Below is a streamlined PRD focused on a minimal viable product (MVP) that handles up to three documents. This version prioritizes simplicity while validating the core idea of leveraging a multimodal LLM API to extract rich, domain-specific information from medical files.

**MVP Product Requirements Document (PRD)**

**Project Name:** Simple Medical Document Extraction MVP  
**Date:** March 2, 2025

**1. Introduction**

**Background**

Medical professionals and healthcare organizations often rely on presentations and reports that contain critical, yet unstructured, information. Manual extraction is inefficient, especially when working with documents like PDFs and PowerPoint decks. This MVP will demonstrate the feasibility of using a multimodal LLM API to extract key medical insights and taxonomy data from a limited number of documents (up to 3).

**Purpose**

To build a proof-of-concept application that can:

* Upload a small set of documents (maximum three).
* Use a multimodal LLM API to extract essential metadata and medical taxonomy.
* Present the extracted data in a simple, structured table format.

**2. Problem Statement**

* **Manual Extraction Inefficiency:** Manually processing even a few documents is time-consuming.
* **Lack of Structure:** Medical documents contain rich insights that are hard to capture without automation.
* **Proof-of-Concept Need:** A simple, focused MVP is required to validate the approach before scaling to larger volumes.

**3. Goals & Objectives**

**Primary Goals**

* **Automated Extraction:** Integrate with a multimodal LLM API to process and extract key data points from each document.
* **Proof-of-Concept:** Demonstrate the extraction process using up to 3 documents.
* **Structured Output:** Generate a basic table view of the extracted information for quick review.

**Objectives**

* Validate the feasibility of using LLMs for medical document analysis.
* Show that rich medical insights and taxonomies (e.g., key terms, diagnoses, treatment mentions) can be reliably extracted.
* Provide a foundation for later scaling and additional features.

**4. Target Users**

* **Medical Data Analysts:** Interested in quick extraction of key information from a few critical documents.
* **Healthcare Professionals:** Looking to automate and validate manual data extraction from presentations and reports.
* **Technical Proof-of-Concept Evaluators:** Early adopters and stakeholders wanting to test the technology before further investment.

**5. User Stories**

* **Document Upload:**
  + *As a user, I want to upload up to three documents (PDF/PPT) so I can quickly test the extraction process.*
* **Extraction via LLM API:**
  + *As a user, I want the system to call the multimodal LLM API to extract a few key fields such as slide/page title, a brief summary, and relevant medical taxonomy (e.g., diagnoses or treatments) so that I can see the rich data captured from the documents.*
* **Results Presentation:**
  + *As a user, I want to see the extracted data displayed in a simple table so I can review and export it easily.*

**6. Functional Requirements**

**6.1 File Ingestion**

* **Upload Interface:** A basic UI that allows users to select and upload up to three documents.
* **Supported Formats:** Ensure compatibility with at least PPT/PPTX and PDF files.

**6.2 Pre-processing & Segmentation**

* **Document Parsing:** Convert documents into a processable format and segment them (e.g., by slides or pages).

**6.3 LLM Extraction Engine**

* **API Integration:** Interface with the multimodal LLM API to extract:
  + Basic metadata (e.g., slide title or page header).
  + A short summary or key insights.
  + Medical taxonomy elements (e.g., diagnoses, treatment options, key terms).
* **Minimal Customization:** Pre-define a simple set of fields to extract without complex user configuration.

**6.4 Data Presentation**

* **Table View:** Display results in a basic table with rows representing each slide or page and columns for each extracted field.
* **Export Capability:** Allow simple export options (e.g., CSV).

**7. Non-functional Requirements**

* **Simplicity:** Focus solely on the core extraction functionality to keep the MVP lean.
* **Performance:** Ensure acceptable response times for processing up to three documents.
* **Security:** Basic security for file uploads and API calls; advanced compliance features can be added in later versions.
* **Usability:** Simple, intuitive interface designed for proof-of-concept evaluation.

**8. MVP Scope**

* **Core Functionality:**
  + Upload up to three documents (PPT/PDF).
  + Process and segment documents.
  + Extract a predefined set of key metadata and medical taxonomy using the LLM API.
  + Display the extracted information in a basic table view.
* **Out of Scope for MVP:**
  + Advanced customization of extraction rules.
  + Scalability to handle large batches of documents.
  + In-depth integration with external dashboards or advanced security measures.

**9. Timeline & Milestones**

* **Week 1:**
  + Define the minimal set of extraction fields.
  + Set up file upload and basic document parsing.
* **Week 2:**
  + Integrate with the multimodal LLM API.
  + Develop extraction logic and table view.
* **Week 3:**
  + Test the MVP with sample documents.
  + Refine UI/UX and address any issues.
* **Week 4:**
  + Final proof-of-concept demo and documentation.

**10. Risks & Mitigations**

* **API Limitations:**
  + *Risk:* The LLM API might return inconsistent results for rich medical content.
  + *Mitigation:* Use a small sample set to adjust extraction parameters and log results for review.
* **Limited Scope Validation:**
  + *Risk:* The MVP may not capture all desired data points.
  + *Mitigation:* Clearly define and communicate that this is a proof-of-concept intended to validate the core extraction process.
* **User Adoption:**
  + *Risk:* Early users may expect more features than provided.
  + *Mitigation:* Set proper expectations through documentation and demos.

This MVP PRD outlines the simplest possible version of your app—a focused, proof-of-concept solution that handles three documents and demonstrates the core capability of extracting rich, medical-specific data using a multimodal LLM API.

|  |  |
| --- | --- |
| **Model ID** | **Model Name** |
| llama-3.3-70b | Llama 3.3 70B |
| llama-3.2-3b | Llama 3.2 3B |
| dolphin-2.9.2-qwen2-72b | Dolphin 2.9.2 Qwen2 72B |
| llama-3.1-405b | Llama 3.1 405B |
| qwen-2.5-coder-32b | Qwen 2.5 Coder 32B |
| deepseek-r1-llama-70b | DeepSeek R1 Llama 70B |
| deepseek-r1-671b | DeepSeek R1 671B |
| qwen-2.5-vl | Qwen 2.5 VL 72B |