**PDF Voice: AI-Powered PDF to Speech Converter with Customizable Audio Settings using MERN Stack**

**Abstract:**

**PDFVoice:** AI-Powered MERN Stack PDF to Speech Converter with

Customizable Audio Settings is an advanced web application that converts PDF

documents into high-quality speech using AI-driven text-to-speech (TTS)

synthesis. Built on the MERN (MongoDB, Express.js, React.js, Node.js) stack,

the system integrates OCR (Optical Character Recognition), Natural Language

Processing (NLP), and deep learning-based TTS models to deliver realistic and

natural speech output. The application allows users to upload PDFs, select a

preferred language, and adjust audio parameters such as volume and pitch.

Using AI-based text extraction, the system accurately extracts text from

documents. The built-in language detection and translation module ensures that

documents in different languages are automatically detected and translated

before speech synthesis. For speech generation, the platform leverages Google

Wave Net, VITS, and SSML (Speech Synthesis Markup Language) to enhance

speech clarity, add natural pauses, and emphasize key phrases.

Users can also download the generated speech file for offline use, making it a

valuable tool for visually impaired individuals, professionals, and multilingual

users. By combining AI-powered OCR, NLP, and deep learning-based TTS ,

PDF Voice provides an efficient, accessible, and customizable speech generation

solution for document-based content consumption..

Dr. N. Sree Divya

**Internal Supervisor**

**IOMP ID:**IT-25-27

**Name 1:** Pendyala Jaya Charan

**Roll No :** 22261A1247

**Name 2:** Chenamoni Vaishnavi

**Roll No :**22261A1214

(Mrs. U.Chaitanya)

**IOMP Supervisor**

