


TextEvolve

Digitizing History with AI/OCR

Team Name : Dynamic Dreamers

NMNT ID : NMNTSTD7116020

College Name : Kathir College of Engineering

Theme: Artificial Intelligence

Problem Statement:

How might we develop an AI or OCR solution to digitize and convert handwritten, old registered documents into a readable and accessible format in regional languages, improving public access and readability of historical records? The solution should allow the output to be downloaded in various formats such as PDF and Word, enabling wider distribution and accessibility.

Existing System:

- Traditional methods of digitization are slow and prone to errors.
- Manual transcription is time-consuming and expensive.
- Existing OCR models have limited accuracy for handwritten and regional texts.
- Lack of language diversity and support for historical scripts.

Proposed System:

- AI-driven OCR solution for handwritten and regional languages.
- Combines a custom ML model with Google OCR API for improved accuracy.
- Fast, scalable, and easy-to-use platform with real-time results.
- Secure data storage and enhanced accessibility.

- **Google OCR API:** For high-accuracy text extraction.
- **TensorFlow, PyTorch, Keras:** For building and training the ML model.
- **MongoDB:** For flexible and scalable document storage.
- **Flask:** For backend service and API integration.
- **React Vite & Tailwind:** For responsive and fast UI.

Custom ML Model:

- Designed specifically for handwritten and regional scripts.
- Improves recognition accuracy over time with continuous training.

Dual OCR Integration:

- Custom model for initial recognition.
- Google OCR API for enhanced accuracy and corrections.

Language Versatility:

- Language Versatility

Robust Data Handling:

- Python Flask backend paired with MongoDB ensures scalable and secure storage.

User-Friendly Platform:

- Modern UI for easy upload and retrieval.
- Real-time results and progress tracking.

Technology Stack

Programming Languages


Database & Deployment


Frameworks & Libraries





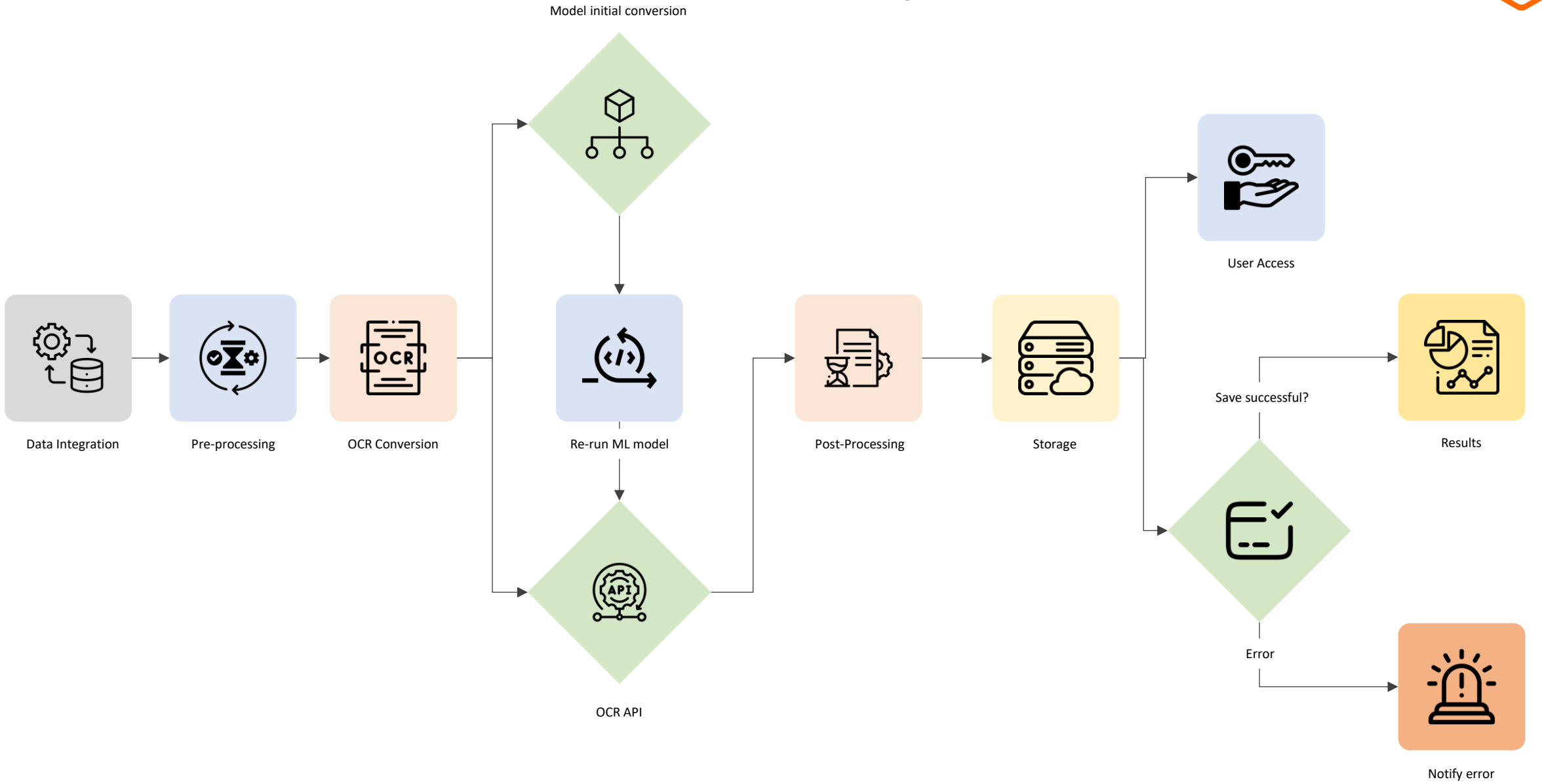

Development Tools



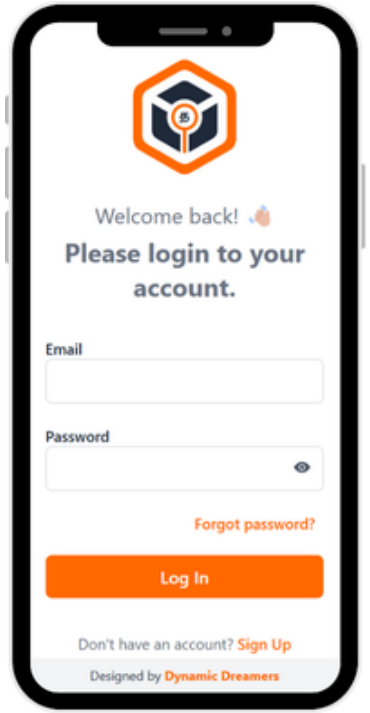


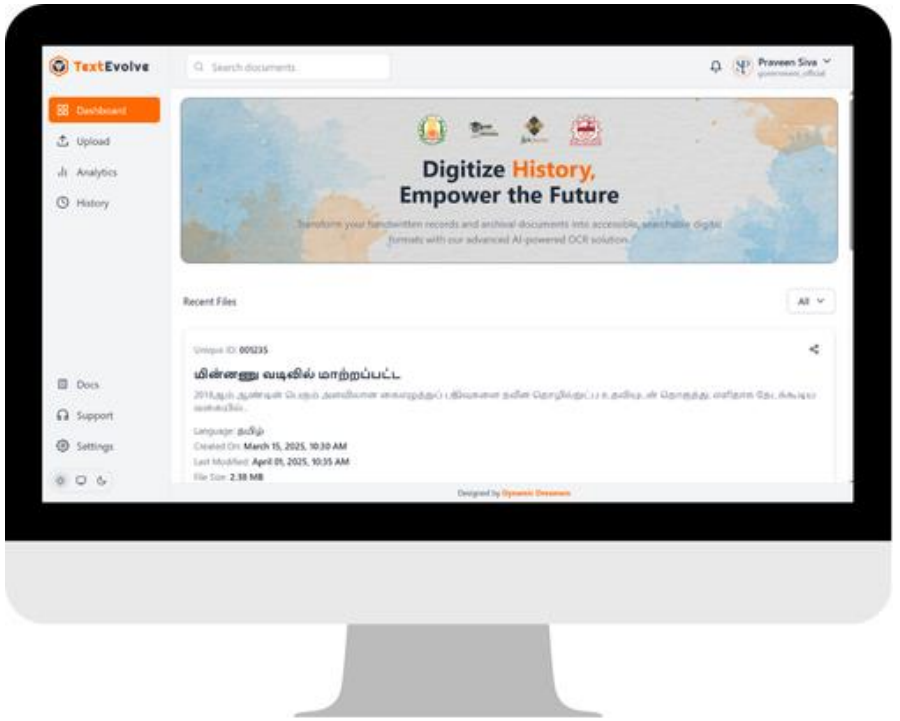


System Architecture & Workflow



Demo / Sample Results





https://squadofcreators.github.io/TextEvolve

Degradation:

Physical documents deteriorate over time.

Inaccessibility:

Limited access to rare or regional texts.

Handwriting Variability:

Handwritten texts present difficulties in standard OCR processing.

Language Barriers:

Many OCR solutions are not optimized for regional languages.

Cultural Preservation: Safeguard historical documents for future generations.

Enhanced Accessibility: Allow researchers, educators, and the public to easily access digitized content.

Improved Research: Facilitate the study of regional languages and historical texts.

Scalability: A robust system that can grow to handle vast archives.

TextEVolve is an AI-driven solution that digitizes handwritten historical documents with high accuracy using a **custom ML model** and **Google OCR API**.

By combining innovation with technology, it preserves cultural heritage, enhances accessibility, and supports research — ensuring valuable historical knowledge is safeguarded for future generations.

Ask Questions

Our Team

Mrs. Kavitha M

Faculty Guide

Yuva Nandhini M

AI/ML Developer

Praveenkumar S

AI/MERN Developer

Sibi Siddharth S

AI/MERN Developer

Uma Maheswari P

AI/ML Developer

Preserving the Past,
Empowering the Future