PROJECT PROPOSAL

1. **Theme :** Artificial Intelligence

2. **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.

3. **College Code & College Name :** 7116 & Kathir College of Engineering

4. **Guide Name, Designation, Mobile No. & Email id :** Kavitha M, Assistant Professor/Artificial Intelligence and Data Science, +918939944709, kavitha@kathir.ac.in

5. Student Team details:

Sl. No.	Student Reg. No.	Name of the Student	Branch	Mobile No.	Email ID
1	711621243036	Praveenkumar S	AI&DS	+916381299183	praveensiva0820@gmail.com
2	711621243046	Sibi Siddharth S	AI&DS	+919629124660	sibisiddharth8@gmail.com
3	711621243053	Uma Maheswari P	AI&DS	+919994387880	uma200427@gmail.com
4	711621243059	Yuva Nandhini M	AI&DS	+918248502477	m.yuvanandhini@gmail.com

6. **Project Summary :** This project aims at developing an AI-based OCR system to digitize old handwritten documents in regional languages with their original digital text. Using the Google Vision API for its high accuracy in text recognition and a custom-built language model based on CNN and RNN, it can work well with complex scripts. The output will be further enhanced by OpenAI's GPT-4 to correct spelling, grammar, and other errors that are likely to be in the output. The user will be able to download the processed text in several web formats, including PDF and Word, depending on both desktop and mobile needs. In this solution it will also be possible for a user to use the application offline, as it does not need an internet connection.

7. **Proposed solution with methodology :**

- **OCR with Google Vision API:** High-precision OCR for scanned handwritten documents will be handled by the Google Vision API. It is quite adaptable, supporting more than 60 languages and integrating easily with different regional languages.

- **CNN-Based Custom Language Model:** To increase accuracy for particular regions, we will create regional models based on CNN and RNN architectures. These models are especially helpful in locations with regions with slow or no internet connectivity since they can be downloaded and used offline.

- **Text Post-Processing with GPT-4:** Once the OCR extracts text from images, OpenAI's GPT-4 will refine the text to make it readable by removing misinterpreted and inconsistent texts. This step will correct errors, increase clarity, and transform handwritten text into coherent phrases to avoid common misunderstandings.

- **Downloadable Output:** The processed text will be available for users to download in Word and PDF and various other formats. This feature helps with research, printing, archiving, and storage needs, making it simpler to use for the public.

- **Web and Mobile Application:**
  A secure web based app with authentication will be built using ReactJS, Vite, Tailwind CSS, and Firebase. The application will be optimized for both desktop and mobile devices, allowing users to upload images, process them, and download the output in the required format. The web app will include offline capabilities to process documents without requiring constant internet access via downloading the models.

- **Backend and Cloud Integration:**
  Google’s Firebase will be used for secure authentication and cloud storage purposes, while Flask will run the backend logic to handle the custom AI model’s execution. This backend will ensure that the users can securely upload documents, process them, and retrieve the results in their preferred format.

8. **Workplan / time schedule indicating the project mile stone :**

No.	Milestone	Duration
1	Research & Requirement Analysis	1 week
2	Configuration and Setup of Google Vision API	1 week
3	Custom Model Development (CNN & RNN)	2 weeks
4	Integration of GPT-4 for Text Correction	1 week
5	Development of Web based Application	3 weeks
6	Testing & Debugging the Application	1 week
7	Final Report, Documentation and Presentation	1 week

9. **Plan of action of implementation :**

Phase 1: Conduct research and gather datasets of handwritten documents in various regional (Indian) languages for model training.

Phase 2: Implement Google Vision API for initial OCR and evaluate its performance and accuracy across different languages and handwriting styles.

Phase 3: Develop and train a custom model using CNN and RNN tailored for regional languages to improve the accuracy and offline capabilities of text extraction process.

Phase 4: Integration of GPT-4 via API to process the extracted text, correcting errors and inconsistencies to improve the readability.

Phase 5: Development of a web based application with secure authentication, ensuring users can upload documents, images and download the processed output in PDF, Word or any other preferred format

Phase 6: Conduct testing, debug issues, and optimize the platform for enhanced usability and performance.

10. **List of facilities available in the college to develop the prototype of the project**

- High performance computers with dedicated graphics card for custom model training.
- High speed Internet connection and access to use cloud APIs and libraries.
- Access to AI/ML libraries and resources to guide the model development process.
- Subscription to firebase and other cloud platforms from student ids.

11. **Nature of Industry support for the project :**

- **Google Vision API:** Provides an advanced OCR service with support for over 60 languages.
- **OpenAI GPT-4:** Access to a powerful AI model for text correction and natural language processing.
- **Firebase:** Cloud storage and backend services for real-time database management and authentication.
- **Cloud Providers (Optional):** Hosting services for deploying the application and models.

12. **Total Cost :**

Estimated Total Project Cost: ₹8750

This includes costs for API usage, cloud storage, software tools, and resources required for training custom models.

13. **Details of Financial assistance required :**

- **API Subscription:** ₹550/month
- **Model Training :** ₹5200/month
- **Cloud Storage & Hosting Costs:** ₹2000/month
- **Miscellaneous Expenses :** ₹1000 (for sample prototype)

14. **Expected outcomes / results :**

Accurate Digitization of Handwritten Documents: Convert old, handwritten documents into a readable, searchable, and editable digital format in regional languages.

Enhanced Accessibility: Provide a user-friendly web and mobile platform for easy access to digitized historical records.

Offline Capabilities: Custom models will allow the solution to work offline, ensuring accessibility in regions with limited internet connectivity.

Downloadable Output (PDF, Word): Users can download the processed text in the desired format (PDF or Word) for further use.

Improved Readability and Accuracy: Post-processing with GPT-4 will ensure that the output text is clear, coherent, and error-free.

Preservation of Historical Records: The solution helps preserve valuable handwritten documents in digital formats, making them more accessible to the public for future use.

UNDERTAKING

1. ALL the students are studying in final year engineering. All the students are registered only once for this scheme.
2. The college will provide the basic infrastructure and other required facilities to the students for timely completion of their projects.
3. The college assumes to undertake the financial and other management responsibilities of the project. We are aware that the amount is to be utilized only for the purpose sanctioned i.e. to meet the expenses for developing the prototype and not for purchase of computer consumables, stationaries, honorarium, overhead etc. Unutilised balance amount will be returned back to the University after the time of completion of the project.

Name and Sign of Student 1	Name and Sign of Student 2	Name and Sign of Student 3	Name and Sign of Student 4

Signature of the Mentor

Signature and seal of the principal