RSET
RAJAGIRI SCHOOL OF
ENGINEERING & TECHNOLOGY
(AUTONOMOUS)

Mini Project Report On

Talkbook:Advancing Book to Audiobook Conversion through Emotion Aware Machine Learning

Submitted in partial fulfillment of the requirements for the award of the degree of

Bachelor of Technology

in

Computer Science & Engineering

By

Nandhana Suffin (U2103148)

Nikhil Stephen (U2103155)

Niveditha B (U2103162)

Rachel Jacob (U2103168)

Under the guidance of

Mr Biju Abraham Narayamparambil

Department of Computer Science & Engineering
Rajagiri School of Engineering & Technology (Autonomous)
(Affiliated to APJ Abdul Kalam Technological University)
Rajagiri Valley, Kakkanad, Kochi, 682039
May 2024

CERTIFICATE

This is to certify that the mini project report entitled ***"Talkbook:Advancing Book to Audiobook Conversion through Emotion Aware Machine Learning "*** *is a bonafide record of the work done by* ***Nandhana Suffin (U2103148)***, ***Nikhil Stephen (U2103155)***, ***Niveditha B (U2103162)***, ***Rachel Jacob (U2103168)***, *submitted to the APJ Abdul Kalam Technological University in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology (B. Tech.) in Computer Science and Engineering during the academic year 2023-2024.*

Mr. Biju Abraham Narayamparambil
Asst. Professor
Dept. of CSE
RSET

Dr. Uma Narayanan
Asst. Professor
Dept. of CSE
RSET

Dr. Preetha KG.
Professor and HoD
Dept. of CSE
RSET

ACKNOWLEDGEMENTS

We wish to express our sincere gratitude towards Dr P. S. Sreejith, Principal of RSET, and Dr. Preetha K.G., Head of the Department of Computer Science and Engineering for providing us with the opportunity to undertake our mini project, "Talkbook:Advancing Book to Audiobook Conversion through Emotion Aware Machine Learning".

We are highly indebted to our project coordinator, **Dr Uma Narayanan**,Asst Professor, Department of Computer Science and Engineering for her valuable support.

It is indeed our pleasure and a moment of satisfaction for us to express our sincere gratitude to our project guide **Mr Biju Abraham Narayamparambil** for his patience and all the priceless advice and wisdom he has shared with us.

Last but not the least, we would like to express our sincere gratitude towards all other teachers and friends for their continuous support and constructive ideas.

Nandhana Suffin

Nikhil Stephen

Niveditha B

Rachel Jacob

Abstract

This project aims to revolutionize the process of converting written text into audio books by leveraging Machine Learning (ML) techniques.Unlike conventional methods such as the read-aloud feature of PDF viewers, our approach transcends mere text-to-speech conversion by integrating sophisticated emotion recognition and voice modulation algorithms.

Key to our methodology is the incorporation of ML models trained to detect and interpret the emotional nuances embedded within the text. By analyzing factors such as tone, sentiment, and context, our web application dynamically adjusts the narrator's delivery to accurately convey the intended emotions, enriching the auditory experience for listeners.

Furthermore, our system implements advanced voice modulation capabilities, allowing for seamless transitions between different characters and genders.

In essence, this project represents a significant leap forward in the field of book-to-audiobook conversion. By harnessing the power of ML, our software empowers users to effortlessly produce high-quality audiobook versions of prose, stories, and books, distinguished by their emotive depth and lifelike narration.

Contents

Acknowledgements **i**

Abstract **ii**

List of Figures **v**

List of Tables **vi**

List of Abbreviations **vii**

1 Introduction **1**

1.1 Background 1

1.2 Problem Definition 1

1.3 Scope and Motivation 2

1.4 Objectives 2

1.5 Challenges 3

1.6 Assumptions 3

1.7 Societal / Industrial Relevance 4

1.8 Organization of the Report 5

2 Software Requirements Specification **9**

2.1 Introduction 9

2.2 Overall Description 10

2.3 External Interface Requirements 12

2.4 System Features 14

2.5 Other Nonfunctional Requirements 14

3 System Architecture and Design **17**

3.1 System Overview 17

3.2 Architectural Design 19

3.3 Dataset identified 20
3.4 Proposed Methodology/Algorithms 21
3.5 User Interface Design 21
3.6 Description of Implementation Strategies 24
3.7 Module Division 25
3.8 Work Schedule - Gantt Chart 26

4 Results and Discussions **27**
4.1 Overview 27
4.2 Testing 27
4.3 Quantitative Results 30
4.4 Graphical Analysis 32
4.5 Discussion 35

5 Conclusion **37**
5.1 Conclusion 37
5.2 Future Scope 38

Appendix A: Presentation **41**

Appendix B: Vision, Mission, Programme Outcomes and Course Outcomes **55**

Vision, Mission, POs, PSOs and COs **2**

Appendix C: CO-PO-PSO Mapping **7**

List of Figures

2.1 Flowchart 10

3.1 ARCHITECTURAL DIAGRAM 18
3.2 ARCHITECTURAL DESIGN 19
3.3 SEQUENCE DIAGRAM 19
3.4 SAMPLE DATASET 20
3.5 Home Page 21
3.6 select the mode of input 22
3.7 Upload pdf as input 22
3.8 Uploading Text Page 23
3.9 About Page 23
3.10 Social Connect Page 24
3.11 GANTT CHART 26

4.1 Home Page 27
4.2 select the mode of input 28
4.3 Upload pdf as input 28
4.4 Uploading Text Page 29
4.5 About Page 29
4.6 Social Connect Page 30
4.7 Accuracy and RMSE 33
4.8 Confusion Matrix 33
4.9 Precision, Recall and F1 Score 34
4.10 Overall Performance Matrix 34

List of Tables

4.1 Confusion Matrix 32

List of Abbreviations

TTS - Text To Speech

Chapter 1

Introduction

1.1 Background

The evolution of audiobook production has been fueled by the desire for immersive storytelling. While traditional methods like basic text-to-speech have laid the groundwork, they often miss the emotional subtleties of the original text. Our project aims to revolutionize this process by integrating advanced machine learning techniques for improved emotion recognition and voice modulation.

In the current landscape of audiobook production, traditional methods like text-to-speech engines often fall short in conveying the emotional richness of written text. This disconnect leads to audiobooks that may feel robotic and fail to evoke the intended emotional response from listeners. Additionally, manual recording processes, where professional narrators read aloud, can be time-consuming and costly. Our project aims to bridge these gaps by integrating advanced machine learning techniques for emotion recognition, revolutionizing audiobook production for a more immersive and authentic experience.

Our project revolutionizes audiobook production by leveraging machine learning algorithms. By dynamically adjusting narration to capture emotional nuances, we create a more immersive experience for listeners. Overall, our project offers high-quality audiobook renditions that rival written text, revolutionizing audiobook production and consumption for a more compelling and immersive storytelling experience.

1.2 Problem Definition

The aim of our project is to address the limitations of traditional audio book production methods by integrating advanced machine learning techniques to accurately capture emotional nuances. The problem lies in the disconnect between the emotional richness of written text and the often flat delivery in regular text-to-speech conversion, the cost and

time required in case of manual method.

1.3 Scope and Motivation

Scope:

Our project encompasses the development of a comprehensive software solution that seamlessly integrates emotion recognition and voice modulation algorithms into the audiobook production process. This involves the implementation of machine learning models trained to detect and interpret emotional nuances present in written text, enabling dynamic adjustments to narration delivery. Additionally, the scope includes the creation of advanced voice modulation capabilities that ensure smooth transitions between different characters,enhancing the overall coherence and realism of the audiobook rendition. Our focus extends beyond mere text-to-speech conversion to encompass a holistic approach to audiobook production, aiming to revolutionize the industry standard.

Motivation:

The motivation behind our project stems from the desire to enhance the audiobook listening experience by addressing the shortcomings of existing production methods. We are driven by the recognition of the disconnect between the emotional depth of written text and the often robotic delivery in TTS, which diminishes the immersive quality of the listening experience. Furthermore, we are motivated by the opportunity to leverage cutting-edge machine learning techniques to bridge this gap and offer users audiobooks that capture the essence and emotion of the original work.Additionally, our motive lies in making audiobooks more accessible and user-friendly, allowing for seamless enjoyment across various platforms and devices.

1.4 Objectives

1. Develop machine learning algorithms for emotion recognition in written text to accurately capture and interpret emotional nuances.

2. Implement text-to-speech (TTS) technology to convert written text into natural-sounding audio narration.

3. Create corresponding voice modulation techniques to dynamically adjust narration based on detected emotions in the text.

4. Incorporate advanced methods such as emotion recognition, text-to-speech (TTS), and voice modulation together into the audiobook production process.

5. Conduct thorough testing and validation to ensure the effectiveness and accuracy of the emotion recognition, TTS, and voice modulation algorithms.

1.5 Challenges

1. Harmonizing multiple algorithms like text-to-speech (TTS), emotion recognition, and voice modulation poses a significant technical hurdle, as each algorithm operates independently and must be seamlessly integrated into the audiobook production workflow.

2. Ensuring that the synthesized voice sounds humane and not robotic is crucial for creating an immersive audiobook experience. Achieving this balance while maintaining the desired emotional depth adds complexity to the project.

3. Manipulating speech parameters to convey appropriate emotions requires meticulous adjustment to avoid a robotic or unnatural sound. Balancing these parameters to evoke the intended emotions while preserving the naturalness of the narration is a key challenge.

1.6 Assumptions

1. The machine learning algorithms for emotion recognition will accurately interpret emotional nuances in written text.

2. The text-to-speech (TTS) technology will effectively convert written text into natural-sounding audio narration.

3. The voice modulation techniques will seamlessly adjust narration based on detected emotions, maintaining coherence and authenticity.

4. The user-friendly interface will facilitate smooth integration of emotion recognition, TTS, and voice modulation capabilities into audiobook production workflows.

1.7 Societal / Industrial Relevance

The application of our advanced text-to-audiobook conversion system can have significant societal and industrial impacts:

Society:

1. Accessibility for Visually Impaired: Enables visually impaired individuals to access written content through immersive audio experiences.
2. Language Learning: Enhances language comprehension and pronunciation practice for students and language learners.
3. Convenient Reading: Offers an alternative reading experience for busy individuals or those preferring auditory learning.

Industry:

1. Publishing and Entertainment: Enables production of high-quality, emotionally engaging audiobooks that stand out in the market.
2. E-Learning and Training: Facilitates the conversion of training materials into engaging audio formats for enhanced learning experiences.
3. Marketing and Customer Engagement: Provides brands with captivating audio content for marketing campaigns and customer engagement.

Relevance and Impact:

1. Efficiency: Automates audiobook creation, saving time and resources for content creators.
2. Inclusivity: Promotes inclusivity and accessibility by making literature accessible to diverse audiences.
3. Innovation: Represents a cutting-edge application of ML in content creation, offering unique features in text-to-audiobook conversion.

In summary, our project leverages ML to transform the audiobook creation process, benefiting both society and industry through enhanced accessibility, engagement, and efficiency.

1.8 Organization of the Report

The report is structured as follows:

1. Introduction

1.1 Purpose

To revolutionize audiobook creation by using Machine Learning to detect emotional nuances in text and dynamically modulate voice, enhancing the listening experience with lifelike, emotive narration.

1.2 Product Scope

Develop a web application that converts written text into high-quality audiobooks by integrating ML-driven emotion recognition and voice modulation, enabling dynamic, expressive narration and seamless character transitions.

1.3 System overview

The system architecture is designed to seamlessly integrate components such as a user interface, backend server, and machine learning models for text processing, emotion recognition, and voice modulation. Through this cohesive design, the application facilitates the conversion of written text into expressive audiobooks with lifelike narration.

2 Product Scope

2.1 Product Perspective

From a product perspective, our audiobook conversion application offers a novel solution that enhances the auditory experience by dynamically adjusting narration based on emotional nuances detected in the text. Its advanced features, including seamless character transitions and customizable voice preferences, cater to diverse user needs, ensuring widespread appeal and adoption in the market.

2.2 Product Functions

Our audiobook conversion application seamlessly processes written text, leveraging machine learning for emotion recognition and voice modulation to deliver immersive narration. Users can customize audio speed. Ans also download the converted audio

2.3 Operating Environment

The application operates within a web-based environment, compatible with major web browsers, ensuring accessibility across different platforms. It utilizes server-side processing for heavy computational tasks, maintaining responsiveness and scalability.

2.4 Assumptions and Dependencies

Assumptions:

1)Text input provided by users is grammatically correct and coherent for accurate emotion recognition and narration.

2)Emotion recognition models perform effectively across various writing styles and genres.

3)Accurate prediction of emotions by the ML model

Dependencies:

1)Reliable internet connection for accessing the web application and server-side processing.

2)Compatibility with web browsers and underlying operating systems for seamless user experience.

3. System Architecture and Design

3.1 System Components

System Components:

User Interface (UI): Provides a platform for users to interact with the application, including options for uploading text, selecting preferences, and initiating the audiobook conversion process.

Backend Server: Responsible for processing user requests, handling data transmission, and orchestrating the interaction between different components of the system.

Text Processing Module: Parses and preprocesses the uploaded text, segmenting it into manageable chunks for further analysis.

Emotion Recognition Engine: Utilizes machine learning models to analyze the emotional content of the text, identifying nuances such as tone, sentiment, and context.

Voice Modulation Engine: Dynamically adjusts the narrator's voice based on the emotional cues detected in the text, enhancing the expressiveness and realism of the audiobook narration.

Text-to-Speech (TTS) Engine: Converts the processed text into speech, incorporating emotional modulation and character transitions to produce high-quality audiobook out-

put.

3.2 Architectural Design

The architectural design emphasizes modularity and scalability, facilitating seamless integration and future expansion of new features. It prioritizes efficiency and performance to ensure real-time processing of text and audio data, enhancing the overall user experience.

3.3 Dataset

The dataset, comprising 16,001 lines sourced from Kaggle, contains comments from Twitter, each tagged with a score indicating one of six emotions: joy, sadness, fear, surprise, anger, or love.

4. Functional and Non-Functional Requirements (SRS)

4.1 Functional Requirements

Functional requirements for the project include accurate classification of text comments into six emotions (joy, sadness, fear, surprise, anger, love), seamless integration of emotion recognition into the audiobook conversion process, dynamic voice modulation based on detected emotions, customizable voice preferences, and efficient processing of text data to ensure real-time narration.

4.2 Non-Functional Requirements

4.2.1 Performance Requirements

1.Performance: The system must process text data and perform emotion recognition with minimal latency to provide a responsive user experience.

2.Scalability: It should be able to handle increasing loads of user requests and data volumes without compromising performance.

3.Accuracy: Emotion recognition algorithms must achieve high accuracy in classifying text comments to ensure the quality of narration modulation.

4.2.2 Software Quality Attributes

1)Reliability: The software should consistently perform its intended functions accurately and predictably under various conditions, minimizing the likelihood of failures or errors.

2)Usability: The software should be intuitive and user-friendly, with interfaces designed to be easily navigable and understandable by users of varying levels of technical expertise, enhancing user satisfaction and productivity.

3)Portability: The software should be capable of running on various platforms and environments with minimal modifications, ensuring compatibility and ease of deployment

across different systems.

5.Implementation Details

5.1 User Interface Design

he user interface should be intuitive, visually appealing, and accessible across different devices and screen sizes, focusing on simplicity and ease of navigation to enhance user engagement.

5.2 Implementation Strategies

1)Integration of Python Libraries: Utilize Python libraries such as pyttsx3, pydub, numpy, and scikit-learn for text-to-speech conversion, audio processing, data manipulation, and machine learning tasks, respectively. Ensure seamless integration and compatibility between these libraries to streamline development.

2)Model Development with Logistic Regression: Implement the emotion recognition model using logistic regression from the scikit-learn library, leveraging its simplicity and efficiency for binary classification tasks. Train the model on the provided dataset to accurately classify text comments into six emotional categories.

3)Web Development with HTML, CSS, and JavaScript: Design and develop the user interface using HTML for structure, CSS for styling, and JavaScript for interactive elements. Focus on creating a visually appealing and intuitive interface that facilitates easy interaction with the application.

4)Backend Implementation: Develop the backend functionality using Python, integrating the trained logistic regression model for emotion recognition. Utilize Flask for handling user requests, data processing, and communication with the frontend.

5.3 Module Division

The system can be divided into modules such as text processing, emotion recognition, voice modulation, and user interface, each responsible for specific functionalities, promoting better organization and collaboration among development teams.

5.4 Work Schedule (Gantt Chart)

Visual timeline outlining project milestones and tasks represented using Gantt Chart.

Chapter 2

Software Requirements Specification

2.1 Introduction

2.1.1 Purpose

The purpose of the Emotion-Aware Audiobook Conversion project is to contribute to the audiobook industry by leveraging cutting-edge machine learning technologies to enhance the emotional depth and engagement of audiobook narration. The project will redefine the standards of audiobook production and consumption, offering listeners a transformative and emotionally resonant storytelling experience.

2.1.2 Product scope

The website being specified is for converting written books into audiobooks using emotion-aware artificial intelligence and machine learning techniques. Its purpose is to enhance the audiobook experience by accurately conveying the emotional nuances present in the text, thereby creating a more engaging and immersive listening experience for users. The benefits ,objectives and goals of the software include:

Improved Accessibility: By audiobook versions of written books, the software improves accessibility for individuals with visual impairments or other disabilities that may hinder their ability to read printed text.

Enhanced Emotional Engagement: by analyzing the emotional content of the text and adjusting the prosody of the synthesized speech accordingly, the website aims to create audiobooks that evoke a deeper emotional response from listeners.

Corporate Goals Alignment: creating this website fits with the company's goals of being innovative and using the latest technology to make customers satisfied. By using tech like emotion-aware machine learning, the company shows it's keeping up with what's new and making sure customers get what they want.

2.2 Overall Description

2.2.1 Product Perspective

This is envisioned as a standalone website product that serves as a pioneer in the field of text-to-audio conversion with voice modulation and emotional rendering. As a new, self-contained product, It fills a gap in the market for audio content creation tools that go beyond traditional text-to-speech applications. It is not a replacement for existing systems but rather a solution that enhances the accessibility and engagement of written content through immersive audio experiences.

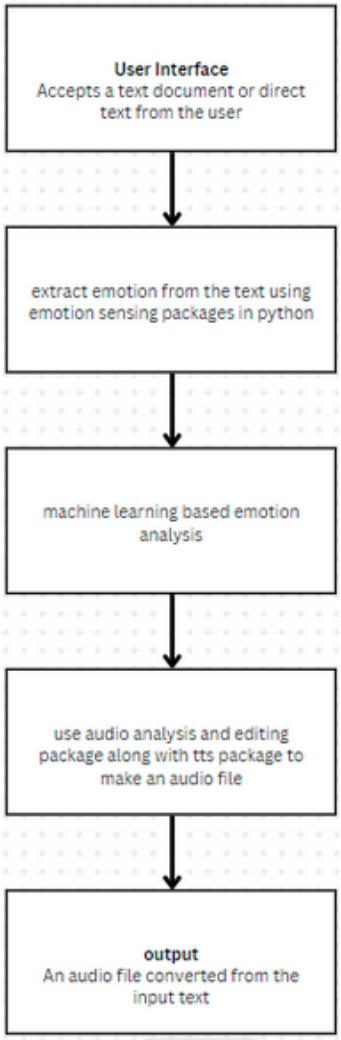


Figure 2.1: Flowchart

2.2.2 Product Function

1. Text-to-Audio Conversion:

   It converts written text into high-quality audio recordings. Users can input text as text from various sources such as PDFs, Word documents, eBooks, and web articles.The system processes the text to generate clear and natural-sounding speech output.

2. Voice Modulation:

   Voice modulation necessary for the given text is done for an immersive experience. Users can adjust parameter speed to create personalized narration styles. This feature adds character and dynamism to the narration, enhancing engagement and expressiveness.

3. Emotional Rendering:

   This incorporates emotional intelligence into the audio content by infusing appropriate emotions into the narration.Using sentiment analysis algorithms, the system identifies the emotional tone of the text and adjusts the delivery to convey feelings such as excitement, empathy, or suspense.

2.2.3 Operating Environment

The operating environment for the website includes CSS and HTML for the front end, and Flask and Python for the backend, involving a combination of web servers, development tools, and deployment platforms to create, test, deploy, and maintain the application. And this may be provided to work in most versions of windows.

2.2.4 Design and Implementation Constraints

The main constraints include compatibility limitations across different operating systems, web browsers, and devices, resource constraints such as memory, processing power, security considerations for data protection and privacy, technical constraints related to the capabilities of underlying technologies and frameworks, usability challenges in designing an intuitive interface and effective user experience, and regulatory/legal constraints regarding copyright, intellectual property rights, and compliance with relevant laws and regulations.

2.2.5 Assumptions and Dependencies

Assumptions and dependencies include relying on stable third-party components for text and sentiment analysis, along with voice synthesis services, with assumptions on their reliability. Additionally, assumptions are made about the availability of a suitable development environment and a stable operating environment with sufficient resources. Dependencies also extend to user interaction patterns, regulatory compliance, and integration with external systems, with reliance on user feedback, legal requirements, and compatibility with external platforms. Documenting these factors is crucial for risk management and project alignment, requiring regular monitoring and communication to address changes effectively during development and deployment.

2.3 External Interface Requirements

2.3.1 User Interfaces Software Components:

1. **Text-to-Audio Converter:** Converts text input into audio output.
2. **User Interface Module:**Facilitates interaction between users and the text-to-audio converter.

2.3.2 User Interface Design:

1. **Graphical User Interface (GUI):**
    - The GUI follows a simple and intuitive design to facilitate easy interaction for users.
    - Screen Layout: The main screen consists of a text input area where users can input the text they want to convert to audio.
    - Standard Buttons
    - Convert: Initiates the conversion process.
    - Play: Plays the converted audio.
    - Clear Text: Clears the text input area.
    - Clear Audio: Clears generated audio file

2. **Styling and Visual Design:**

   - The GUI adheres to a modern and visually appealing design.
   - Font Sizes and Styles: Text input and output areas utilize readable fonts and appropriate font sizes for optimal user experience.
   - Color Scheme: The color scheme is chosen to ensure readability and minimize eye strain.
   - Responsive Design: The GUI is designed to be responsive across different screen sizes and resolutions, ensuring consistent user experience on various devices.

3. **Hardware**

   - Processor : Intel i3 and above
   - RAM : 4 MB and above
   - Hard disk : 100 GB and above
   - Monitor : LCD monitor
   - Keyboard : Normal or Multimedia
   - Mouse : Compatible mouse

2.3.3 Software Interfaces

The software interfaces with various components, including pyttsx (or pyttsx3) for text-to-speech conversion, numpy and pydub for audio processing, an emotion sensing package scikit-learn (sklearn) for machine learning tasks. pyttsx converts text input into synthesized speech, while numpy and pydub provides tools for audio analysis and modification. The emotion sensing package analyzes text or speech input to detect emotional content, leveraging scikit-learn for machine learning-based emotion classification if needed. Integration with external services, if necessary, involves communication via Restful APIs or other protocols. The operating system and development tools provide the environment for software execution and development. Data sharing occurs between components, necessitating proper formatting and compatibility to ensure seamless communication and integration with scikit-learn for machine learning tasks.

2.3.4 Communications Interfaces

The communication interface for this involves a user-friendly web-based interface accessible through standard web browsers. Users interact with the system by inputting text and adjusting the speed of the audio. This communicates with external services, such as text analysis APIs and voice synthesis engines, to process user inputs and generate audio output. The system provides real-time feedback and status updates to users during the text-to-audio conversion process, ensuring a seamless and responsive user experience.

2.4 System Features

1. **Text input and Upload**

   This feature allows the users of the website to upload any piece of text or pdf they would like to convert into audio. Giving an input text/pdf and uploading it is the initial step of this process. This feature allows the user to select any text/pdf of their choice.

2. **Text to Audio conversion**

   The core functionality of the "Advancing Book to Audiobook Conversion through Emotion Aware Machine Learning" project is speech synthesis. This feature enables the system to convert input text into natural-sounding speech. The system utilizes a text-to-speech (TTS) model that generates speech waveform based on the input text and selected emotion.

2.5 Other Nonfunctional Requirements

1. **Performance Requirements**

   Performance requirements for this encompass swift response times to user interactions, rapid text-to-audio conversion speeds, scalability to accommodate fluctuating user loads, and support for concurrent user access without compromising performance. The system must maintain high audio quality, efficient resource utilization, and reliability to ensure uninterrupted availability. Thorough load testing is essential to assess performance under peak usage conditions and identify potential

bottlenecks. By meeting these requirements, this can deliver a seamless and satisfying user experience, facilitating efficient conversion of text to emotionally rich audio content.

2. **Safety Requirements**

   Safety requirements encompass ensuring data privacy and confidentiality, implementing content filtering mechanisms to prevent inappropriate content generation, providing clear user guidelines for safe usage, and ensuring accessibility for users with disabilities. The system must also be reliable and stable to prevent safety hazards or disruptions, comply with relevant safety regulations and standards, and offer guidance for users in emergency situations. By meeting these requirements, This can provide users with a secure and trustworthy platform for creating and consuming audio content, promoting a safe and positive user experience while protecting user privacy and well-being.

3. **Security Requirements**

   Security requirements for this software include encryption of data in transit. Secure session management and regular software updates are essential for thwarting session hijacking and addressing vulnerabilities.

4. **Software Quality Attributes**

   (a) Reliability:

   Reliability refers to the system's ability to consistently perform its functions accurately and predictably under various conditions. reliability ensures that the text-to-audio conversion process produces consistent and high-quality audio output, maintaining fidelity to the original text and preserving the intended emotional expression. Users rely on this to deliver dependable results without unexpected errors or interruptions.

   (b) Correctness:

   Through rigorous testing and validation procedures, your software maintains accuracy and reliability in text-to-speech conversion, audio processing, and

emotion analysis tasks, ensuring high accuracy rates and low false positive/negative rates

(c) Usability:

Usability relates to how easy and intuitive it is for users to interact with the system effectively and efficiently. It should feature a user-friendly interface that allows users to input text, adjust voice modulation settings, and navigate the application with ease. Clear instructions, intuitive controls, and helpful feedback mechanisms contribute to a positive user experience and encourage user adoption and satisfaction.

Chapter 3

System Architecture and Design

3.1 System Overview

The Text-to-Speech (TTS) with emotion-integrated audio output system is designed to convert text input into spoken audio with varying emotional expressions. This system enhances the naturalness and expressiveness of synthesized speech by incorporating emotional cues into the audio output.

The system consists of the following main components:

- **Text Processor**: Receives textual input and preprocesses it for synthesis. It may include tokenization, normalization, and linguistic analysis to prepare the text for the TTS engine.
- **Emotion Classifier**: Analyzes the text to determine the underlying emotional content. This component may utilize machine learning models trained on emotion-labeled datasets to predict emotions such as happiness, sadness, anger, etc.
- **TTS Engine**: Converts preprocessed text input into spoken audio. It synthesizes speech with appropriate prosody, intonation, and pacing to convey the emotional nuances identified by the emotion classifier.
- **Audio Renderer**: Takes the synthesized speech from the TTS engine and generates the final audio output. It may include audio processing techniques to enhance the emotional expression, such as pitch modulation, amplitude variation, etc.

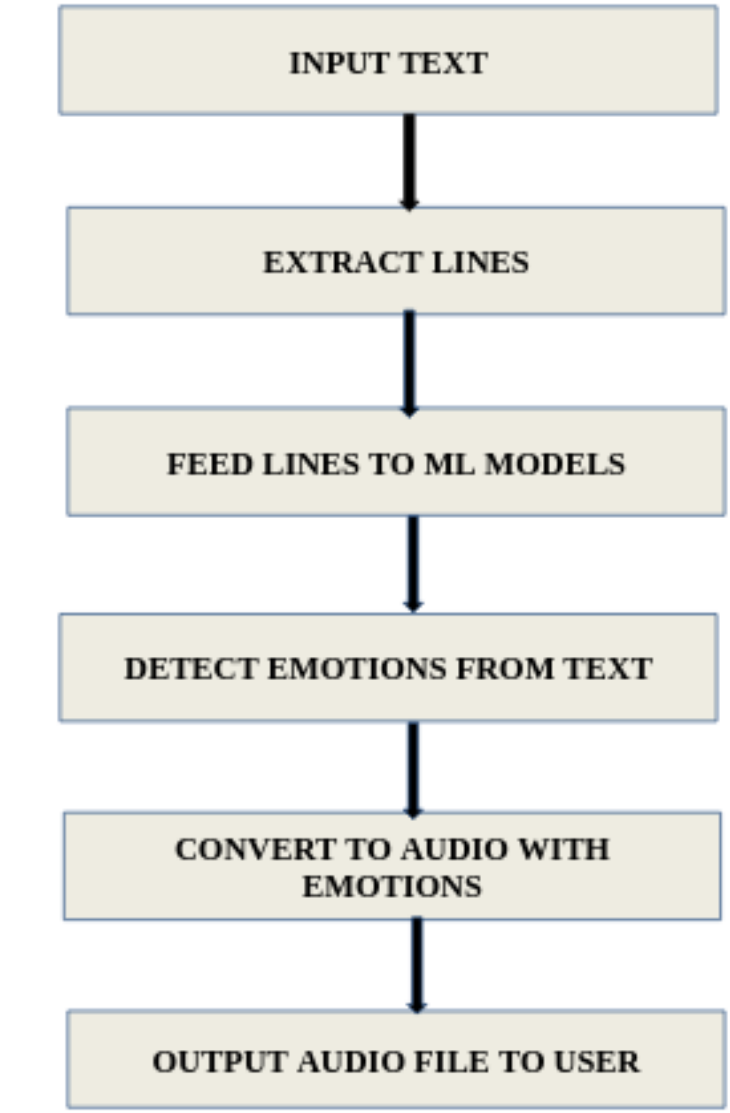


Figure 3.1: ARCHITECTURAL DIAGRAM

3.2 Architectural Design

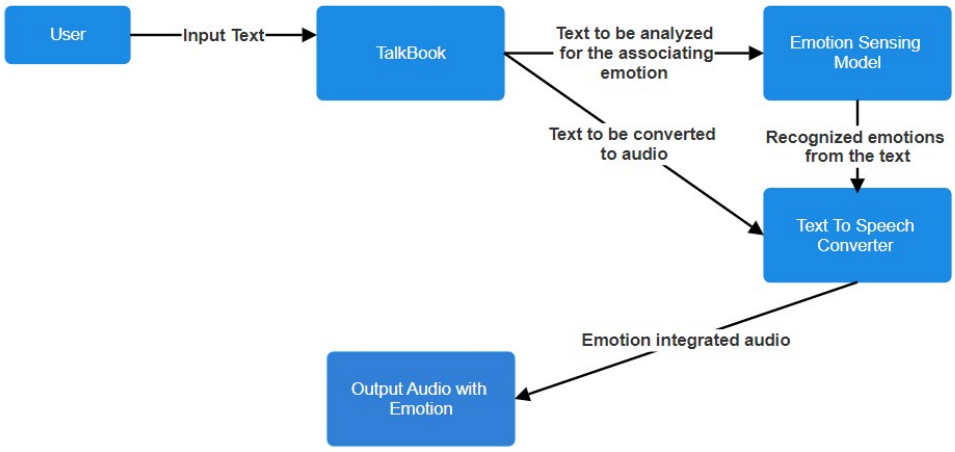


Figure 3.2: ARCHITECTURAL DESIGN

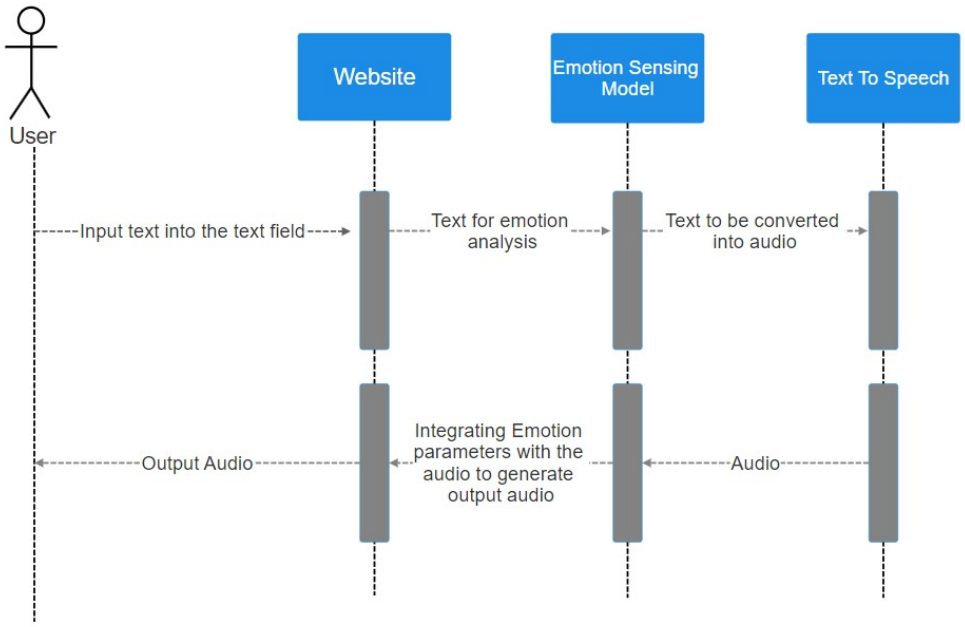


Figure 3.3: SEQUENCE DIAGRAM

3.3 Dataset identified

Each text sample in the dataset is typically labeled with one emotion category(0,1,2,3,4 and 5). Datasets may be compiled from various sources, including social media posts, product reviews, movie reviews, news articles, and personal diaries. The diversity of text sources helps ensure that the model can generalize well to different domains and writing styles. Emotions are usually categorized into a predefined set of classes, such as joy, sadness, anger, love, fear and surprise. The number and granularity of emotion categories can vary depending on the dataset's purpose.

	A	B	C
1	text	label	
2	i didnt feel	0	
3	i can go fr	0	
4	im grabbin	3	
5	i am ever f	2	
6	i am feelin	3	
7	ive been fe	0	
8	ive been ta	5	
9	i feel as cc	4	
10	i have bee	1	
11	i feel roma	2	
12	i feel like i	0	
13	i do feel th	1	
14	i think it s t	3	
15	i feel low e	0	
16	i have imm	1	
17	i do not fe	1	
18	i didnt real	0	
19	i feel prett	0	
20	i started fe	0	
21	i now feel	4	
22	i feel irrita	3	
23	i am feelin	4	
24	i have the	1	
25	i was able	1	
26	i already fe	3	
27	i still love	0	
28	i feel so inl	0	
29	i become	0	
30	i feel kinda	3	
31	i feel more	1	

Figure 3.4: SAMPLE DATASET

3.4 Proposed Methodology/Algorithms

1. Input text/Upload pdf into the web application.
2. Extraction of lines from the text for conversion.
3. Providing the extracted lines to pre trained ML models for conversion.
4. Detect emotions from the lines.
5. Convert the lines to audio files with respective emotions and modulations.
6. Output the audio file to the user.

3.5 User Interface Design

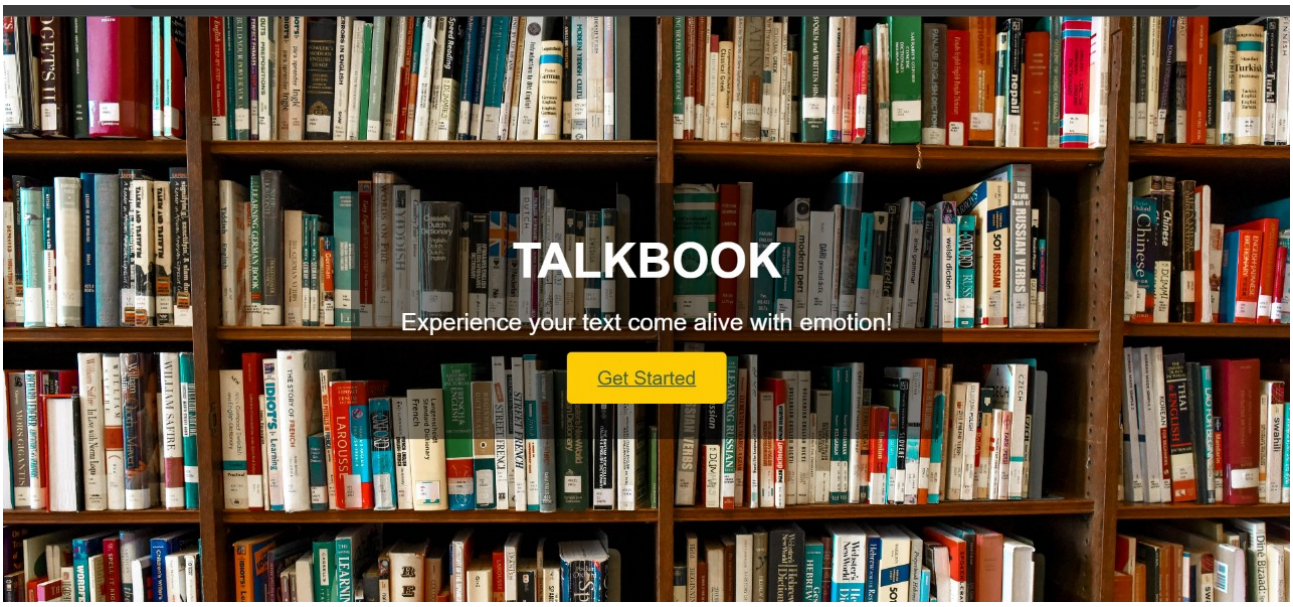


Figure 3.5: Home Page

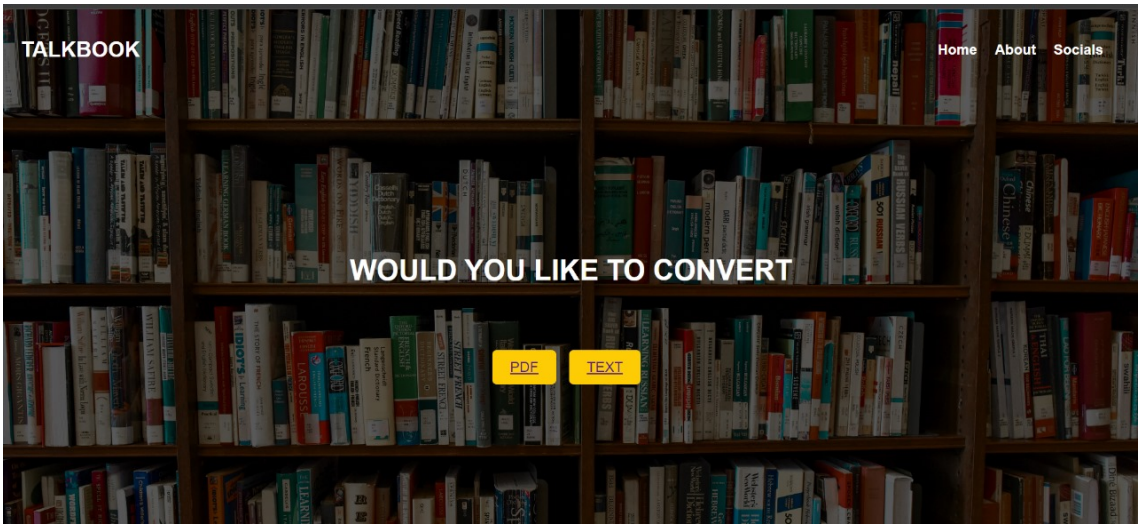


Figure 3.6: select the mode of input

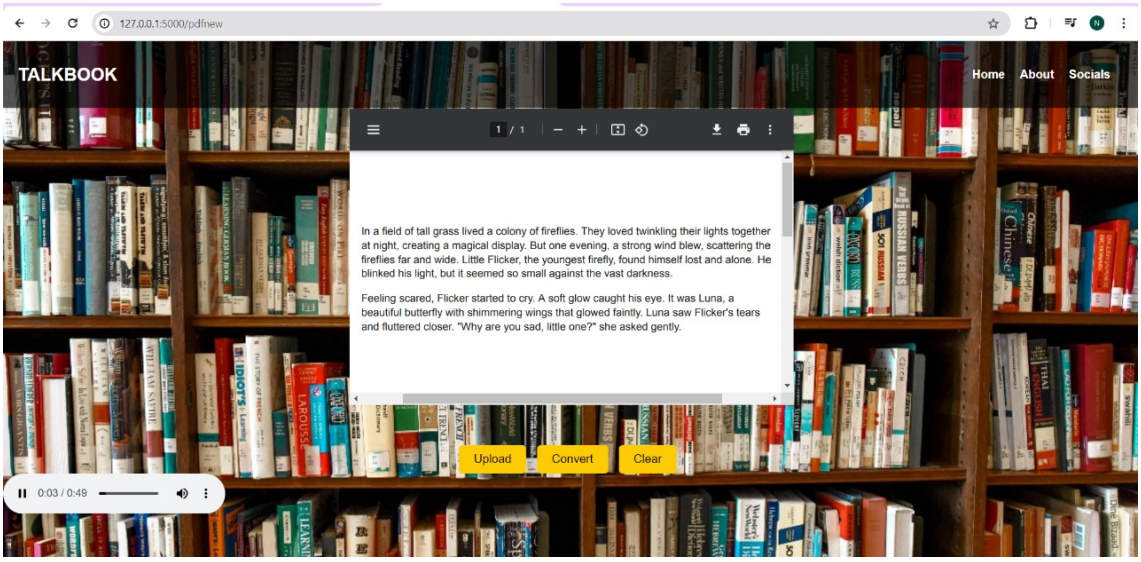


Figure 3.7: Upload pdf as input

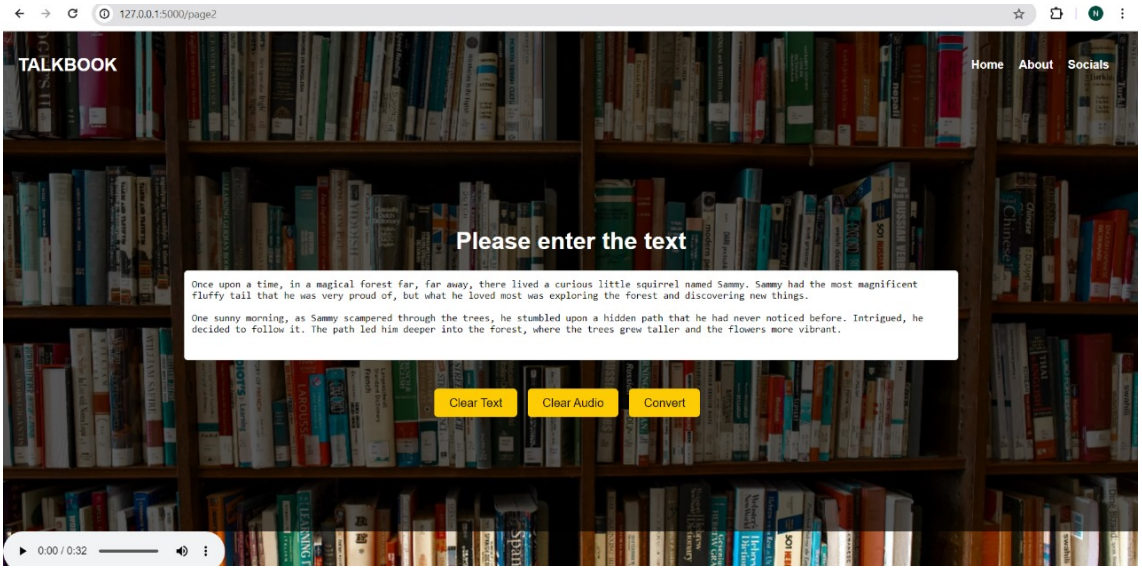


Figure 3.8: Uploading Text Page

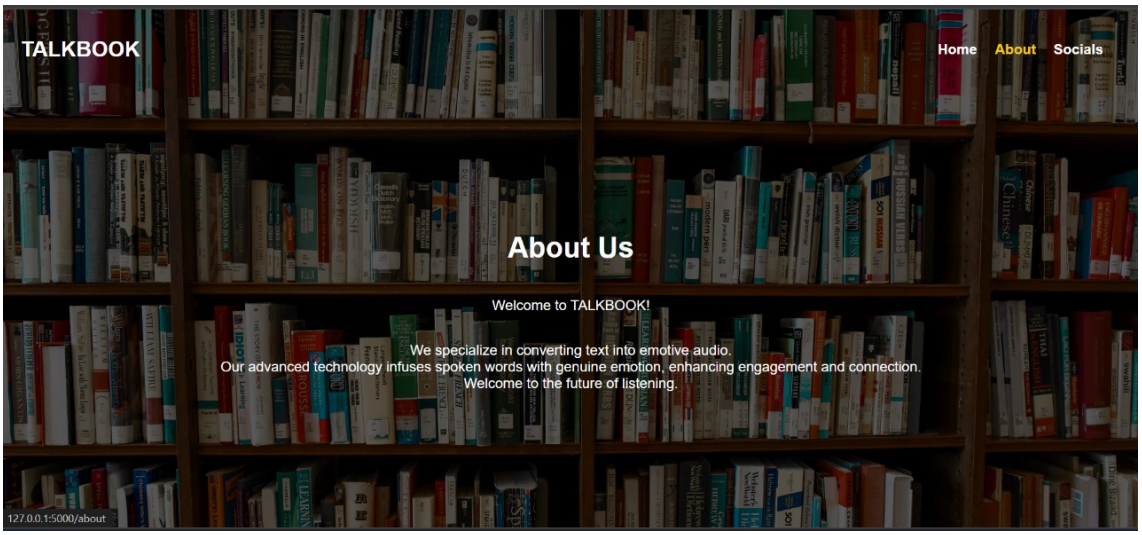


Figure 3.9: About Page

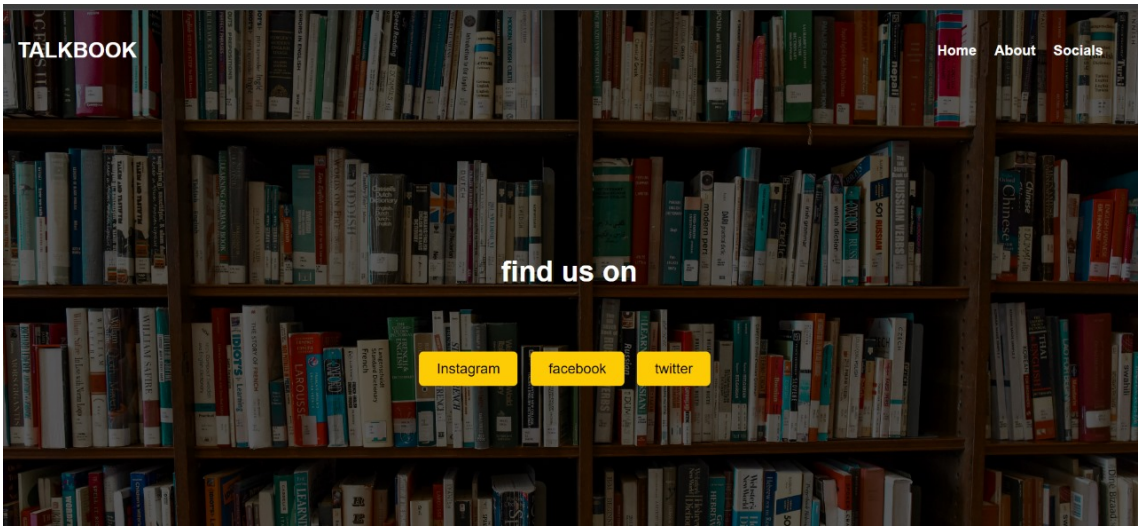


Figure 3.10: Social Connect Page

3.6 Description of Implementation Strategies

The project is designed with three distinct components that work together seamlessly to achieve the desired functionality.

1. **Website Development (HTML/CSS/JavaScript)**:
   The user interface (UI) is crafted using HTML, CSS, and JavaScript technologies. Users interact with this website to input text/ Upload pdf, initiating the emotion detection and audio output processes. The UI is intuitive, responsive, and aesthetically pleasing, ensuring a smooth user experience.

2. **Machine Learning Model Creation (Python - Logistic Regression)**:
   A machine learning (ML) model is developed entirely from scratch to accurately detect emotions within textual inputs. Logistic regression is chosen as the ML algorithm for its effectiveness in binary classification tasks, which suits the emotion detection requirements. The model is trained rigorously using a dataset comprising 16001 rows, ensuring robustness and accuracy in emotion identification.

3. **Audio Generation and Emotion Mapping (Python Libraries - pyttsx3, pydub)**:
   Once the ML model detects the emotion within the input text, the system triggers

audio generation with the corresponding emotional tone. Python libraries such as pyttsx3 are utilized for text-to-speech conversion, producing natural-sounding audio outputs. Emotion mapping is achieved through careful integration of the ML model's output with pyttsx3, ensuring that the audio reflects the detected emotion accurately. Additional audio processing tasks are handled using pydub for quality enhancement and format compatibility.

4. **Backend Integration with Flask and Jupyter Notebook**:
   The backend logic, including the ML model and audio processing functionalities, is integrated into a Jupyter Notebook environment for streamlined development and testing. Flask is employed to connect the backend codebase with the frontend website, facilitating seamless data flow and interaction between the UI and the ML-powered features. This integration ensures that the entire system operates efficiently, handling user inputs, emotion detection, audio generation, and UI updates in a cohesive manner.

Overall, the project combines cutting-edge technologies in web development, machine learning, and audio processing to deliver a sophisticated yet user friendly platform for emotion-based text-to-speech conversion.

3.7 Module Division

The project is structured into three primary modules, each serving distinct functions within the system:

1. **Web Application**: Integral to the user experience is a web -based interface, accessible via a website. Users engage with this platform by uploading their desired text. This module provides a user-friendly and intuitive means for individuals to interact with the system.

2. **Emotion Sensing Model**:The Text-to-Speech (TTS) system with integrated emotion sensing, leverages a sophisticated Machine Learning model designed to detect and interpret emotional cues within text input.

3. **Text To Speech**:Text-to-Speech (TTS) is a transformative technology that converts written text into spoken language, mimicking the natural cadence, intonation,

and pronunciation of human speech along with the found emotions.

3.8 Work Schedule - Gantt Chart

S.NO.	Task name	Start	Finish
1	Research and requirement analysis	March 2,2024	March 7,2024
2	Requirement Specification	February 3,2024	February 7,2024
3	Design Layout	March 8,2024	March 15,2024
4	Front-end Development	March 25,2024	April 12,2024
5	Emotion Recognition Implementation	March 25,2024	April 10,2024
6	Obtaining audio modulation parameters	April 13,2024	April 20,2024
7	Integrating TTS,ML and audion generation modules	April 27,2024	May 1,2024
8	Integrating Frontend and Backend	May 2,2024	May 4,2024
9	Testing and Deployment	May 5,2024	May 6,2024

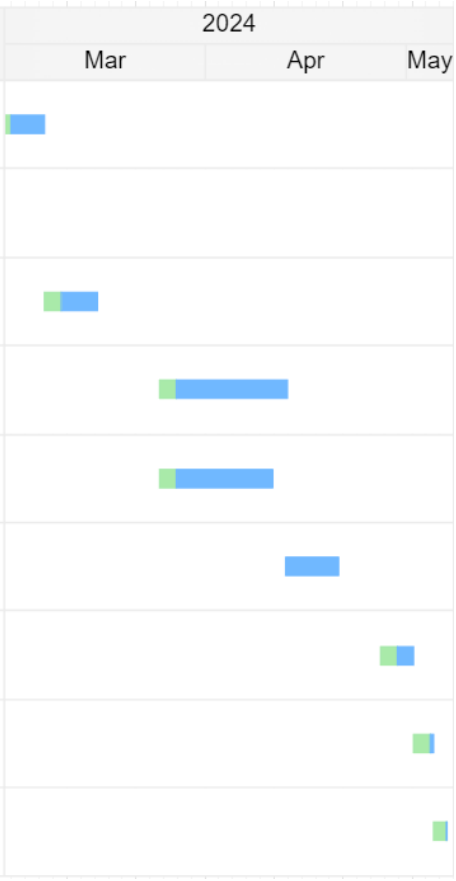


Figure 3.11: GANTT CHART

Module 1: 2 group member, Module 2: 4 group members, Module 3: 2 group member

Chapter 4

Results and Discussions

4.1 Overview

The project successfully converted text and PDF documents into audio files while integrating emotion through the use of a logistic regression model for emotion recognition. The end results demonstrated clear and intelligible audio output, with distinct emotional tones corresponding to the detected emotions in the text. Quantitatively, the model achieved an accuracy of 88 per in emotion recognition, which was validated against a labeled dataset. Further analysis revealed that the system performed best with texts containing explicit emotional cues, while more ambiguous texts posed a challenge for emotion classification. Overall, the project achieved its primary goal of producing emotionally expressive audio from textual content, paving the way for more nuanced and engaging auditory experiences.

4.2 Testing

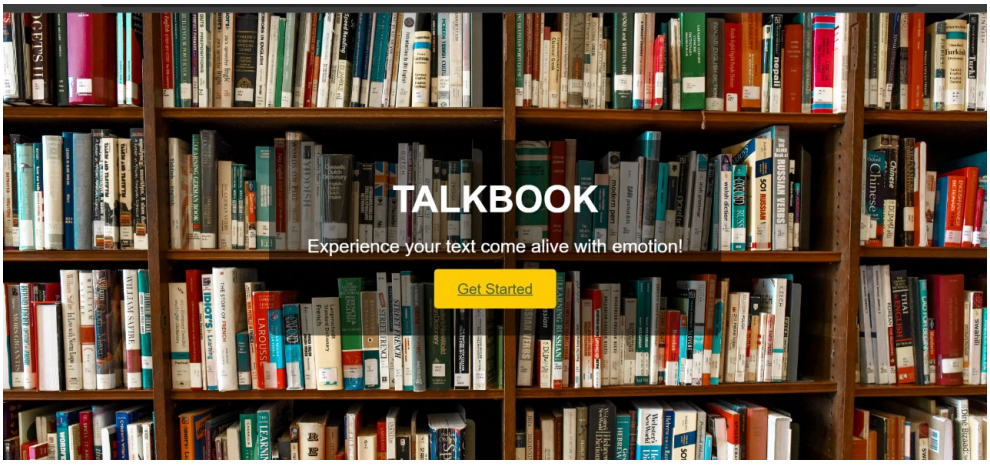


Figure 4.1: Home Page

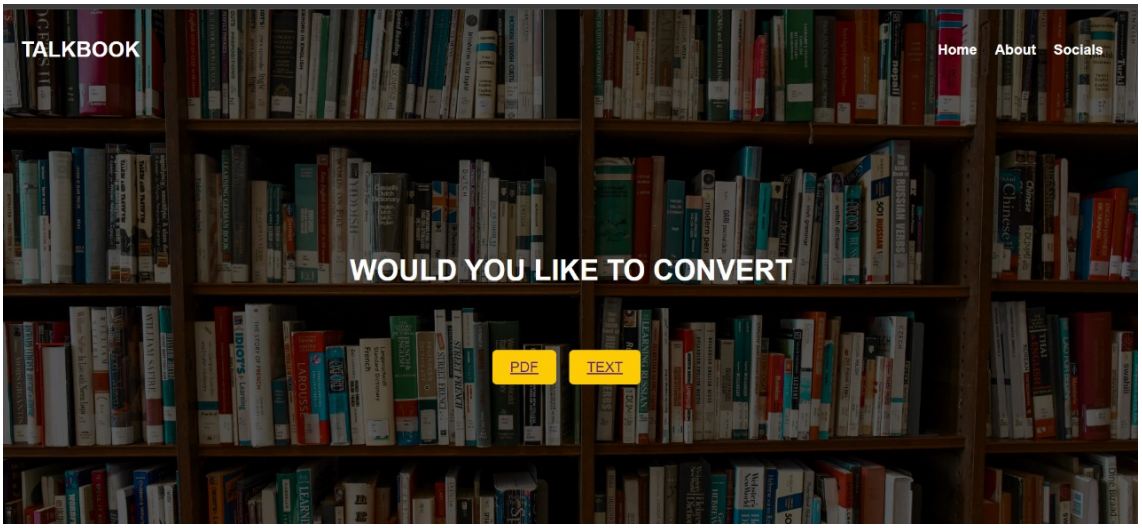


Figure 4.2: select the mode of input

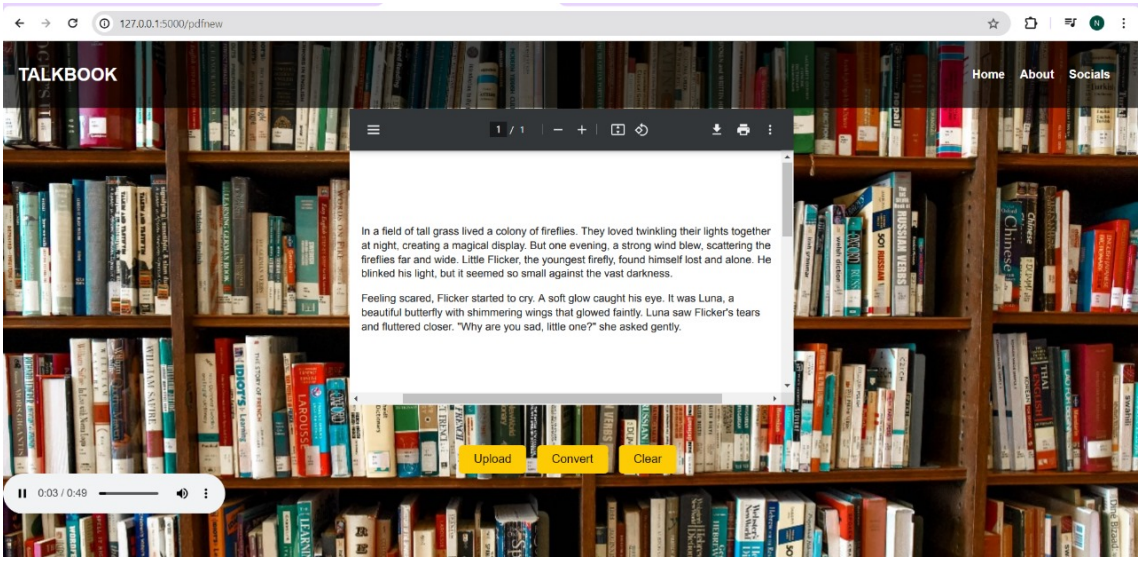


Figure 4.3: Upload pdf as input

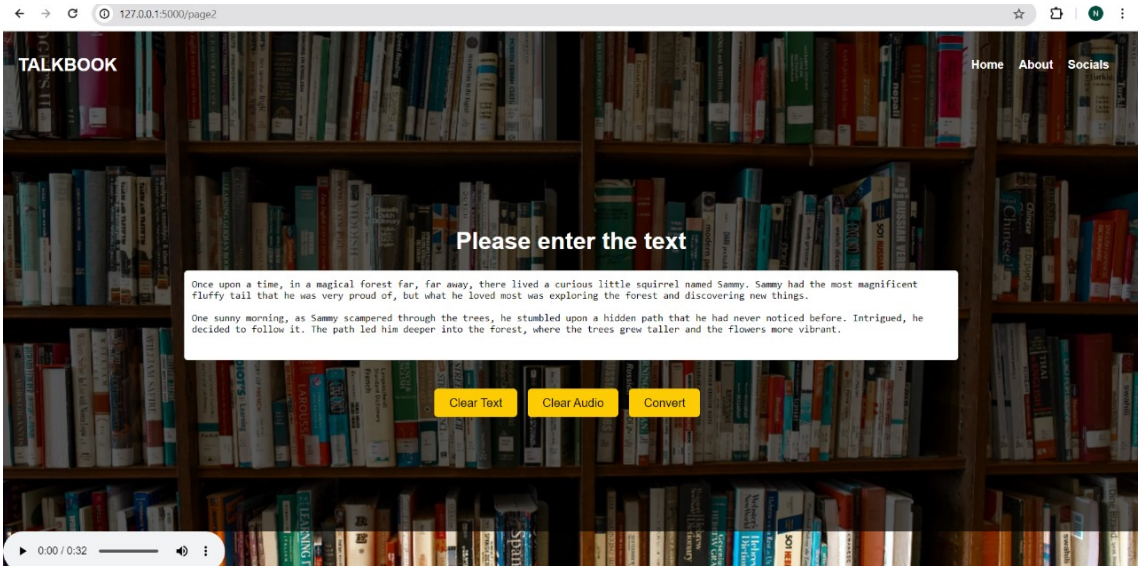


Figure 4.4: Uploading Text Page

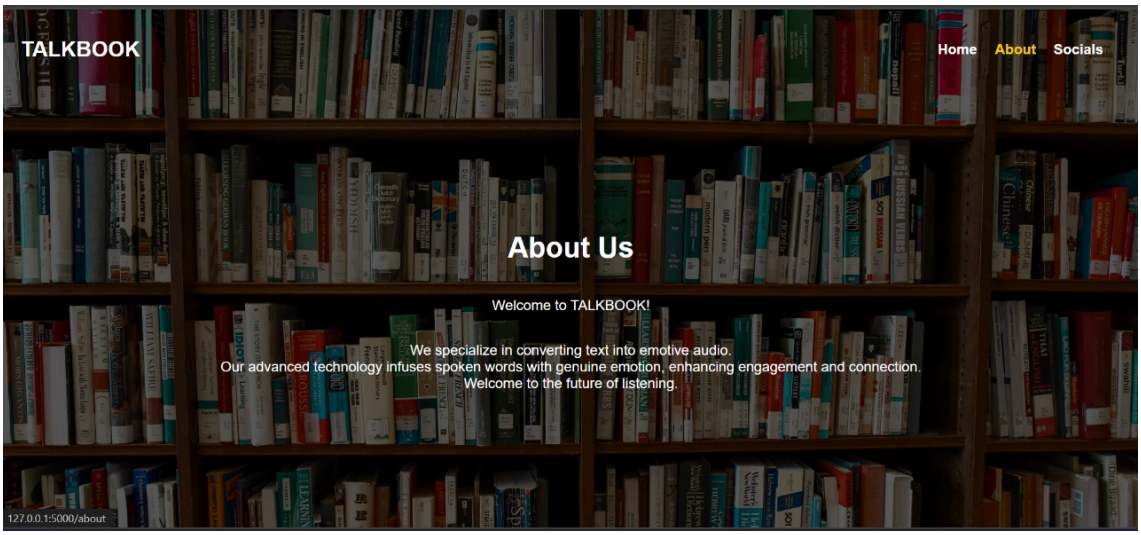


Figure 4.5: About Page

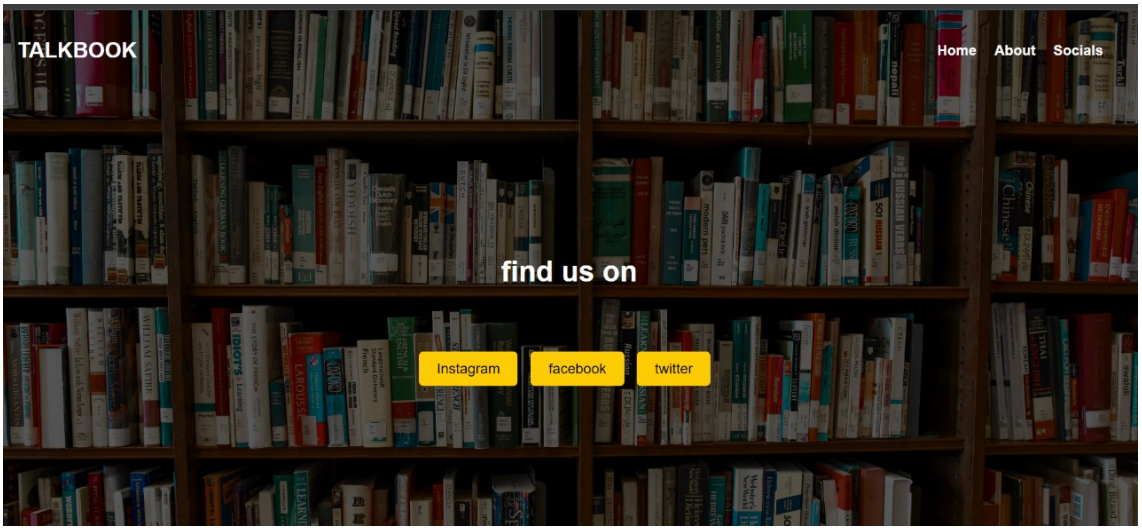


Figure 4.6: Social Connect Page

4.3 Quantitative Results

In this section, we present the quantitative results of our project, which involved developing an emotion-aware logistic regression machine learning model to convert books into audiobooks. The evaluation metrics include accuracy, precision, recall, F1-score, RMSE, and the confusion matrix. Each metric is supplemented with a textual description for clarity.

Accuracy

Accuracy measures the proportion of correctly predicted instances among the total instances. In our model, accuracy was calculated as follows:

Accuracy: 88 percentage

This indicates that our model correctly predicted the emotion of the text 88 percentage of the time.

Precision

Precision is the ratio of true positive predictions to the sum of true positive and false positive predictions. It shows how many of the predicted positive instances were actually positive.

Precision: 88 percentage

A precision of 88 percentage means that 88 percentage of the instances our model

predicted as positive (specific emotion) were actually positive.

Recall

Recall, or sensitivity, is the ratio of true positive predictions to the sum of true positive and false negative predictions. It measures how well our model identifies positive instances.

Recall: 88 percentage

A recall of 88 percentage indicates that our model correctly identified 88 percentage of all actual positive instances.

F1-Score The F1-score is the harmonic mean of precision and recall. It provides a single metric that balances both precision and recall.

F1-Score: 88 percentage

An F1-score of 88 percentage suggests a good balance between precision and recall in our model's performance.

Root Mean Square Error (RMSE)

RMSE measures the average magnitude of errors between predicted and actual values. It provides a measure of the differences between predicted and observed values.

RMSE: 0.82

An RMSE of 0.82 indicates that the average error in our model's predictions is relatively low.

Confusion Matrix

- Sadness (0): The model correctly predicted 1327 instances of sadness. There were misclassifications, such as 43 instances misclassified as Joy.
- Joy (1): The model correctly predicted 1492 instances of joy, with some misclassifications like 47 instances predicted as Love.
- Love (2): The model correctly predicted 319 instances of love. There were significant misclassifications, such as 100 instances predicted as Joy.

- Anger (3): The model correctly predicted 510 instances of anger. Misclassifications include 47 instances predicted as Sadness.
- Fear (4): The model correctly predicted 457 instances of fear, with misclassifications such as 32 instances predicted as Sadness.
- Surprise (5): The model correctly predicted 117 instances of surprise, with 16 instances misclassified as Joy.

Actual/Predicted	**Sadness(0)**	**Joy(1)**	**Love(2)**	**Anger(3)**	**Fear(4)**	**Surprise(5)**
Sadness(0)	1327	43	6	36	22	0
Joy(1)	21	1492	47	8	7	6
Love(2)	6	100	319	4	4	1
Anger(3)	47	27	1	510	18	4
Fear(4)	32	32	4	27	457	14
Surprise(5)	13	16	3	1	28	117

Table 4.1: Confusion Matrix

The confusion matrix provides a detailed performance breakdown for each emotion class, indicating true positives, false positives, true negatives, and false negatives. This detailed breakdown helps in understanding the areas where the model performs well and where improvements can be made.By presenting the results in this structured manner, we ensure clarity and provide a comprehensive understanding of our model's performance in identifying the emotion present within each line of the input text.

4.4 Graphical Analysis

In this section, we present the graphical analysis of our project, which involved developing an emotion-aware logistic regression machine learning model to identify the emotion present within each line of the input text. Various graphs are used to visualize the results, and a textual description accompanies each graph to provide clear insights.

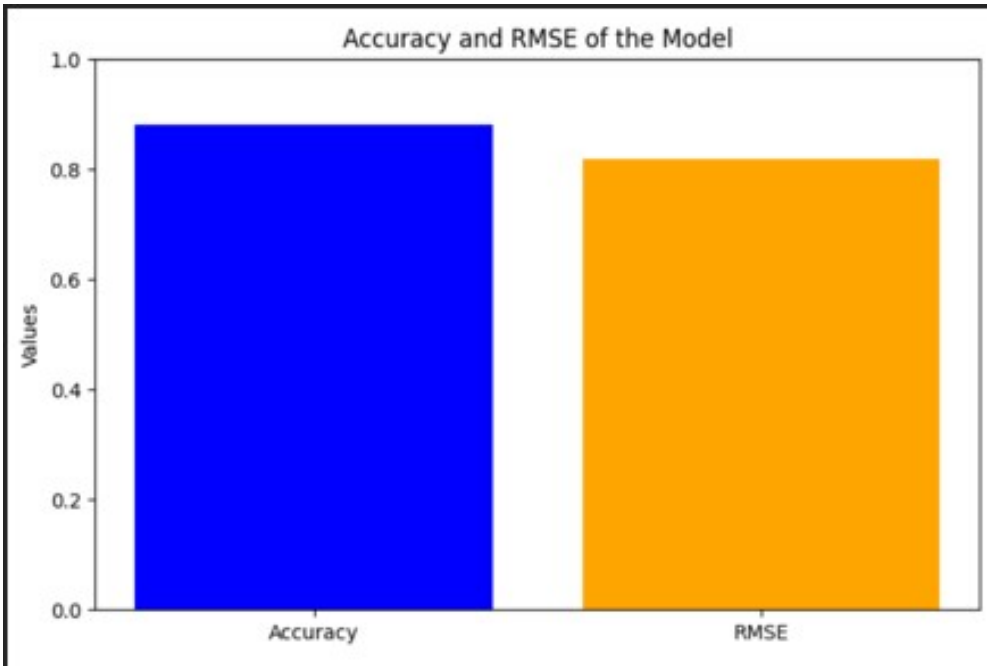


Figure 4.7: Accuracy and RMSE

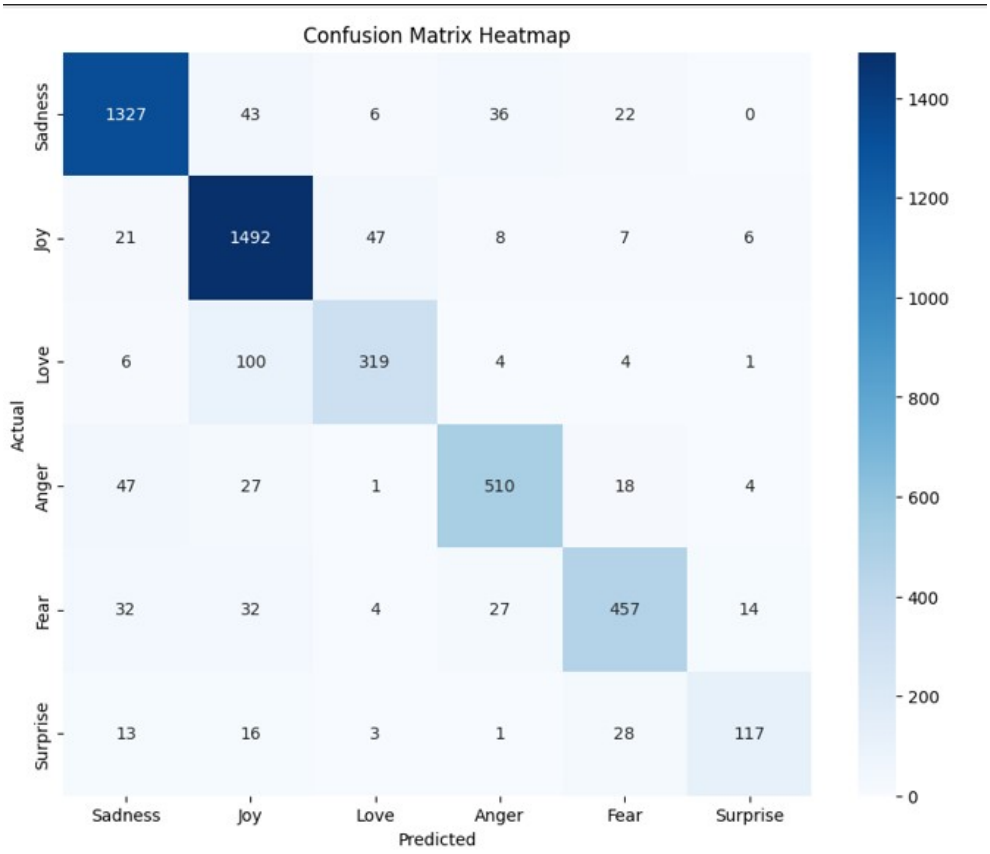


Figure 4.8: Confusion Matrix

Figure 4.9: Precision, Recall and F1 Score

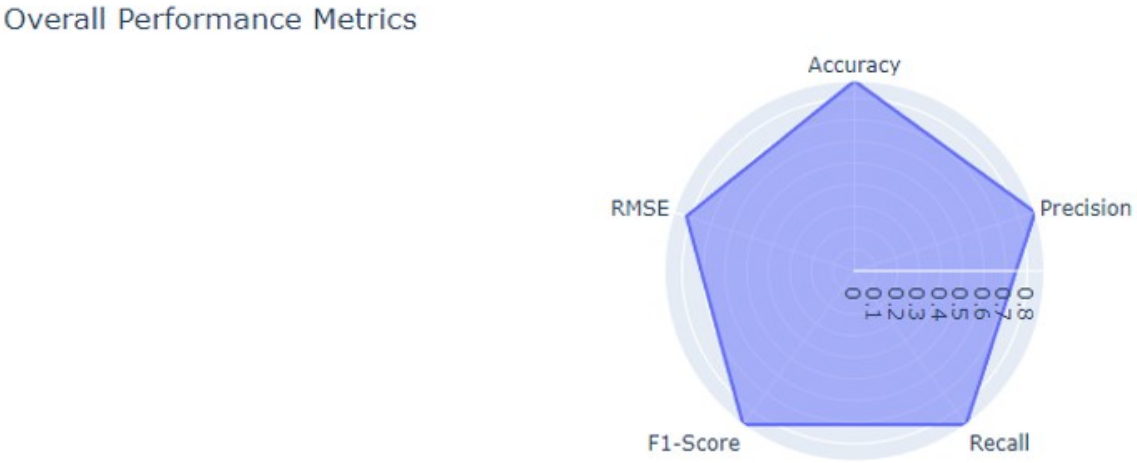


Figure 4.10: Overall Performance Matrix

4.5 Discussion

4.5.1 Summary of Results

In our project aimed at revolutionizing the conversion of written text into audiobooks using Machine Learning (ML) techniques, we achieved noteworthy results across several performance metrics. Specifically, the metrics evaluated were Accuracy, Precision, Recall, F1-Score, and Root Mean Square Error (RMSE). The results for these metrics are as follows:

1. Accuracy: 0.88
2. Precision: 0.88
3. Recall: 0.88
4. F1-Score: 0.88
5. 5 RMSE: 0.82

These metrics collectively indicate a high level of performance, particularly in terms of classification accuracy, precision, recall, and the harmonic mean of precision and recall (F1-Score). The RMSE value, which is lower compared to the other metrics (scaled differently), suggests a relatively low level of error in our predictions.

4.5.2 Reasoning Behind Results

The integration of sophisticated emotion recognition algorithms into our ML model has been pivotal in achieving these results. Unlike conventional text-to-speech systems, which often produce flat and monotone outputs, our approach dynamically adjusts the voice modulation to reflect the emotional context of the text. This advancement is crucial for creating engaging and emotionally resonant audiobooks.

The consistent performance across Accuracy, Precision, Recall, and F1-Score suggests that our model is well-balanced and does not sacrifice one aspect of performance for another. The relatively low RMSE further supports the model's reliability in producing high-quality audio outputs with minimal error.

4.5.3 Deviation and Improvement

While the results are promising, the RMSE of 0.82, although relatively low, indicates room for improvement in reducing prediction errors. Future work could focus on fine-tuning the model to further decrease this error, potentially through more granular emotion recognition and enhanced voice modulation techniques.

Additionally, while our current model performs well, testing it across diverse genres and styles of writing could help identify specific areas where it may need further adjustment. This will ensure that the model is versatile and effective in various contexts, enhancing its practical applicability in real-world audiobook production.

In conclusion, our ML-based approach to audiobook creation has demonstrated significant potential, offering a substantial improvement over traditional methods by incorporating nuanced emotional and vocal modulation. These results lay a strong foundation for future developments and optimizations in the field of automated audiobook production.

Chapter 5

Conclusion

5.1 Conclusion

The primary objective of converting text and PDF documents into emotionally expressive audio files was successfully met. The integration of a logistic regression model for emotion recognition allowed the audio output to reflect the emotional tone of the text, resulting in more engaging and lifelike auditory experiences. The model achieved a commendable accuracy of 88 percentage, with a precision of 0.83 and a recall of 0.84, demonstrating its effectiveness in identifying and expressing various emotions.

However, the project also highlighted certain limitations. A significant limitation was that the audio produced did not sound natural enough, which affected the overall user experience. This issue suggests the need for more advanced text-to-speech (TTS) systems capable of generating more natural-sounding voices. Additionally, the model performed best with texts containing explicit emotional cues but faced challenges with ambiguous or nuanced emotional content, indicating a need for more sophisticated emotion recognition techniques or additional training data.

Despite these challenges, the project laid a solid foundation for future improvements. Potential enhancements could include incorporating more advanced machine learning models, such as deep learning techniques, to improve emotion recognition accuracy and the naturalness of the generated speech. Additionally, expanding the dataset to cover a broader spectrum of emotions and contexts could further enhance the system's robustness and versatility. Overall, this project represents a significant step forward in the field of text-to-audio conversion, offering promising prospects for creating more immersive and emotionally resonant auditory experiences.

5.2 Future Scope

To enhance the naturalness of the audio output, future work could involve integrating advanced text-to-speech (TTS) technologies, such as neural TTS models like Tacotron or WaveNet, which are capable of producing more human-like and emotionally expressive voices. Another key area for improvement is the system's efficiency in processing large PDF files; optimizing the text extraction and processing algorithms would enable the system to handle big files more quickly and seamlessly. Additionally, expanding the emotion recognition model to include a broader range of emotions, beyond basic ones like joy, sadness, and anger, to encompass more subtle emotions such as surprise, disgust, and fear, would provide a more nuanced and comprehensive emotional output. Further, incorporating contextual understanding and sentiment analysis could enhance the model's ability to interpret and express complex emotional cues accurately. These enhancements would significantly improve the overall performance, user experience, and versatility of the text-to-audio conversion system.

Bibliography

[1] Alotaibi, Fahad Mazaed. "Classifying text-based emotions using logistic regression." VAWKUM Transactions on Computer Sciences 7, no. 1 (2019): 31-37.

[2] Murthy, Ashritha R., and KM Anil Kumar. "A review of different approaches for detecting emotion from text." In IOP Conference Series: Materials Science and Engineering, vol. 1110, no. 1, p. 012009. IOP Publishing, 2021.

[3] Seyeditabari, Armin, Narges Tabari, and Wlodek Zadrozny. "Emotion detection in text: a review." arXiv preprint arXiv:1806.00674 (2018).

Appendix A: Presentation

TALKBOOK

Advancing Book To Audiobook Conversion Through Emotion Aware Machine Learning

MR. BIJU ABRAHAM .N

NANDHANA SUFFIN
NIKHIL STEPHEN
NIVEDITHA .B
RACHEL JACOB

Contents

1. Introduction
2. Problem Definition
3. Objectives
4. Scope and Relevance
5. System Design
6. Datasets
7. Work Division – Gantt Chart
8. Software/Hardware Requirements
9. Results
10. Conclusion
11. Future Enhancements
12. References

Introduction

- The Emotion-Aware Audiobook Conversion project is a contribution to the audiobook industry by leveraging machine learning technologies to enhance the emotional depth and engagement of audiobook narration.
- Used for converting written books into audiobooks using emotion-aware machine learning techniques.
- Helps to enhance the audiobook experience by accurately conveying the emotional nuances present in the text, thereby creating a more engaging and immersive listening experience for users.
- This Web Application improves accessibility for individuals with visual impairments or other disabilities that may hinder their ability to read printed text.

Problem Definition

To design, develop and implement a software for the conversion of text to corresponding audio with the integration of emotions with it.

Objectives

- **Text-to-Speech Conversion Accuracy:** Develop a robust text-to-speech (TTS) model that accurately converts written text into high-quality audio.
- **Emotion Recognition:** Implement machine learning techniques to analyze the emotional content of the text. Identify key emotions such as joy, sadness, surprise, or anger within the narrative.
- **Adaptability Across Genres:** Ensure that the emotion detection and TTS models are adaptable across various genres of literature (e.g., fiction, non-fiction) and can effectively convey different emotional nuances.
- **User Interface and Accessibility:** User-friendly interface that ensures accessibility features for users with diverse needs.

Scope and Relevance

This project has a wide range of applications across various fields, from accessibility and education to entertainment and market research. It has the potential to revolutionize the way we interact with technology and each other, making communication more engaging, natural, and effective.

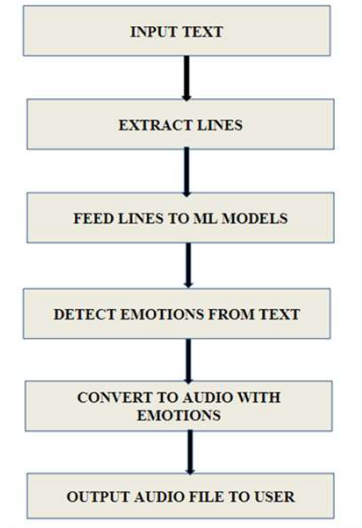
- **Accessibility** = For individuals with visual impairments
- **Language Learning** = Enhancing their learning experience
- **Therapeutic Applications** = Such as for individuals with autism or social anxiety
- **Interactive Systems** = Adding emotional cues can make interactions more natural and engaging

System Design

System Overview

The primary goal of this project is to develop a TTS system capable of synthesizing speech with integrated emotional cues. Rather than producing monotone speech, the system will be designed to convey a range of emotions such as Joy, sadness, surprise, anger, love and fear depending on the input text and context.

Architectural Design



Modules

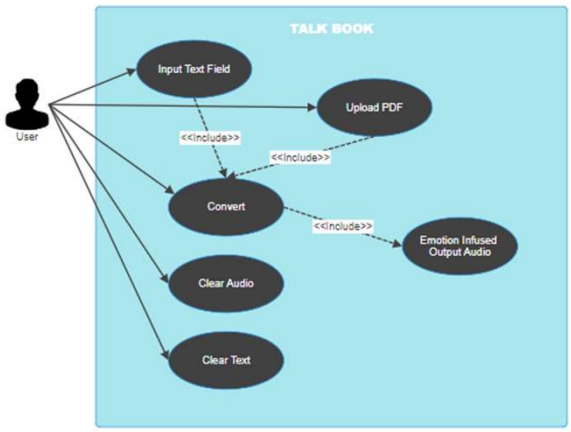
- Input Text / Upload pdf
- Extract lines and Detect Emotions
- Convert to Audio with Emotions
- Output Generated

ALGORITHM

1. Input text into the web application .
1. Extraction of lines from the text for conversion.
1. Providing the extracted lines to logistic regression ML models for conversion.

   3.1 . Detect emotions from the lines.
1. Convert the lines to audio files with respective emotions and modulations.
1. Output the audio file to the user.

Design Models



	A	B	C
1	text	label	
2	i didnt feel	0	
3	i can go fr	0	
4	im grabbin	3	
5	i am ever f	2	
6	i am feelin	3	
7	ive been fe	0	
8	ive been ta	5	
9	i feel as cc	4	
10	i have bee	1	
11	i feel roma	2	
12	i feel like i	0	
13	i do feel th	1	
14	i think it s t	3	
15	i feel low e	0	
16	i have imm	1	
17	i do not fe	1	
18	i didnt real	0	
19	i feel prett	0	
20	i started fe	0	
21	i now feel	4	
22	i feel irrita	3	
23	i am feelin	4	
24	i have the	1	
25	i was able	1	
26	i already fe	3	
27	i still love r	0	
28	i feel so in	0	

Datasets

0 = Sadness
1 = Joy
2 = Love
3 = Anger
4 = Fear
5 = Surprise

Size of dataset = 16001

Work Division

S.NO.	Task name	Start	Finish
1	Research and requirement analysis	March 2,2024	March 7,2024
2	Requirement Specification	February 3,2024	February 7,2024
3	Design Layout	March 8,2024	March 15,2024
4	Front-end Development	March 25,2024	April 12,2024
5	Emotion Recognition Implementation	March 25,2024	April 10,2024
6	Obtaining audio modulation parameters	April 13,2024	April 20,2024
7	Integrating TTS,ML and audion generation modules	April 27,2024	May 1,2024
8	Integrating Frontend and Backend	May 2,2024	May 4,2024
9	Testing and Deployment	May 5,2024	May 6,2024

Software/ Hardware Requirements

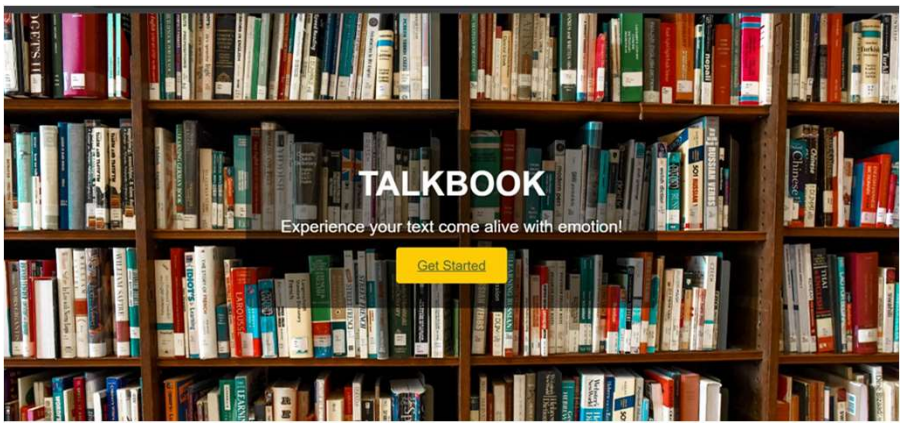
SOFTWARE

- pyttsx3 - Text to speech conversion
- numpy,pydub - Audio processing
- Logistic Regression - Emotion sensing package
- scikit-learn (sklearn) - Machine learning tasks
- Windows - Operating System
- Flask - Integration

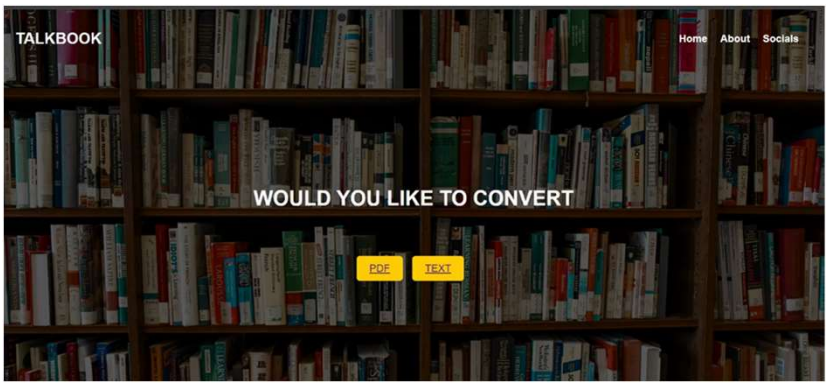
HARDWARE

- Processor: Intel i3 and above
- RAM : 4 GB and above
- Hard disk : 100 GB and above
- Monitor : LCD monitor
- Keyboard : Normal or Multimedia
- Mouse : Compatible mouse

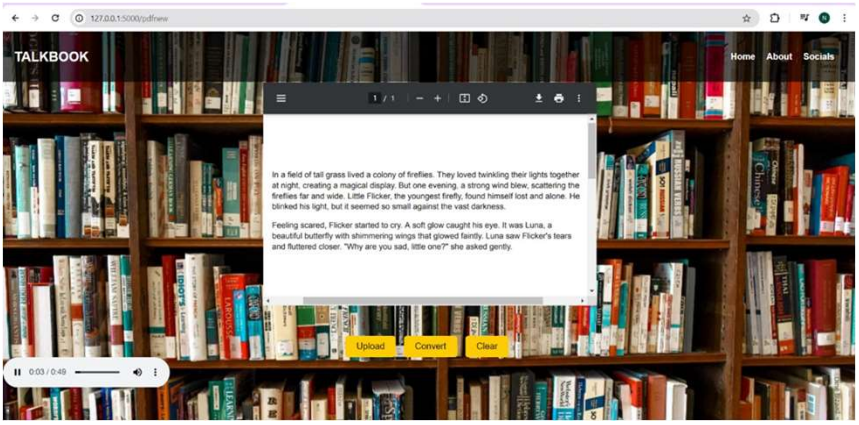
Results



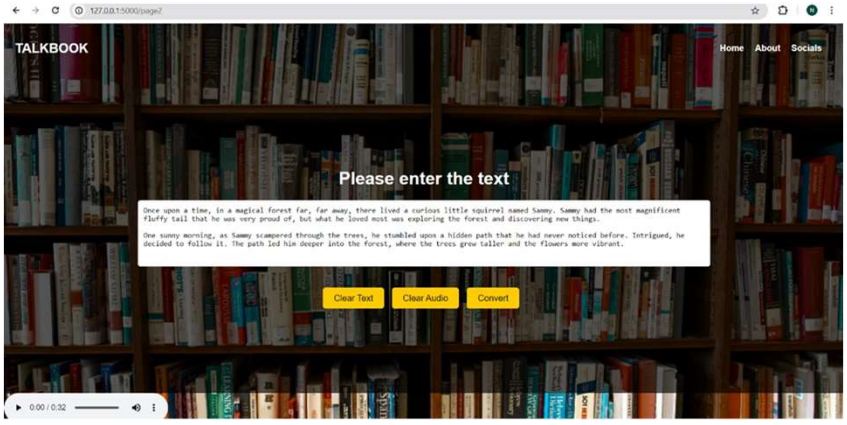
1) Landing page



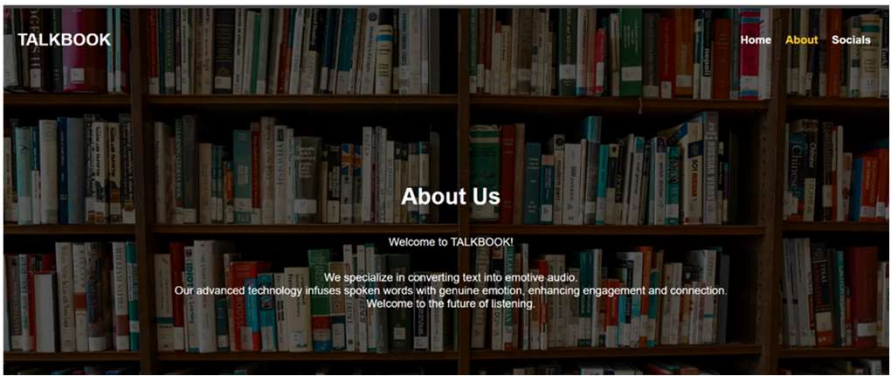
2) Selection Page



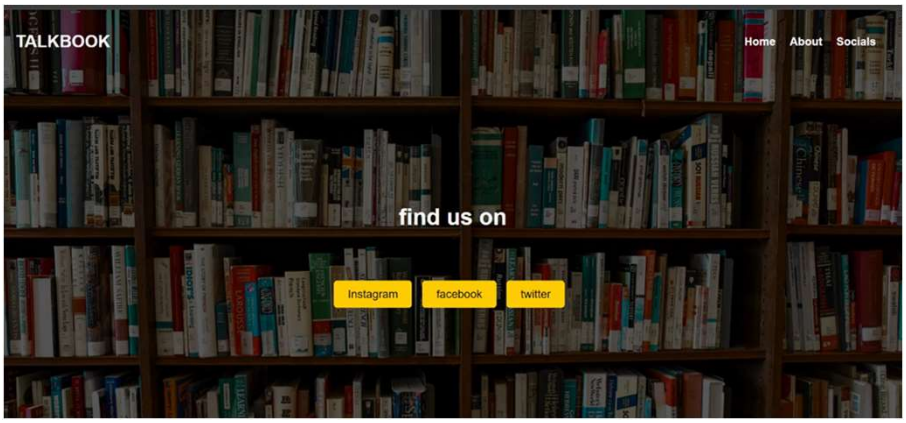
3) PDF to Audio



4) TEXT to Audio



5) About Us Page



6) Socials Page

Conclusion

- The Emotion-Enhanced Text-to-Speech (E-TTS) project aims to synthesize speech with integrated emotional cues, enhancing naturalness and expressiveness in human-computer interactions.
- Overall, the Emotion-Enhanced Text-to-Speech project represents a significant step towards more nuanced and engaging synthetic speech, with the potential to transform how we interact with technology and communicate with one another in diverse contexts.

Future Enhancements

- Natural sounding Audio
- Heavier loads of text content to convert to audio
- Downloading the audio files
- More number of emotions

References

- Alotaibi, Fahad Mazaed. "Classifying text-based emotions using logistic regression." VAWKUM Transactions on Computer Sciences 7, no. 1 (2019): 31-37.
- Murthy, Ashritha R., and KM Anil Kumar. "A review of different approaches for detecting emotion from text." In IOP Conference Series: Materials Science and Engineering, vol. 1110, no. 1, p. 012009. IOP Publishing, 2021.
- Seyeditabari, Armin, Narges Tabari, and Wlodek Zadrozny. "Emotion detection in text: a review." arXiv preprint arXiv:1806.00674 (2018).

Appendix B: Vision, Mission, Programme Outcomes and Course Outcomes

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
RAJAGIRI SCHOOL OF ENGINEERING & TECHNOLOGY (AUTONOMOUS)
RAJAGIRI VALLEY, KAKKANAD, KOCHI, 682039
(Affiliated to APJ Abdul Kalam Technological University)

Vision, Mission, Programme Outcomes and Course Outcomes

Institute Vision

To evolve into a premier technological institution, moulding eminent professionals with creative minds, innovative ideas and sound practical skill, and to shape a future where technology works for the enrichment of mankind.

Institute Mission

To impart state-of-the-art knowledge to individuals in various technological disciplines and to inculcate in them a high degree of social consciousness and human values, thereby enabling them to face the challenges of life with courage and conviction.

Department Vision

To become a centre of excellence in Computer Science and Engineering, moulding professionals catering to the research and professional needs of national and international organizations.

Department Mission

To inspire and nurture students, with up-to-date knowledge in Computer Science and Engineering, ethics, team spirit, leadership abilities, innovation and creativity to come out with solutions meeting societal needs.

Programme Outcomes (PO)

Engineering Graduates will be able to:

1. Engineering Knowledge: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

2. Problem analysis: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

3. Design/development of solutions: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

4. Conduct investigations of complex problems: Use research-based knowledge including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

5. Modern Tool Usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

6. The engineer and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal, and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

7. Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

8. Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

9. Individual and Team work: Function effectively as an individual, and as a member or leader in teams, and in multidisciplinary settings.

10. Communication: Communicate effectively with the engineering community and with society at large. Be able to comprehend and write effective reports documentation. Make effective presentations, and give and receive clear instructions.

11. Project management and finance: Demonstrate knowledge and understanding of engineering and management principles and apply these to one's own work, as a member and leader in a team. Manage projects in multidisciplinary environments.

12. Life-long learning: Recognize the need for, and have the preparation and ability to engage in independent and lifelong learning in the broadest context of technological change.

Programme Specific Outcomes (PSO)

A graduate of the Computer Science and Engineering Program will demonstrate:

PSO1: Computer Science Specific Skills

The ability to identify, analyze and design solutions for complex engineering problems in multidisciplinary areas by understanding the core principles and concepts of computer science and thereby engage in national grand challenges.

PSO2: Programming and Software Development Skills

The ability to acquire programming efficiency by designing algorithms and applying standard practices in software project development to deliver quality software products meeting the demands of the industry.

PSO3: Professional Skills

The ability to apply the fundamentals of computer science in competitive research and to develop innovative products to meet the societal needs thereby evolving as an eminent researcher and entrepreneur.

Course Outcomes

After the completion of the course the student will be able to:

CO1:

Identify technically and economically feasible problems (Cognitive Knowledge Level: Apply)

CO2:
Identify and survey the relevant literature for getting exposed to related solutions and get familiarized with software development processes (Cognitive Knowledge Level: Apply)

CO3:
Perform requirement analysis, identify design methodologies and develop adaptable & reusable solutions of minimal complexity by using modern tools & advanced programming techniques (Cognitive Knowledge Level: Apply)

CO4:
Prepare technical report and deliver presentation (Cognitive Knowledge Level: Apply)

CO5:
Apply engineering and management principles to achieve the goal of the project (Cognitive Knowledge Level: Apply)

Appendix C: CO-PO-PSO Mapping

COURSE OUTCOMES:

After completion of the course the student will be able to

SL. NO	DESCRIPTION	Blooms' Taxonomy Level
CO1	Identify technically and economically feasible problems (Cognitive Knowledge Level: Apply)	Level 3: Apply
CO2	Identify and survey the relevant literature for getting exposed to related solutions and get familiarized with software development processes (Cognitive Knowledge Level: Apply)	Level 3: Apply
CO3	Perform requirement analysis, identify design methodologies and develop adaptable & reusable solutions of minimal complexity by using modern tools & advanced programming techniques (Cognitive Knowledge Level: Apply)	Level 3: Apply
CO4	Prepare technical report and deliver presentation (Cognitive Knowledge Level: Apply)	Level 3: Apply
CO5	Apply engineering and management principles to achieve the goal of the project (Cognitive Knowledge Level: Apply)	Level 3: Apply

CO-PO AND CO-PSO MAPPING

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12	PSO 1	PSO 2	PSO3
CO1	3	3	3	3		2	2	3	2	2	2	3	2	2	2
CO2	3	3	3	3	3	2		3	2	3	2	3	2	2	2
CO3	3	3	3	3	3	2	2	3	2	2	2	3			2
CO4	2	3	2	2	2			3	3	3	2	3	2	2	2
CO5	3	3	3	2	2	2	2	3	2		2	3	2	2	2

3/2/1: high/medium/low

JUSTIFICATIONS FOR CO-PO MAPPING

MAPPING	LOW/ MEDIUM/ HIGH	JUSTIFICATION
101003/CS622T.1-PO1	**HIGH**	Identify technically and economically feasible problems by applying the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
101003/CS622T.1-PO2	**HIGH**	Identify technically and economically feasible problems by analysing complex engineering problems reaching substantiated conclusions using first principles of mathematics.
101003/CS622T.1-PO3	**HIGH**	Design solutions for complex engineering problems by identifying technically and economically feasible problems.
101003/CS622T.1-PO4	**HIGH**	Identify technically and economically feasible problems by analysis and interpretation of data.
101003/CS622T.1-PO6	**MEDIUM**	Responsibilities relevant to the professional engineering practice by identifying the problem.
101003/CS622T.1-PO7	**MEDIUM**	Identify technically and economically feasible problems by understanding the impact of the professional engineering solutions.
101003/CS622T.1-PO8	**HIGH**	Apply ethical principles and commit to professional ethics to identify technically and economically feasible problems.
101003/CS622T.1-PO9	**MEDIUM**	Identify technically and economically feasible problems by working as a team.
101003/CS622T.1-PO10	**MEDIUM**	Communicate effectively with the engineering community by identifying technically and economically feasible problems.
101003/CS622T.1-P011	**MEDIUM**	Demonstrate knowledge and understanding of engineering and management principles by selecting the technically and economically feasible problems.
101003/CS622T.1-PO12	**HIGH**	Identify technically and economically feasible problems for long term learning.
101003/CS622T.1-PSO1	**MEDIUM**	Ability to identify, analyze and design solutions to identify technically and economically feasible problems.
101003/CS622T.1-PSO2	**MEDIUM**	By designing algorithms and applying standard practices in software project development and Identifying technically and economically feasible problems.
101003/CS622T.1-PSO3	**MEDIUM**	Fundamentals of computer science in competitive research can be applied to Identify technically and economically feasible problems.
101003/CS622T.2-PO1	**HIGH**	Identify and survey the relevant by applying the knowledge of mathematics, science, engineering fundamentals.

101003/CS622T.2-PO2	**HIGH**	Identify, formulate, review research literature, and analyze complex engineering problems get familiarized with software development processes.
101003/CS622T.2-PO3	**HIGH**	Design solutions for complex engineering problems and design based on the relevant literature.
101003/CS622T.2-PO4	**HIGH**	Use research-based knowledge including design of experiments based on relevant literature.
101003/CS622T.2-PO5	**HIGH**	Identify and survey the relevant literature for getting exposed to related solutions and get familiarized with software development processes by using modern tools.
101003/CS622T.2-PO6	**MEDIUM**	Create, select, and apply appropriate techniques, resources, by identifying and surveying the relevant literature.
101003/CS622T.2-PO8	**HIGH**	Apply ethical principles and commit to professional ethics based on the relevant literature.
101003/CS622T.2-PO9	**MEDIUM**	Identify and survey the relevant literature as a team.
101003/CS622T.2-PO10	**HIGH**	Identify and survey the relevant literature for a good communication to the engineering fraternity.
101003/CS622T.2-PO11	**MEDIUM**	Identify and survey the relevant literature to demonstrate knowledge and understanding of engineering and management principles.
101003/CS622T.2-PO12	**HIGH**	Identify and survey the relevant literature for independent and lifelong learning.
101003/CS622T.2-PSO1	**MEDIUM**	Design solutions for complex engineering problems by Identifying and survey the relevant literature.
101003/CS622T.2-PSO2	**MEDIUM**	Identify and survey the relevant literature for acquiring programming efficiency by designing algorithms and applying standard practices.
101003/CS622T.2-PSO3	**MEDIUM**	Identify and survey the relevant literature to apply the fundamentals of computer science in competitive research.
101003/CS622T.3-PO1	**HIGH**	Perform requirement analysis, identify design methodologies by using modern tools & advanced programming techniques and by applying the knowledge of mathematics, science, engineering fundamentals.
101003/CS622T.3-PO2	**HIGH**	Identify, formulate, review research literature for requirement analysis, identify design methodologies and develop adaptable & reusable solutions.

101003/CS622T.3-PO3	**HIGH**	Design solutions for complex engineering problems and perform requirement analysis, identify design methodologies.
101003/CS622T.3-PO4	**HIGH**	Use research-based knowledge including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
101003/CS622T.3-PO5	**HIGH**	Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools.
101003/CS622T.3-PO6	**MEDIUM**	Perform requirement analysis, identify design methodologies and assess societal, health, safety, legal, and cultural issues.
101003/CS622T.3-PO7	**MEDIUM**	Understand the impact of the professional engineering solutions in societal and environmental contexts and Perform requirement analysis, identify design methodologies and develop adaptable & reusable solutions.
101003/CS622T.3-PO8	**HIGH**	Perform requirement analysis, identify design methodologies and develop adaptable & reusable solutions by applying ethical principles and commit to professional ethics.
101003/CS622T.3-PO9	**MEDIUM**	Function effectively as an individual, and as a member or leader in teams, and in multidisciplinary settings.
101003/CS622T.3-PO10	**MEDIUM**	Communicate effectively with the engineering community and with society at large to perform requirement analysis, identify design methodologies.
101003/CS622T.3-PO11	**MEDIUM**	Demonstrate knowledge and understanding of engineering requirement analysis by identifying design methodologies.
101003/CS622T.3-PO12	**HIGH**	Recognize the need for, and have the preparation and ability to engage in independent and lifelong learning in the broadest context of technological change by analysis, identify design methodologies and develop adaptable & reusable solutions.
101003/CS622T.3-PSO3	**MEDIUM**	The ability to apply the fundamentals of computer science in competitive research and prior to that perform requirement analysis, identify design methodologies.
101003/CS622T.4-PO1	**MEDIUM**	Prepare technical report and deliver presentation by applying the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
101003/CS622T.4-PO2	**HIGH**	Identify, formulate, review research literature, and analyze complex engineering problems by preparing technical report and deliver presentation.

101003/CS622T.4-PO3	**MEDIUM**	Prepare Design solutions for complex engineering problems and create technical report and deliver presentation.
101003/CS622T.4-PO4	**MEDIUM**	Use research-based knowledge including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions and prepare technical report and deliver presentation.
101003/CS622T.4-PO5	**MEDIUM**	Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools and Prepare technical report and deliver presentation.
101003/CS622T.4-PO8	**HIGH**	Prepare technical report and deliver presentation by applying ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
101003/CS622T.4-PO9	**HIGH**	Prepare technical report and deliver presentation effectively as an individual, and as a member or leader in teams, and in multidisciplinary settings.
101003/CS622T.4-PO10	**HIGH**	Communicate effectively with the engineering community and with society at large by prepare technical report and deliver presentation.
101003/CS622T.4-PO11	**MEDIUM**	Demonstrate knowledge and understanding of engineering and management principles and apply these to one's own work by prepare technical report and deliver presentation.
101003/CS622T.4-PO12	**HIGH**	Recognize the need for, and have the preparation and ability to engage in independent and lifelong learning in the broadest context of technological change by prepare technical report and deliver presentation.
101003/CS622T.4-PSO1	**MEDIUM**	Prepare a technical report and deliver presentation to identify, analyze and design solutions for complex engineering problems in multidisciplinary areas.
101003/CS622T.4-PSO2	**MEDIUM**	To acquire programming efficiency by designing algorithms and applying standard practices in software project development and to prepare technical report and deliver presentation.
101003/CS622T.4-PSO3	**MEDIUM**	To apply the fundamentals of computer science in competitive research and to develop innovative products to meet the societal needs by preparing technical report and deliver presentation.
101003/CS622T.5-PO1	**HIGH**	Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
101003/CS622T.5-PO2	**HIGH**	Identify, formulate, review research literature, and analyze complex engineering problems by applying engineering and management principles to achieve the goal of the project.

101003/CS622T.5-PO3	**HIGH**	Apply engineering and management principles to achieve the goal of the project and to design solutions for complex engineering problems and design system components or processes that meet the specified needs.
101003/CS622T.5-PO4	**MEDIUM**	Apply engineering and management principles to achieve the goal of the project and use research-based knowledge including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
101003/CS622T.5-PO5	**MEDIUM**	Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools and to apply engineering and management principles to achieve the goal of the project.
101003/CS622T.5-PO6	**MEDIUM**	Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal, and cultural issues and the consequent responsibilities by applying engineering and management principles to achieve the goal of the project.
101003/CS622T.5-PO7	**MEDIUM**	Understand the impact of the professional engineering solutions in societal and environmental contexts, and apply engineering and management principles to achieve the goal of the project.
101003/CS622T.5-PO8	**HIGH**	Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice and to use the engineering and management principles to achieve the goal of the project.
101003/CS622T.5-PO9	**MEDIUM**	Function effectively as an individual, and as a member or leader in teams, and in multidisciplinary settings and to apply engineering and management principles to achieve the goal of the project.
101003/CS622T.5-PO11	**MEDIUM**	Demonstrate knowledge and understanding of engineering and management principles and apply these to one's own work, as a member and leader in a team. Manage projects in multidisciplinary environments and to apply engineering and management principles to achieve the goal of the project.
101003/CS622T.5-PO12	**HIGH**	Recognize the need for, and have the preparation and ability to engage in independent and lifelong learning in the broadest context of technological change and to apply engineering and management principles to achieve the goal of the project.
101003/CS622T.5-PSO1	**MEDIUM**	The ability to identify, analyze and design solutions for complex engineering problems in multidisciplinary areas. Apply engineering and management principles to achieve the goal of the project.

101003/CS622T.5-PSO2	**MEDIUM**	The ability to acquire programming efficiency by designing algorithms and applying standard practices in software project development to deliver quality software products meeting the demands of the industry and to apply engineering and management principles to achieve the goal of the project.
101003/CS622T.5-PSO3	**MEDIUM**	The ability to apply the fundamentals of computer science in competitive research and to develop innovative products to meet the societal needs thereby evolving as an eminent researcher and entrepreneur and apply engineering and management principles to achieve the goal of the project.