

Medical Report Analyzer Documentation

Project Description

The Medical Report Analyzer is a Streamlit-based web application designed to extract, process, and analyze text from medical reports in various formats. It leverages natural language processing and computer vision technologies to provide users with a structured analysis of medical documents, highlighting key medical insights, abnormalities, and recommended actions. This tool aims to help patients and healthcare professionals quickly understand medical reports without requiring specialized medical knowledge.

Key Features

- **Multi-format Document Support:** Processes PDF, DOCX, and image files (JPG, PNG)
- **Advanced Text Extraction:** GPT-4o Vision for enhanced image-based text extraction
- **Structured Medical Analysis:** Provides a comprehensive breakdown of medical reports organized by key categories
- **Interactive Web Interface:** User-friendly Streamlit interface for easy document uploading and analysis
- **Analysis Export:** Ability to download the generated analysis as a text file
- **Original Text Preservation:** Maintains original document structure including tables, bullets, and formatting

Technologies Used

- **Streamlit:** For the web application interface
- **OpenAI GPT-4o and GPT-4 Turbo:** For enhanced text extraction and medical analysis
- **PyMuPDF (fitz):** For PDF text extraction
- **python-docx:** For DOCX file processing
- **Pillow:** For image handling
- **Python standard libraries:** For file operations and handling

How It Works

1. Document Upload and Text Extraction

- User uploads a medical document through the Streamlit interface
- The application identifies the file type (PDF, DOCX, or image)

- For PDFs: PyMuPDF extracts text from all pages
- For DOCX: python-docx extracts text from paragraphs and tables
- For images:
 - GPT-4o Vision analyzes the image with enhanced understanding of medical document structures

2. Medical Report Analysis

- The extracted text is sent to GPT-4 Turbo with a specialized prompt
- The analysis is structured into six key sections:
 1. **Disease Identification:** Diagnosed conditions and severity levels
 2. **Abnormal Measurements:** Values outside normal ranges
 3. **Recommended Actions:** Key suggestions for improving health
 4. **Medical Consultation:** Specialist referral needs
 5. **Summary:** Key findings and required next steps

3. Results Presentation

- The analysis is displayed in the Streamlit interface using markdown formatting
- The original extracted text is available in an expandable section
- The user can download the analysis as a text file

Potential Use Cases

- **Patient Empowerment:** Helping patients understand complex medical reports and terminology
- **Healthcare Provider Efficiency:** Enabling quick review of external medical documents
- **Medical Education:** Training students to identify key components in medical reports
- **Research Data Extraction:** Efficiently processing large volumes of medical documents
- **Telehealth Support:** Enhancing remote healthcare consultations with rapid document analysis
- **Healthcare Coordination:** Improving communication between different healthcare providers by standardizing report interpretation

Summary

The Medical Report Analyzer is a powerful tool that bridges the gap between complex medical documentation and user understanding. By combining traditional document processing techniques with advanced AI capabilities, it transforms raw medical reports into structured, actionable insights. The application's flexibility in handling various document formats and its user-friendly interface make it accessible to both healthcare professionals and patients, ultimately contributing to better informed healthcare decisions and improved patient outcomes.