AI-Powered Medical Report Analyzer using Langchain

Project Objectives

The initiative sought to develop an innovative intelligent system capable of assessing medical reports in both text (PDF) and image formats (scans/photos) utilizing advanced AI technologies. The main aims included automating the retrieval of essential information from lab reports and prescriptions, applying AI methods to analyze and interpret medical images like X-rays and MRIs, and creating a medical chatbot to offer real-time, interactive assistance for diagnoses and patient inquiries. Ultimately, the objective was to use AI to improve the accessibility, accuracy, and efficiency of healthcare services, thus enhancing overall healthcare delivery.

Problem Statement

The existence of unstructured medical data—including PDFs, handwritten notes, and image scans—leads to considerable inefficiencies and risks in healthcare, such as delays in manual interpretation, variations in standards among institutions, and increased diagnostic errors resulting from incomplete or unclear records. Patients encounter difficulties in accessing or comprehending their data, as unstructured formats like free-text notes or non-digital images do not lend themselves to automated analysis. These challenges arise because 80% of healthcare data is unstructured, which complicates interoperability and places an additional burden on clinicians with time-consuming documentation tasks that contribute to burnout. Tackling these issues necessitates an AI-powered system that can process various data types (text, images, handwritten content), standardize information extraction, and offer user-friendly interfaces for both healthcare professionals and patients to enhance accuracy, minimize delays, and improve data accessibility.

Existing Solutions

Current methods for managing healthcare data consist of tools such as Tesseract OCR and Google Vision for extracting text from images, electronic health record (EHR) systems that

require manual data input, general chatbots offering basic health information without incorporating the context of reports, and specialized applications for processing DICOM medical images. Although these technologies meet specific demands, they remain disjointed, tailored to particular domains, and lack a unified approach to integrating text, images, and conversational understanding. This lack of integration hinders their ability to provide cohesive and all-encompassing healthcare solutions.

Key Uniqueness

- Multimodal Input Capability: Able to handle both PDF documents and medical images in JPG/PNG format.
- Complete Automation: Combines OCR, NLP, computer vision, and chatbot features into one cohesive application.
- Integration with SAM: Implements Meta's Segment Anything Model for broader medical image segmentation.
- Medical Chatbot Functionality: Powered by ChatDoctor and Gemini AI to comprehend report details and respond to inquiries.
- Instant Visualization: Creates graphs for laboratory results and emphasizes segments of images automatically.
- No Manual Tagging Required: An entirely AI-driven solution that effectively manages noisy, real-world reports.

Tools

Domain	Technologies
OCR	Tesseract OCR, Microsoft LayoutLMv3
NLP	Hugging Face Transformers (Mistral-7B, BioBERT), ChatDoctor
LLM Chatbot	Google Gemini Pro, nics-lab/ChatDoctor
Image Segmentation	Meta's Segment Anything Model (SAM), OpenCV, PIL

Visualization	Matplotlib, Seaborn, Streamlit UI
PDF & Image Parsing	pdfplumber, PyMuPDF, Pytesseract
Dev & Deployment	VS Code, GitHub, Streamlit Cloud

Key Phases of Development

The development of the project comprised four essential stages, each focused on specific elements of the multimodal medical analysis system:

1. **Analysis of Text Reports**
   - PDF Extraction: Employed LayoutLMv3 alongside pdfplumber to retrieve organized data from unstructured PDF lab reports and prescriptions.
   - Diagnostic Extraction: Leveraged NLP models to pinpoint key test results, diagnoses, and clinical suggestions from the extracted text.
   - Visualization: Created interactive charts (such as blood glucose trends and lipid profiles) using Matplotlib/Seaborn for tracking lab parameters.

2. **Medical Image Assessment**
   - Image Segmentation: Combined Meta's Segment Anything Model (SAM) with OpenCV/PIL to automatically identify areas of interest in X-rays, CT scans, and MRIs.
   - Pathology Detection: Designed pipelines to highlight irregularities in the segmented images (such as fractures and tumors) utilizing computer vision techniques.
3. **Chatbot Integration**
   - Context-Aware AI: Integrated Google Gemini Pro with ChatDoctor to support queries specific to reports (e.g., "Clarify my elevated CRP levels").
   - Diagnostic Support: Programmed the chatbot to offer advice based on symptoms while clarifying its non-emergency applicability.

4. **Frontend Development**
   - Streamlit UI: Created a user-friendly interface for uploading PDFs/images and displaying parsed information.
   - Real-Time Display: Developed auto-refresh panels that showcase segmented images, lab graphs, and summaries of diagnostics.
   - Integrated Chat: Included the medical chatbot directly within the interface for a smooth user experience.

Project Outcomes

- Enhanced Precision in Text Extraction: Achieved over 90% accuracy in extracting text from actual lab reports.
- AI-Enhanced Image Segmentation: Executed AI-driven segmentation for medical images without the need for manual labeling.
- Improvements in Efficiency: Significantly decreased the time required to read and summarize medical documents.
- Integrated Multimodal Functionality: Effectively combined OCR, NLP, computer vision, and LLM-based chatbot features into one cohesive platform.
- User-Centric Design: Created an easy-to-use interface that is accessible to patients, physicians, and researchers.

Key Learnings

- AI Domain Integration: Merging Natural Language Processing (NLP), Computer Vision (CV), and chat functionalities leads to the creation of highly effective tools.
- Standardizing Medical Reports: The variation in medical report formats can be mitigated by utilizing tools like SAM and LayoutLMv3, which help in standardizing these formats.
- Conversational Healthcare Assistants: Advanced Language Models (LLMs) such as Gemini and Mistral support the creation of conversational healthcare assistants.
- Using Streamlit for Implementation: Streamlit allows for swift deployment and development of user interfaces.
- Equilibrating Healthcare Solutions: Practical healthcare solutions must find a balance between accuracy, usability, and transparency.

Github Link: Medi Ai (Github Code)
““‘https://github.com/Kevinjojo2003/MediAi””

Project Live Link : https://mediai.streamlit.app/