LLM Research Paper Analyzer

Time Limit: 3 Hours

Problem Statement

Build an intelligent research assistant that processes academic papers, generates summaries, and identifies research gaps using LLM analysis.

What You'll Build

A web application that uploads research papers, extracts content, provides LLM-generated summaries, and identifies potential research gaps and future work opportunities.

Resources Provided

- **Research Papers**: Download 2-3 PDFs from https://arxiv.org/list/cs.AI/recent
- **Backup Data**: https://jsonplaceholder.typicode.com/posts (use as mock papers if needed)
- **LLM Access**: Use any LLM API (OpenAI, Anthropic, Google, Hugging Face, or local models)

Core Features (Implement 3 out of 4)

1. Paper Upload & Content Extraction

- Upload PDF research papers through web interface
- Extract and clean text content from academic PDFs
- Display paper metadata (title, authors, abstract)

2. LLM-Powered Summarization

- Generate concise summaries of research papers
- Extract key findings and main contributions
- Identify research methodology and results

3. Research Gap Analysis

- Use LLM to identify limitations in current research
- Suggest potential future work directions
- Highlight unexplored areas and opportunities

4. Paper Search & Discovery

- Search through processed papers by keywords
- Filter papers by research area or methodology
- Compare multiple papers for gap analysis

Technical Requirements

- **Frontend**: Paper upload interface, summary display, gap analysis results
- **Backend**: PDF processing, LLM integration, content analysis
- **Database**: SQLite for papers, summaries, and gap analysis
- **LLM Integration**: Advanced prompting for academic content analysis

Success Criteria

- Successfully extract meaningful content from academic PDFs
- Generate accurate, concise summaries of research papers
- Identify realistic research gaps and future work suggestions
- Provide searchable database of processed papers

Time Management Tips

- **Hour 1**: PDF upload, text extraction, basic paper storage
- **Hour 2**: LLM summarization, content analysis, database integration
- **Hour 3**: Research gap detection, search functionality, testing

Focus on LLM prompt engineering - quality analysis requires well-crafted prompts for academic content!