A

Minor Project Report on

## Multi Linguistic PDF to Audio Converter And Compressor

Submitted in partial fulfillment of the requirements for the degree of

BACHELOR OF ENGINEERING

IN

### Computer Science & Engineering

### Artificial Intelligence & Machine Learning

by

Ayush Bhandari (23106117)

Vedant Aher (23106123)

Aparajit Aherkar (23106131)

Jayraj Dhole ( 23106114)

Under the guidance of

## Prof. Kanchan Wankhede



### Department of Computer Science & Engineering

### (Artificial Intelligence & Machine Learning)

**A. P. Shah Institute of Technology**

**G. B. Road, Kasarvadavali, Thane (W)-400615**

**University Of Mumbai**

**2024-2025**

## A. P. SHAH INSTITUTE OF TECHNOLOGY

## CERTIFICATE

This is to certify that the project entitled “**Multi Linguistic PDF to Audio Convertor And Compressor”** is a bonafide work of Ayush Bhandari (23106117), Vedant Aher (23106123), Aparajit Aherkar (23106131), Jayraj Dhole (23106114) submitted to the University of Mumbai in partial fulfillment of the requirement for the award of **Bachelor of Engineering** in **Computer Science & Engineering (Artificial Intelligence & Machine Learning).**

|  |  |
| --- | --- |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| Prof. Kanchan Wankhede | Dr. Jaya Gupta |
| Mini Project Guide | Head of Department |

## 

## A. P. SHAH INSTITUTE OF TECHNOLOGY

## Project Report Approval

This Mini project report entitled “**Multi Linguistic PDF to Audio Convertor And Compressor*”*** by **Ayush Bhandari, Vedant Aher, Aparajit Aherkar and Jayraj Dhole**is approved for the degree of ***Bachelor of Engineering*** in ***Computer Science &Engineering***, (AIML) ***2024-25.***

##### External Examiner:

##### Internal Examiner:

Place: APSIT, Thane

Date:

**Declaration**

##### We declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

|  |  |  |  |
| --- | --- | --- | --- |
| Ayush Bhandari | Vedant Aher | Aparajit Aherkar | Jayraj Dhole |
| (23106117) | (23106123) | (23106131) | (23106114) |

#### ABSTRACT

In today's digital age, the widespread use of Portable Document Format (PDF) files has made information sharing and storage more efficient. However, challenges persist for certain groups, particularly those with visual impairments or reading disabilities, who may find it difficult to access text-based content. Furthermore, large PDF files with extensive formatting and embedded images often become cumbersome to share, particularly in areas with limited internet bandwidth or storage capacity.

This project presents a comprehensive solution through the development of a Multi Linguistic PDF to Audio Converter and Compressor. The system is designed to address two key issues: accessibility and file size management. First, the system converts the textual content of PDF documents into high-quality audio, supporting a wide range of languages. This feature significantly enhances the accessibility of information, allowing users, including those with visual impairments, to listen to content in their preferred language.

Second, the project incorporates an advanced PDF compression algorithm, which reduces the size of the original PDF document without compromising its integrity. This compression ensures that large files are made more manageable for storage and distribution, particularly in environments where bandwidth and

**Keywords:** PDF to Audio Conversion, Multilingual Support, PDF Compression, Accessibility, Text-to-Speech, Natural Language Processin

**Index**

|  |  |  |  |
| --- | --- | --- | --- |
| Index | | | Page no. |
| Chapter-1 | | |  |
|  | Introduction | | 1 |
|  |  |  |  |
| Chapter-2 | | |  |
|  | Literature Survey | |  |
|  | 2.1 | History | 2 |
|  | 2.1 | Review | 2-4 |
|  |  |  |  |
| Chapter-3 | | |  |
|  | Problem Statement | | 5 |
|  |  |  |  |
| Chapter-4 | | |  |
|  | Experimental Setup | |  |
|  | 4.1 | Hardware setup | 6 |
|  | 4.2 | Software Setup | 6 |
|  |  |  |  |
| Chapter-5 | | |  |
|  | Proposed system and Implementation | |  |
|  | 5.1 | Block Diagram of proposed system | 7 |
|  | 5.2 | Description of Block diagram | 7 |
|  | 5.3 | Implementation | 8-9 |
|  | Figures | |  |
|  | 5.1 | Block diagram of proposed system | 7 |
|  | 5.2 | Home page | 8 |
|  | 5.3 | Pdf to audio converter page | 8 |
|  | 5.4 | Pdf compressor page | 9 |
| Chapter-6 | | |  |
|  | Conclusion | | 10 |
|  |  |  |  |
| References | | | 11 |
|  |  |  |  |

### INTRODUCTION

In the modern digital age, the widespread adoption of Portable Document Format (PDF) has revolutionized how information is created, shared, and stored across various sectors, from education and business to government and healthcare. PDFs are favored for their ability to preserve the formatting and integrity of documents across different platforms and devices, making them a universal standard for digital documentation. However, despite their widespread use, PDFs present significant challenges, particularly in terms of accessibility and file management. For individuals with visual impairments or learning disabilities, accessing the textual content within PDFs can be a daunting task, as these documents are predominantly visual in nature. Moreover, the ever-increasing complexity of PDFs—often embedded with high-resolution images, graphics, and intricate layouts—leads to large file sizes that are cumbersome to store, share, and download, especially in areas with limited internet bandwidth or digital storage capabilities.

This project seeks to address these challenges by developing a Multi L3inguistic PDF to Audio Converter and Compressor, a comprehensive tool designed to enhance both the accessibility and manageability of PDF documents. The system is engineered to convert the textual content of PDFs into high-quality audio, supporting multiple languages to cater to a diverse user base. This feature is particularly beneficial for visually impaired users, language learners, and anyone who prefers auditory consumption of information. By enabling users to listen to the content in their preferred language, the system significantly enhances the inclusivity and reach of digital documents.

In addition to improving accessibility, the system incorporates a robust PDF compression module that effectively reduces the file size of PDFs without compromising the quality of the text or embedded media. This compression feature is crucial in making large PDF files more manageable for storage and easier to share, particularly in environments with bandwidth constraints or limited storage resources. The integration of natural language processing (NLP) techniques ensures accurate language detection and text processing, while advanced text-to-speech (TTS) engines generate clear, natural-sounding audio. The PDF compression algorithm is designed to optimize file size while maintaining the visual fidelity of the document, ensuring that the original content remains intact and accessible.

By combining these functionalities, the Multi Linguistic PDF to Audio Converter and Compressor offers a versatile and user-friendly solution that addresses the dual needs of accessibility and file efficiency. The system is poised to have broad applications across various fields, from education, where it can support language learning and provide accessible materials for students, to professional settings where document sharing and storage efficiency are paramount. This project not only aims to bridge the gap between digital documentation and accessibility but also strives to create a more inclusive digital environment where information is easily accessible and shareable, regardless of the user's physical abilities or technological constraints. Through the successful implementation of this project, we hope to contribute to a more connected and accessible world, where digital documents are no longer barriers but gateways to information for everyone.

#### LITERATURE SURVEY

###### 2.1-HISTORY

The adoption of Portable Document Format (PDF) has become a cornerstone of digital documentation due to its ability to maintain consistent formatting across different devices and platforms. However, the challenges associated with the accessibility and storage of PDFs, particularly for individuals with visual impairments or in regions with limited digital infrastructure, have prompted extensive research in areas such as text-to-speech (TTS) conversion and file compression. Early developments in TTS technology began with rudimentary speech synthesis systems that produced robotic and unnatural speech. Over time, advancements in natural language processing (NLP) and machine learning have enabled the creation of more sophisticated TTS engines capable of generating natural-sounding audio in multiple languages. These advancements have made it possible to convert text into high-quality speech, thus improving accessibility for users who rely on auditory consumption of information.

Simultaneously, the field of file compression has seen significant progress, particularly in the context of managing large digital documents like PDFs. Initial compression techniques focused primarily on reducing the size of image files within PDFs, which often contribute to the bulk of the file size. Over the years, more advanced compression algorithms have been developed, allowing for the efficient reduction of file size while preserving the quality of the document’s content. This evolution in both TTS and compression technologies has laid the groundwork for the integration of these capabilities into comprehensive solutions that address both accessibility and file management challenges in digital documents

#### 2.2-LITERATURE REVIEW

1. **Survey of Existing Solutions**
2. **PDF TO AUDIO CONVERTER (Mr. Manohar M.) ( IRJMEETS Volume:02 /Issue:12 /December -2020):**

The PDF to Audio Converter project provides an alternative was to access the PDF books for 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 application can be used to read any PDF which has page numbers. The following application can be used to convert text from PDF to audio using Python predefined libraries.

1. **Convert PDF to Audiobook using Optical Character Recognition and Machine Learning. (** **Mohammad Adil Sheikh) ( JETIR January 2021, Volume 8, Issue 1) :**

Audiobooks are Favorable for everyone who is always zealous. It is just not possible to buy store them in a bookshelf in your house. Audiobooks can also be a great way to ease your eyes, a rest from the constant charge of digital screens. It can not only solve the problems for millennials but can also be very beneficial tool for visually impaired person. The power to convert any documented file to audiobook is nothing but a pure gift to society.

1. **Advance Pdf To Audio Converter (Vanshika Sharma) ( EATP 2024,Volume 30, Issue 5) :**

The Advance PDF to Audio Converter project introduces innovative solution for enhancing accessibility to PDF books, catering to individuals with visual impairments, those seeking convenience, and a diverse range of readers. This tool empowers users by enabling them to listen to their preferred PDF content while engaging in daily activities. Notably, the application supports the conversion of PDFs with page numbers, ensuring versatility in its utility. Leveraging Python and predefined libraries such as pyttsx3 and PyPDF2,this converter seamlessly transforms text into audio, merging technology with user-friendly design. The integration of a user-friendly interface enhances the overall experience, making the conversion process straightforward.

**2. Analysis of Issues in Surveyed Solutions:**

Despite the availability of these tools, several issues persist:

**Lack of Integration**: Existing tools typically specialize in either text-to-speech conversion or PDF compression, but not both. Users must employ multiple tools to achieve both accessibility and file management goals.

**Limited Multilingual Support:** While some TTS engines support multiple languages, the quality of speech generation can vary, and there is often limited support for complex languages or dialects.

**3. Survey of Users and Potential User Requirements**

Users of these solutions, particularly those in the following categories, have specific requirements and expectations:

**Visually Impaired Users:** Require seamless and high-quality audio conversion to access PDF content.

**Educational Institutions:** Need tools that support multiple languages and dialects, and which can be used to convert educational materials into audio for diverse student populations.

**Users in Low-Bandwidth Areas:** Require efficient PDF compression to facilitate easy sharing and storage of documents, especially in regions with limited digital infrastructure.

**General Users:** Prefer tools that are easy to use, with integrated features that do not require switching between multiple applications.

**4. General Observations**

Overall, the literature and existing solutions highlight a significant gap: the absence of a unified tool that seamlessly integrates PDF-to-audio conversion with PDF compression. Most solutions in the market either address accessibility through TTS or manage file size through compression, but not both. Users are often forced to compromise between accessibility and efficiency, which can be cumbersome and inefficient. The integration of these functionalities into a single, user-friendly tool could significantly enhance the accessibility and usability of PDFs, particularly for visually impaired users and those in resource-constrained environments. The proposed project aims to bridge this gap by developing a comprehensive solution that meets both needs simultaneously.

#### PROBLEM STATEMENT

In the current digital era, PDF files have become a ubiquitous format for sharing and storing information across various fields, including education, business, and government. However, these documents often present significant challenges for users, particularly those with visual impairments or those in regions with limited digital infrastructure. The primary issues revolve around the accessibility of text content within PDFs and the often large file sizes that make these documents difficult to share or store efficiently.

While PDFs are designed to preserve formatting and layout across different devices, they are inherently visual, making them difficult for visually impaired users to access. Although screen readers and other assistive technologies exist, they often fall short in providing a seamless reading experience, especially when the text is embedded within images or complex layouts. Furthermore, most existing text-to-speech (TTS) solutions are limited in their support for multiple languages, reducing their effectiveness in a globalized world where documents are often multilingual.

Given these challenges, there is a clear need for a comprehensive solution that addresses both the accessibility and file management issues associated with PDF documents. This project proposes the development of a Multi Linguistic PDF to Audio Converter and Compressor. The proposed system will integrate advanced text extraction and language detection techniques to convert the content of PDFs into high-quality audio across multiple languages, thereby improving accessibility. Additionally, the system will incorporate efficient compression algorithms to reduce the size of the original PDFs, making them easier to share and store without compromising on quality. This dual functionality aims to provide a seamless, user-friendly tool that enhances both the accessibility and efficiency of managing PDF documents.

#### EXPERIMENTAL SETUP

#### 4.1 Hardware Setup

The development and testing of the system were carried out on a computer with the following specifications:

* **Processor:** Intel Core i5 or higher
* **RAM:** 8 GB or more
* **Storage:** 256 GB SSD or higher
* **Operating System:** Windows 10 or Linux Ubuntu 20.04 LTS

#### 4.2 Software Setup

#### The project utilized the following software tools and libraries:

#### Programming Language: Python 3.8

#### NLP Libraries: NLTK, SpaCy, and transformers for language detection and processing.

#### TTS Engines: gTTS (Google Text-to-Speech) for audio conversion, with potential upgrades to more advanced engines like Tacotron.

#### PDF Handling: PyMuPDF for PDF text extraction and pdfsizeopt for compression.

#### GUI Framework: Tkinter for building the user interface.

#### PROPOSED SYSTEM & IMPLEMENTATION

#### 5.1 Block diagram of proposed system

#### 

#### Figure 5.1 Block diagram of proposed system

#### 5.2 Description of block diagram

#### PDF Importer: This module is responsible for reading and importing the PDF document into the system.

#### Language Detector: Automatically detects the language of the text within the PDF, which is crucial for selecting the appropriate TTS engine.

#### Text Extractor: Extracts text content from the PDF, ensuring it is clean and ready for conversion.

#### TTS Engine: Converts the extracted text into audio, supporting multiple languages to cater to a diverse audience.

#### PDF Compressor: Reduces the file size of the original PDF using advanced compression algorithms without significant loss of quality.

#### Output Module: The final module presents the user with both the compressed PDF and the corresponding audio file, ready for use or distribution.

#### 5.3 Implementation

#### Home Page:

****

Figure 5.2 Home Page

Pdf to Audio Converter Page:

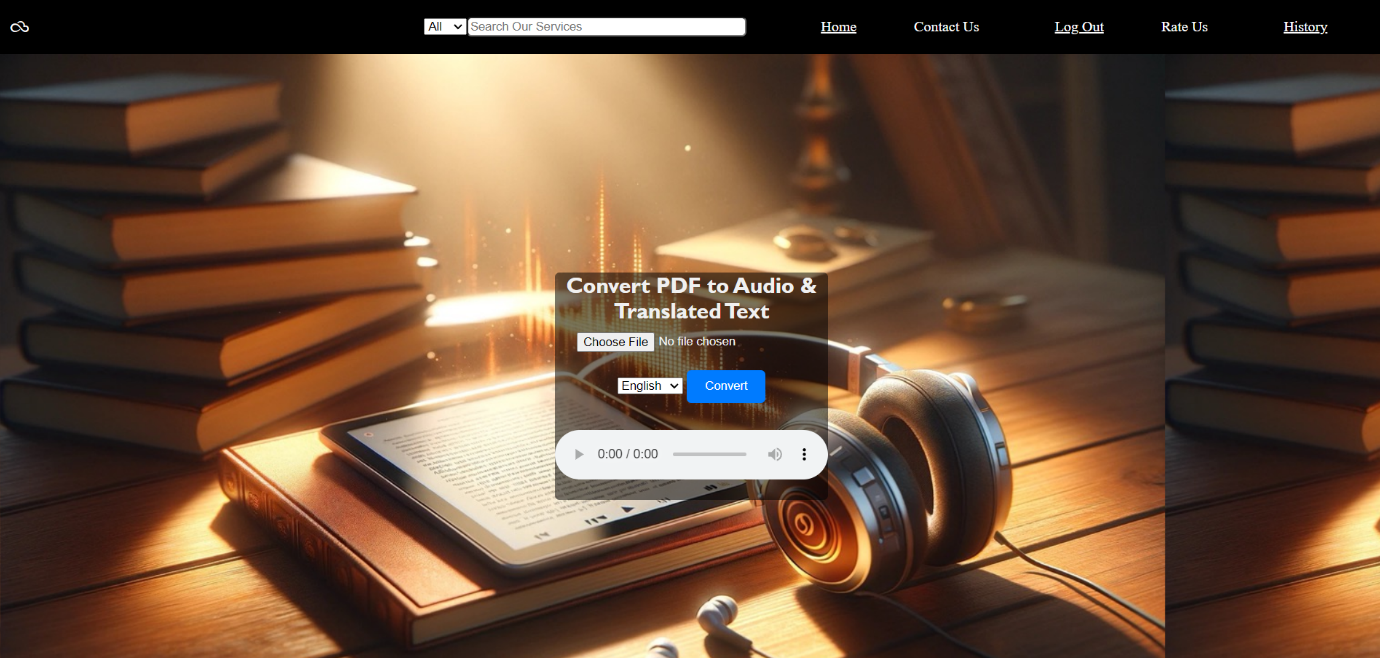
****

Figure 5.3 Pdf to Audio Converter Page

Pdf Compressor Page:

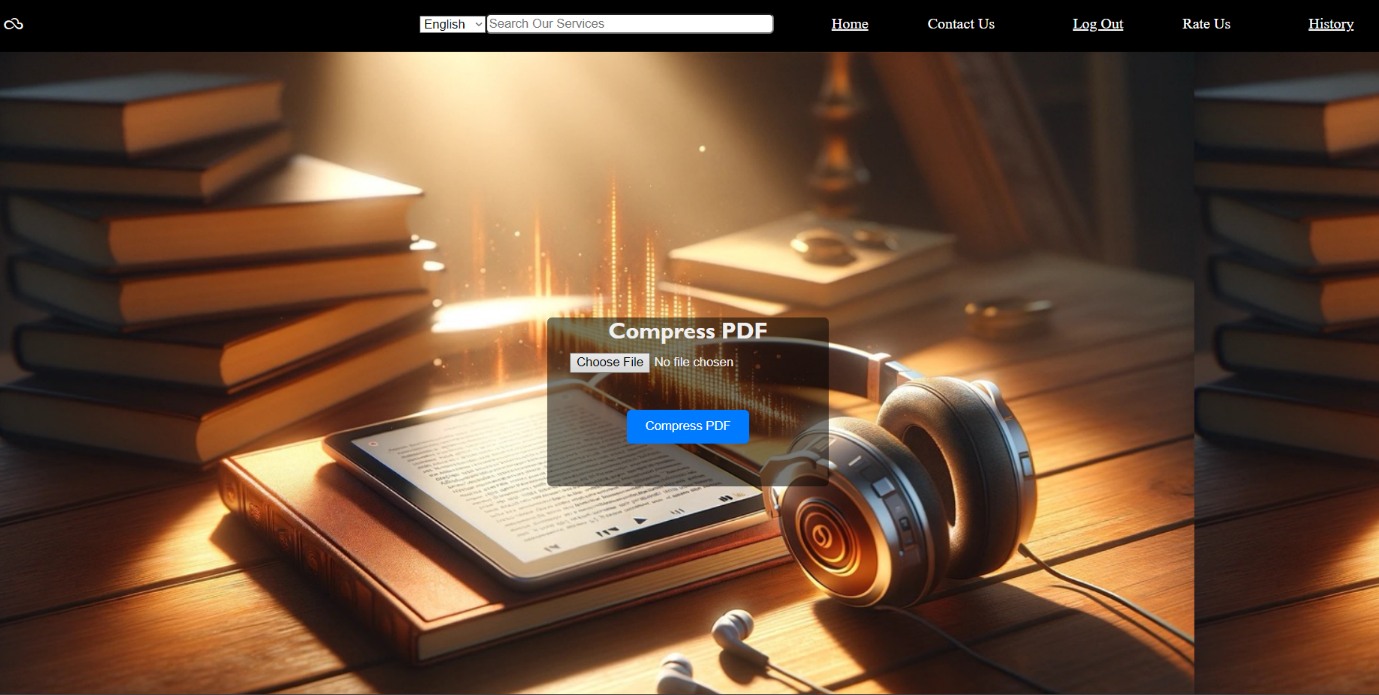
****

Figure 5.4 Pdf compressor page

**The implementation Text phase involved the following steps:**

1. **Extraction**: Using libraries like PDF.JS to accurately extract text from PDF files.
2. **Language Detection:** Implementing language detection using GOOGLE CLOUD SERVICES the langdetect library to ensure proper pronunciation and voice selection.
3. **Speech Synthesis:** Integrating Google TTS for initial testing, followed by more advanced TTS systems like Tacotron for better quality in different languages.
4. **Audio Compression:** Employing PDF.COcompress the audio without losing quality.
5. **User Interface**: Developing a simple user interface using HTML and CSS , where users can select files, choose languages, and manage output settings.

**5.4** **Advantages**

The system offers a combined solution for PDF compression and text-to-audio conversion, making it highly versatile and practical.

#### ****5.5 Application****

#### Ideal for educational purposes, accessibility improvements, and efficient document sharing, especially in regions with limited internet bandwidth.

1. **CONCLUSION**

#### 6.1 Summary of Work

This project successfully developed a comprehensive tool that addresses both the accessibility and storage challenges associated with large PDF documents. By integrating multi-linguistic text-to-speech conversion with PDF compression, the system offers a unique and practical solution for users worldwide.

#### 6.2 Future Scope

Future developments could include:

* **Expanded Language Support:** Adding more languages and dialects, including those with complex scripts.
* **Enhanced Compression Techniques:** Exploring more advanced compression algorithms to further reduce file size without sacrificing quality.
* **Mobile Application Development:** Creating a mobile version of the tool to increase accessibility and convenience for users on the go.

# REFERENCES

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. M. Bells, J. E. Lhotak and N. M. A., "Facilitating text-to-speech conversion of a domain name or a network address containing a domain name," US Patent, Google Patents, 2012, pp. 8,185,396.
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>