SAPI), LAME software is used [5].	The software is developed by python programming language and third-party components [5].
2.	Mentor Hamiti, Agni Dika., 2010	They have used Acoustic files of special letters and basic units albanian language [6].	In this paper, they presented that generating speech from written text in Albanian language, which helps the user those who have difficulties with their eye vision [6].
3.	Itunuoluwa Isewon, AJelili Oyelade, Olufunke	Natural Language Processing (NLP) and Digital Signal Processing (DSP) technology are used [7].	The software is called TextToSpeech Robot, it is a simple application that has text to speech functionality. The system was developed using Java programming language. Java programming

	Oladipupo., 2014		language is used because it is an independent platform [7].
4.	S. Venkateswarlu et. al, 2016	Optical Character Recognition (OCR) and Text to Speech Synthesizer (TTS) in Raspberry pi are used[8].	The present paper has introduced a real time cost beneficial technique. It has the concept of optical character recognition and text to speech synthesizer. It basically has two modules, image processing module and voice processing [8].
5.	Mr. Manohar M.,2020	Python predefined libraries are used [1].	This application can read the text from a PDF that the user has selected which has page numbers [1].

3.PROPOSED SYSTEM

In this current busy routine people do not find time to read a book, or to convert the PDF file into an MP3 player using third-party applications or web applications. Even I have a directory at which I store pdf books that I plan on reading, but I never do. So, I thought hey, why do not I make them audiobooks and listen to them while I do something else 😊! In this system, we are developing a GUI application using python to convert the PDF file into audio format and read it out to the user. The application is more user-friendly as it does not require any audio file or MP3 player. The user will have to select the PDF file which the user wants to listen to.

The problem Statement of this project is:

To create a PDF to Audiobook Converter using different Tkinter and python files, functions, and definitions. The main packages used in this audiobook converter are pytsx3 and PyPDF2. Pytsx3 is a python library for text to speech. This is the reason which helps the machine to speak to us. This reads out the selected PDF into audio in a suitable voice, speed of reading with volume adjustments. This also contains a notepad. Pytsx3 extracts the text from the PDF. This is a pure-python library built as a PDF toolkit [9]. This can extract document information, split, and merge documents page by page. This section(issues) describes the drawback of the older versions of audiobooks. To overcome the(issues) the project PDF to Audio Converter has been developed to extract data from the pdf selected by the user, and to extract the data from the pdf, convert it to audio format to read out loud.

PDF to Audio Converter is a GUI application containing play, pause(buttons), and label to display

text which allows the user to select the PDF files and the user has to click on the play button in order to extract text from the PDF file and will read the text. The application has been developed in a way that until the speaker reads out the extracted text the player cannot be paused. The GUI is also provided with a label to display the text, the text will only be displayed only after reading the extracted text.

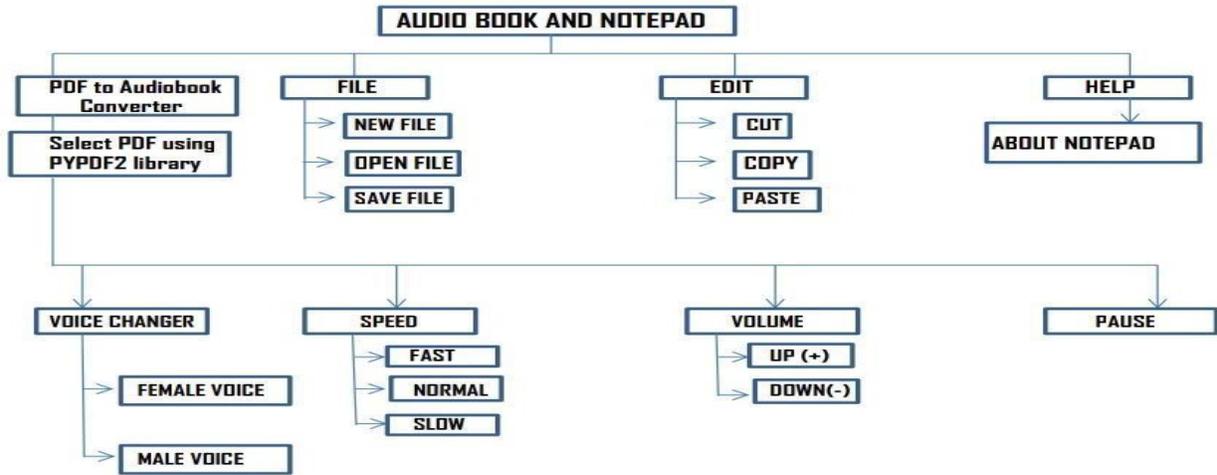
With PDF being the most used document format globally, there is a need to convert the text in PDF formats into Audio signals. These can be utilized for various purposes, e.g., in the educational system, car navigation, announcements in railway stations, response services in telecommunications, and email reading [2]. Furthermore, people with vision disabilities cannot view or read PDF files and this is a major setback. This research addresses the problems in converting PDF text into speech. One is how to improve the naturalness of synthetic speech in PDF-based text into an Audio system.

The Workflow of the project is:

- In this PDF to Audio Converter the user needs to select any PDF file from the desired location by pressing the open pdf.
- After selecting the PDF file, we have to select the type of voice we want like a female voice or a male voice.
- After selecting the type of voice, we want the program to read out the pdf in the respective voice we have selected.
- We can tune the speed and volume of speech.
- We can also write notes.
- To exit the program, we press the exit button.

Given below is the flow chart:

Proposed system - Block Diagram



Working of the proposed system is described as:

From the block diagram, we can understand the basic functionality of the project. The basic feature in our project is to convert a .pdf file into a .mp3 file. For this, we have used two libraries (i.e. pyttsx3 and pyPDF2).

1. pyttsx3 – It is a Python library for Text to Speech. It has many functions which will help the machine to communicate with us. It will help the machine to speak to us
 2. PyPDF2 – It will help with the text from the PDF and is capable of extracting document information, splitting documents page by page, merging documents page by page, etc[9].
- First, the program will import the PyPDF2 and pyttsx3 modules.
 - Then to open the Pdf fill we have used PdfFileReader().
 - Then to select the page to be read we used the getPage() method.

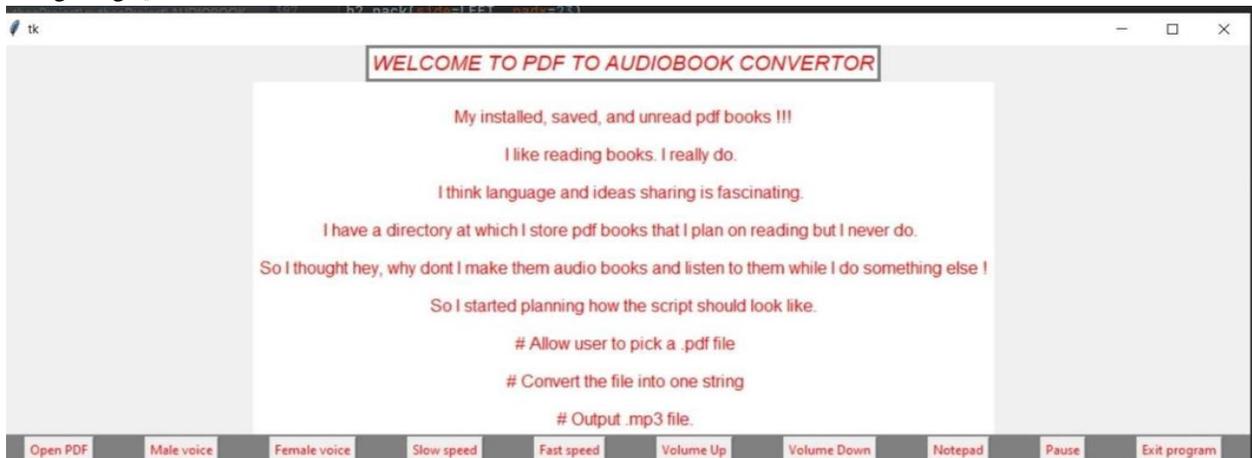
- To extract the text from the page we have used extract text.
- Then by using pyttsx3 it will convert text to speech.
- Then from the pyttx3 object, We used the say() and run wait() methods to speak out the text.

We have added a few extra features to make our project more user-friendly like voice changer, speed, volume, and a notepad.

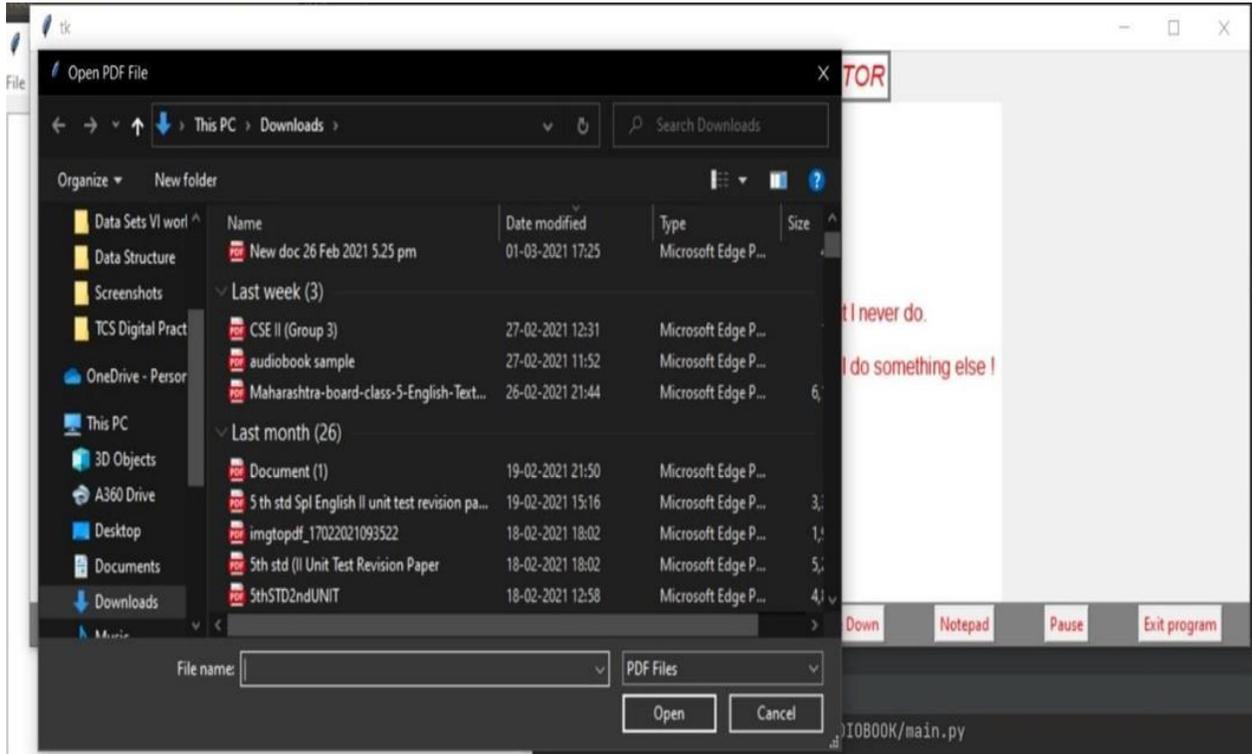
4.RESULTS

Observed Outcome of the Project:

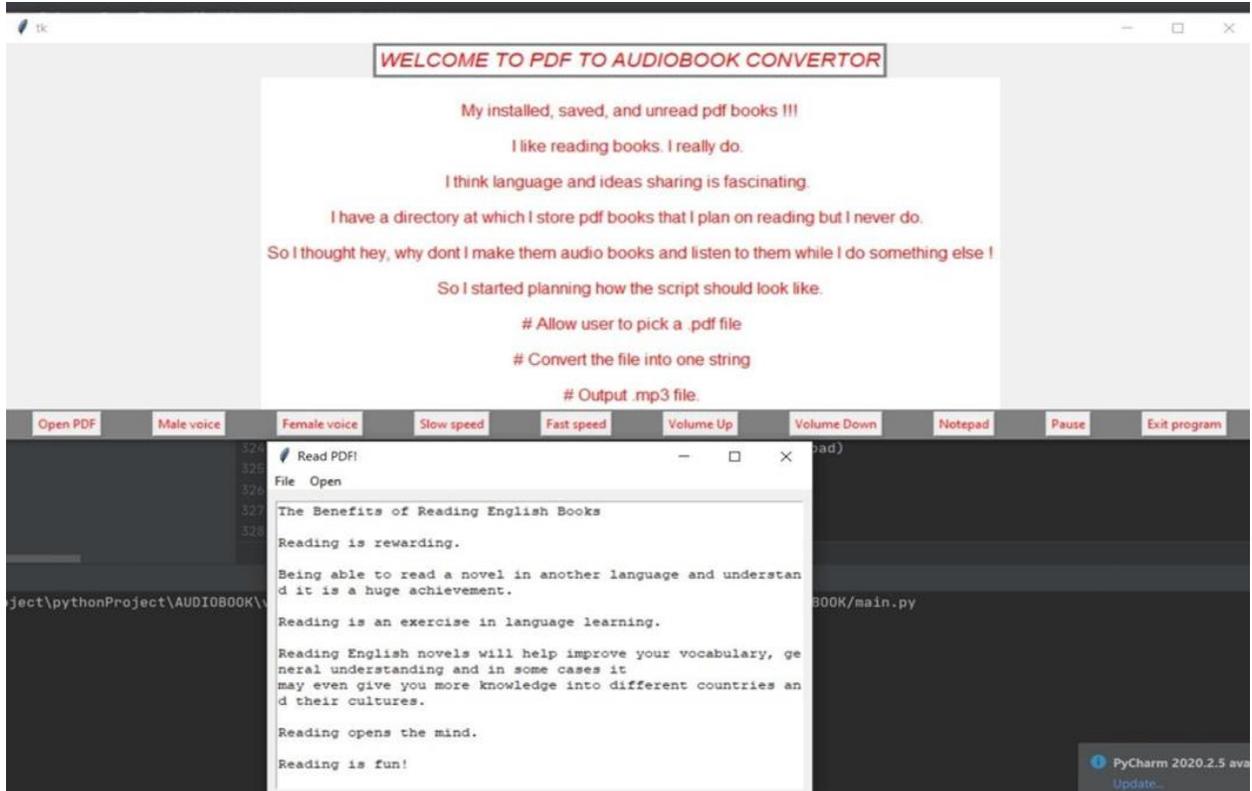
The main tab is a GUI Screen that contains buttons like open pdf, Male voice, Female voice, Slow speed, Fast Speed, Volume Up, Volume Down, Notepad, Pause, Exit Program.



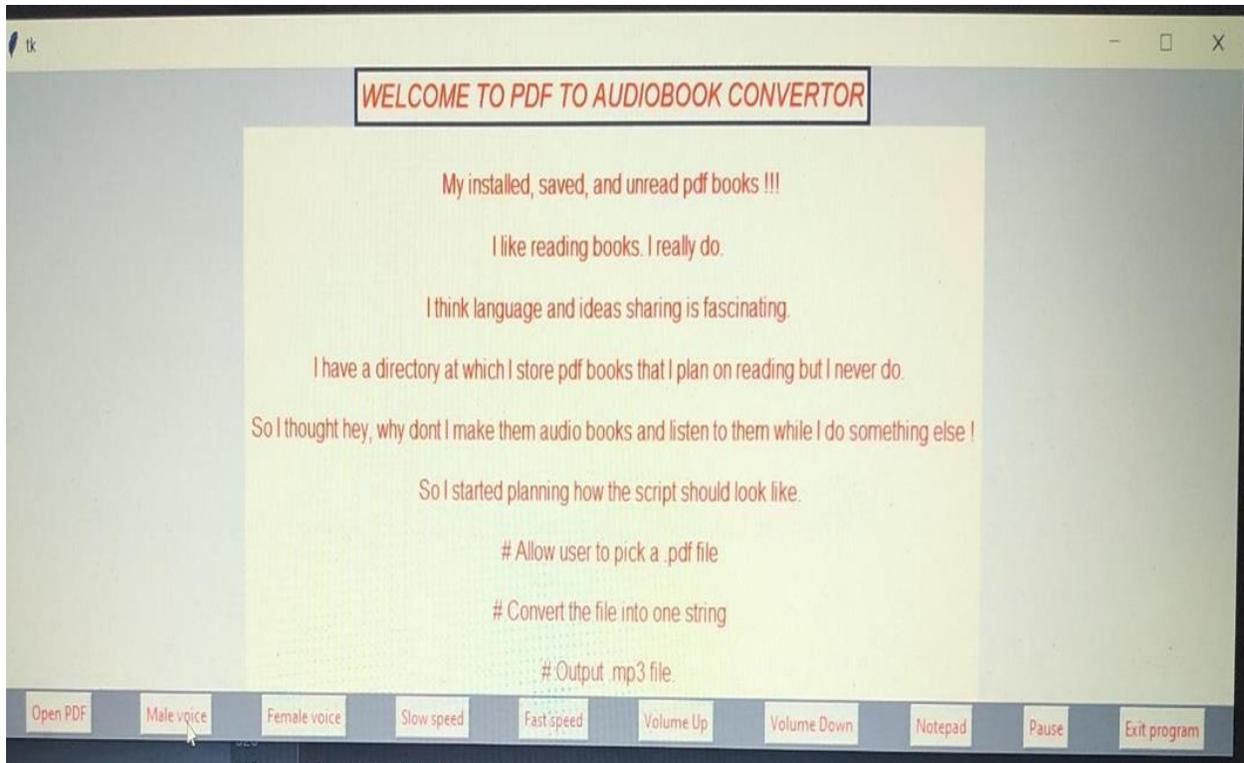
Step 1: While clicking on the Open Pdf button it will allow the user to open the pdf file to be read out. then user have to select the pdf-



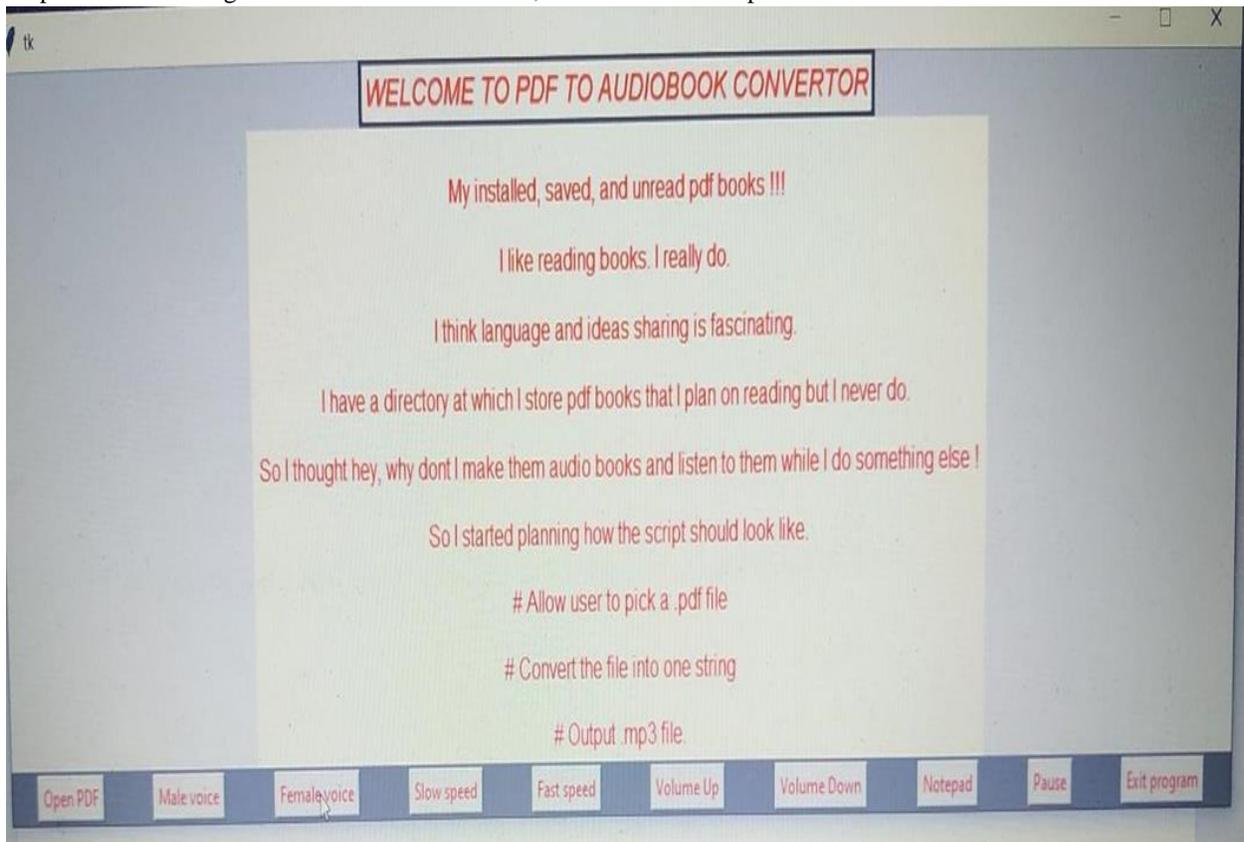
Step 2: After selecting the pdf file, the pdf file will be shown on screen as shown below –



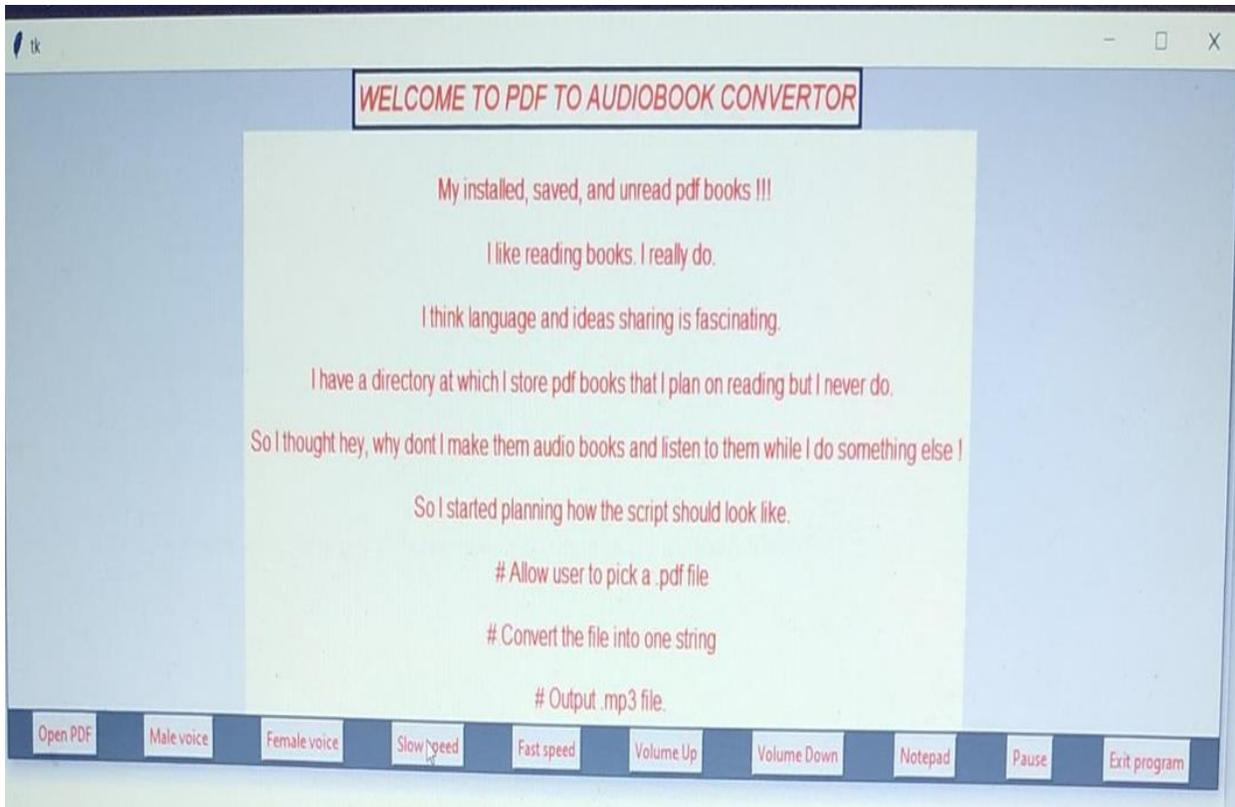
Step 3: While clicking on the Male voice button, it will read out the pdf in Male audio-



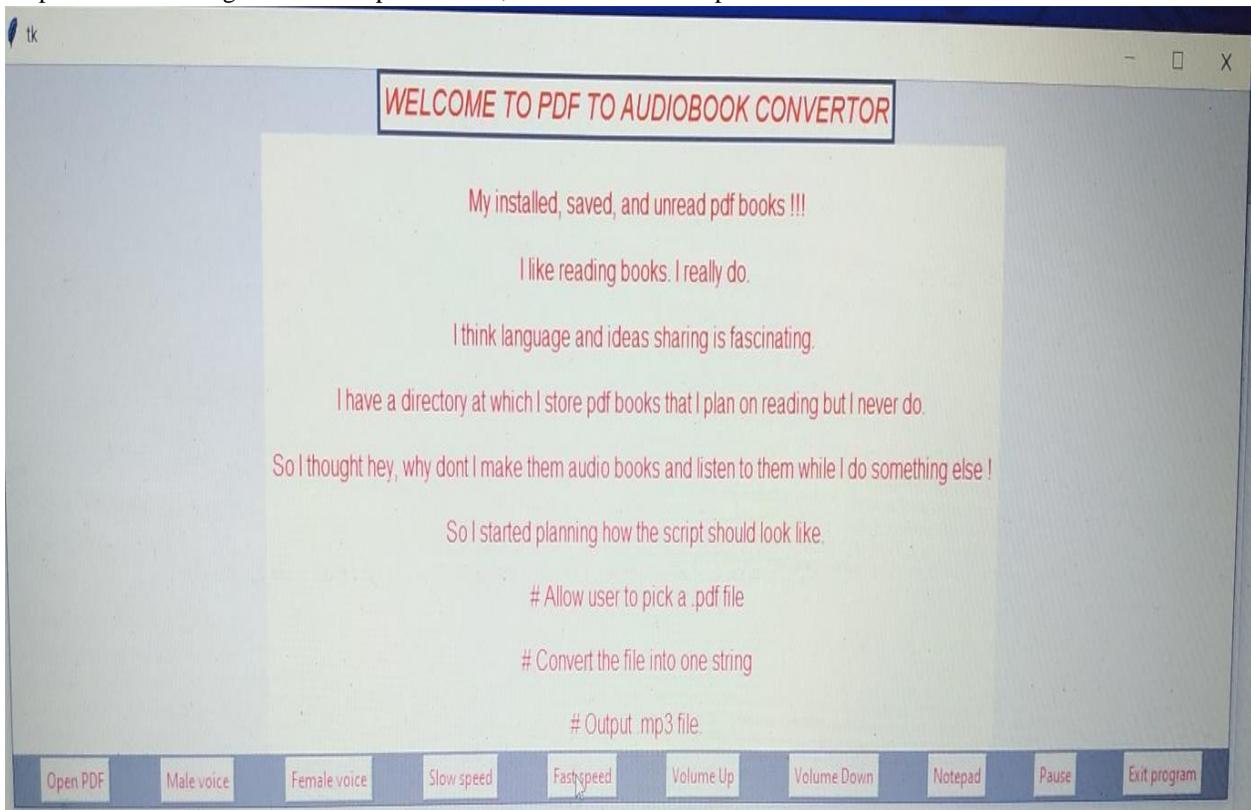
Step 4: While clicking on the Female voice button, it will read out the pdf in Female audio-



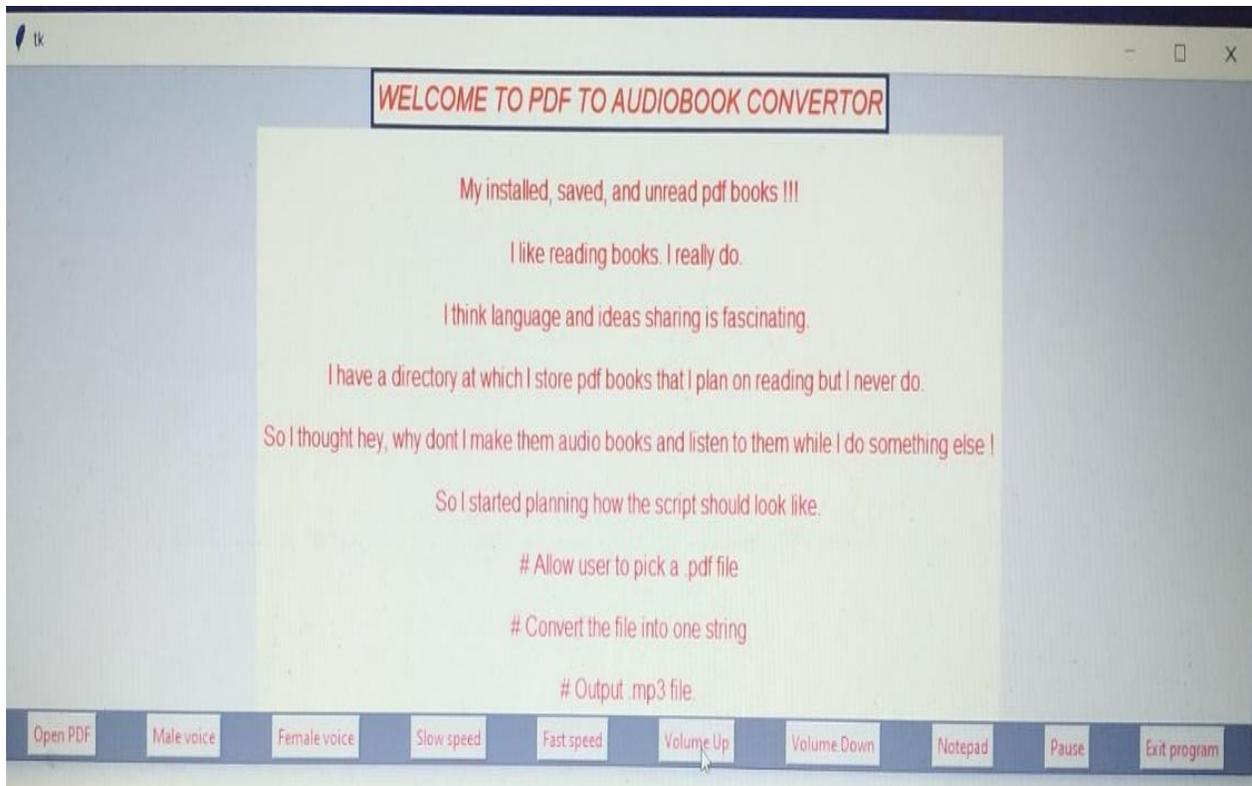
Step 5: While clicking on the Slow speed button, it will read out the pdf in Slow motion-



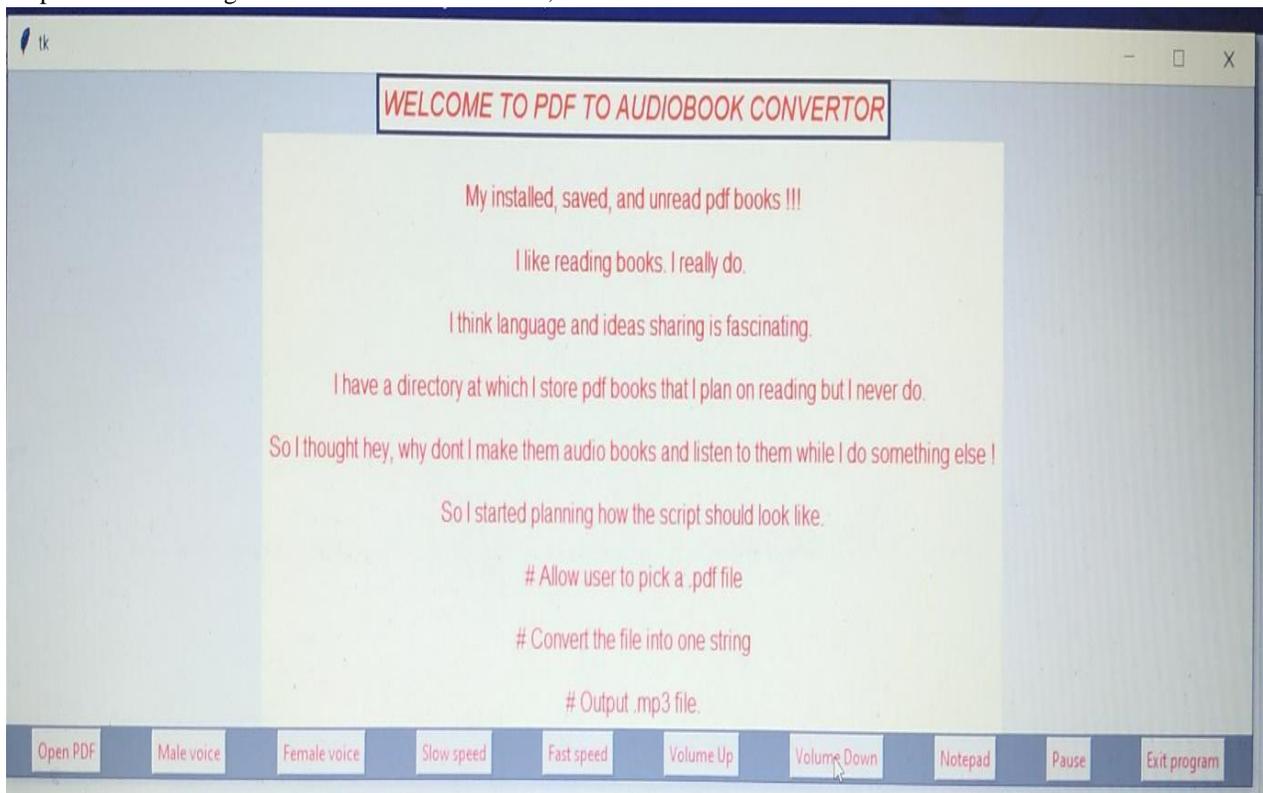
Step 6: While clicking on the Fast speed button, it will read out the pdf in Fast motion -



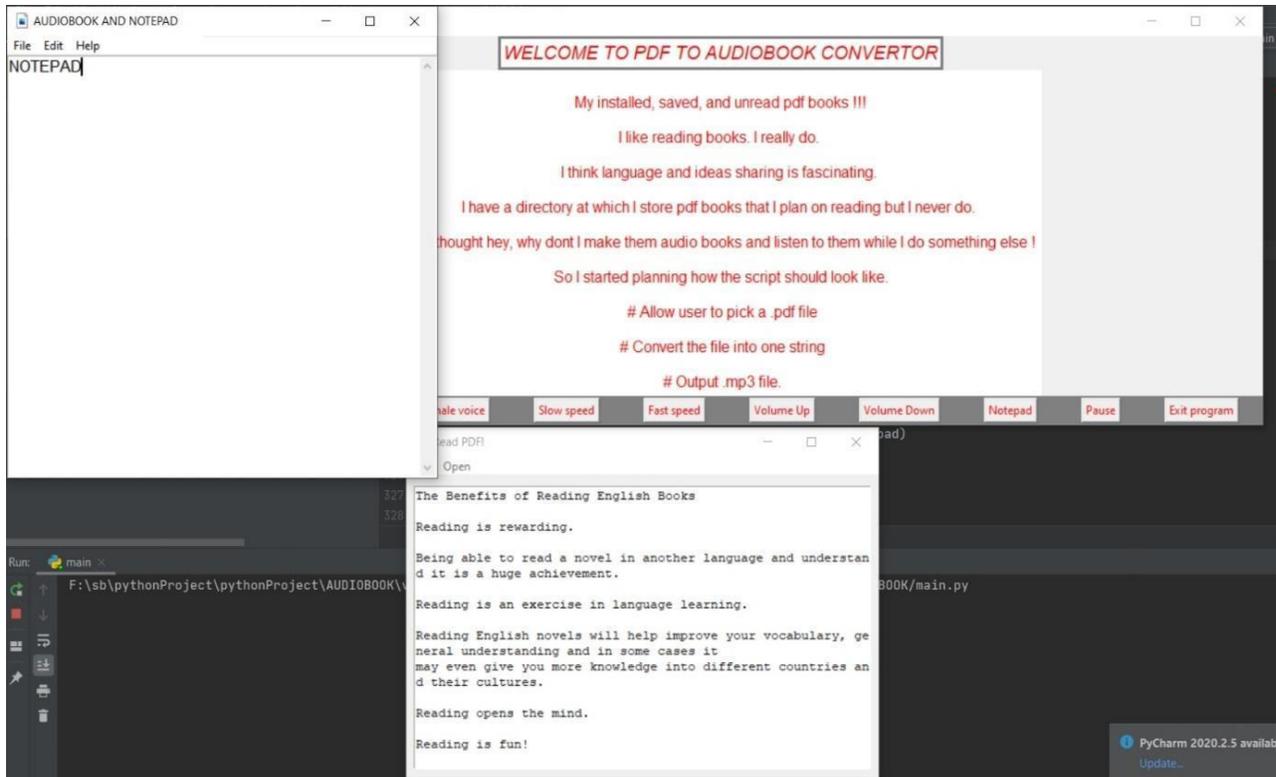
Step 7: While clicking on the Volume Up button, it will increase the volume -



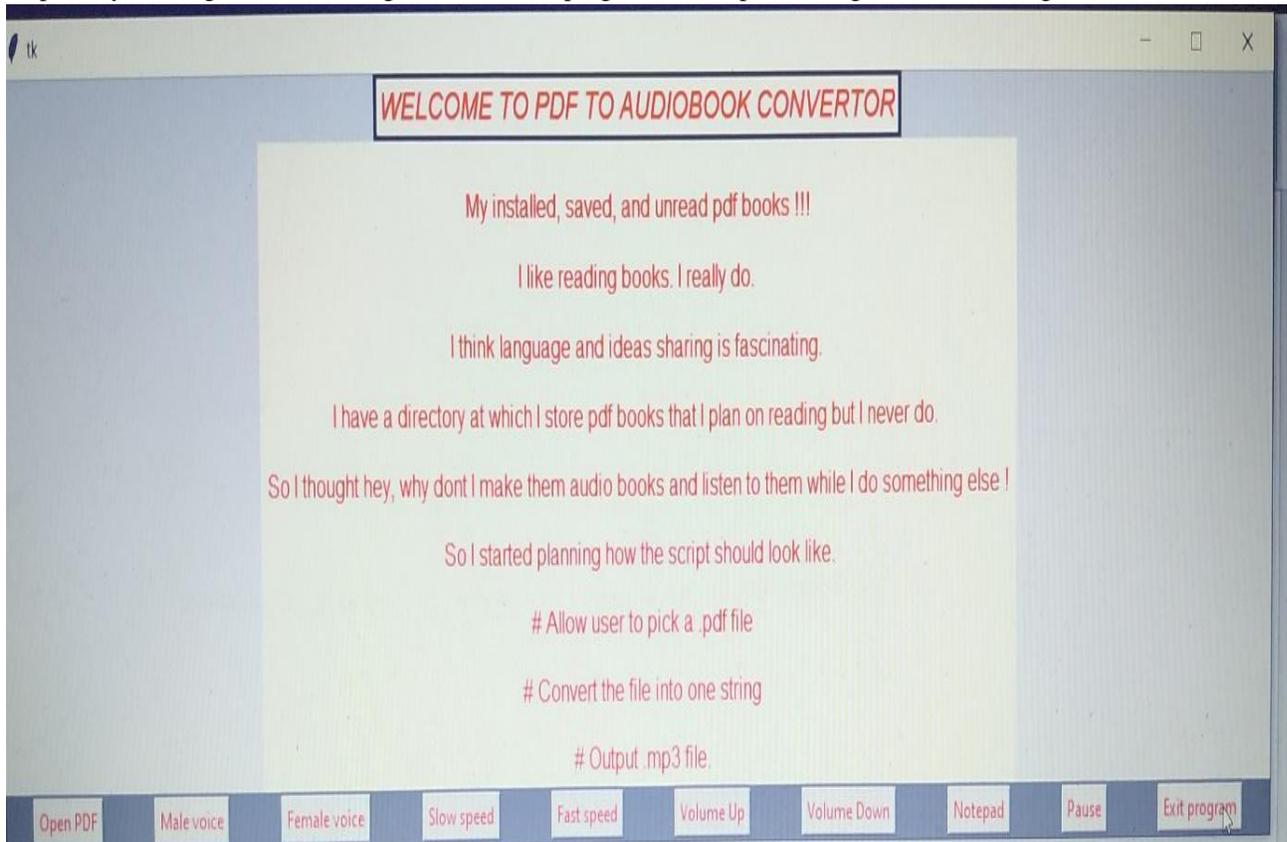
Step 8: While clicking on the Volume Down button, it will decrease the volume -



Step 9: While clicking on the Notepad button it will allow the user to note down important points and the user can save the file -



Step 10:By clicking on the Exit Program button, the program will stop executing and the tab will get closed-



5. CONCLUSION AND FUTURE WORK

It was seen that this code performs really well in reading straightforward PDF text files. Should enable users to select the desired PDF and convert it to audio and display text in, so the user can understand that particular text has been read. Should enable students with reading disabilities. The success of this research project is significant given the broad use of audiobooks in literacy and library programs across the United States. Teachers and school librarians may also use these findings as a rationale for adding audiobooks to the list of reading strategies used successfully with struggling readers. We are interested in future research on the use of audiobooks with struggling readers who are younger and older than those who participated in the study and on audiobook usage with English Language Learners [10].

At this point, the code does not have a stop feature, I intend to add those and do more interesting things with the application of Machine Learning in the audiobook. With the help of machine learning, we can add the features that will recognize the voice of the user and implement the function as the user wants [10]. This feature will help mostly for the disabled users like the blind, handicap.

REFERENCE

- [1] M. (2020). PDF TO AUDIO CONVERTOR. PDF TO AUDIO CONVERTOR, 02(12 DEC 2020), 563-566.
- [2] Final Year Research Project Topics & Materials In PDF & Doc | iproject from IPROJECT. (n.d.). Design and implementation of pdf to audio system computer science proj. Retrieved March 09, 2021, from <https://iproject.com.ng/computer-science/design-and-implementation-of-pdf-to-audio-system/index.html>
- [3] Pdf. (2021, March 08). Retrieved March 09, 2021, from <https://en.wikipedia.org/wiki/PDF>
- [4] 7 ways Audiobooks benefit students who struggle with reading. (n.d.). Retrieved March 09, 2021, from <https://learningally.org/Solutions-for-School/-7-Ways-Audiobooks-Benefit-Students-Who-Struggle-With-Reading>
- [5] Miller, D. (n.d.). Pdf to mp3 converter. Retrieved March 09, 2021, from http://www.cs.unc.edu/Research/assist/et/projects/mp3/final_pdf2mp3.html

- [6] Hamiti, M., & Dika, A. (2010). Learning opportunities through generating speech from written texts. *Procedia - Social and Behavioral Sciences*, 2(2), 4319-4324. doi:10.1016/j.sbspro.2010.03.686
- [7] Isewon, I., Oyelade, J., & Oladipupo, O. (2014). Design and implementation of text to speech conversion for visually impaired people. *International Journal of Applied Information Systems*, 7(2), 25-30. doi:10.5120/ijais14-451143
- [8] Chan, N. (n.d.). Text-to-speech conversion. doi:10.5353/th_b3120958
- [9] Kumar, S. (2020, October 12). Build your Own audiobook in 7 lines of Python code. Retrieved March 09, 2021, from <https://towardsdatascience.com/build-your-own-audiobook-in-7-lines-of-python-code-bfd805fca4b1>
- [10] Bhatti, S. (2020, March 30). Designing an audiobook using python. Retrieved March 09, 2021, from <https://medium.com/analytics-vidhya/designing-an-audiobook-using-python-cfa03b5f0759>