Wakeel

AI Legal Assistant for Islamic, Family and Personal Law

PROJECT PROPOSAL

Amna Raza Khan 24280055 | Amna Shahnawaz 23030024 | Hisham Fawad 24280041 | Yahya Habib 24280001

Problem Statement

Legal assistance is a fundamental right, yet navigating legal jargon and drafting documents in English remains costly and time-consuming. In Pakistan, most people speak Urdu, but legal documents like affidavits are primarily in English, creating a barrier to access. Understanding and applying Islamic/Family Law is important, yet many struggle to find clear, accessible guidance. This gap leaves many without proper legal support, making justice and legal services inaccessible to a large segment of the population.

Solution Overview

We bring you Wakeel, an LLM-powered chatbot designed for seamless legal assistance, document drafting, and AI-driven consultation. Users can chat via Urdu speech to search for case laws, judgments, and legal information, while lawyers can generate legal notices and contracts in English with AI-assisted drafting. Wakeel provides Islamic declarations and rulings in accordance with Pakistan's Muslim Family and Personal Law, ensuring compliance and accuracy.

Wakeel's Objectives

The law bot aims to provide an intuitive and user-friendly platform that enables Urdu speech-based legal search, making case laws and legal references easily accessible while ensuring compliance with Pakistan's family and Islamic legal framework. It will serve as an essential tool for legal research and education, benefiting both professionals and students. By offering AI-assisted legal drafting, the platform will streamline documentation processes, improve efficiency, and save time for lawyers. Additionally, it will promote legal awareness among the general public, bridging the knowledge gap through informational resources.

Wakeel's Goals

- ***Enhance Legal Accessibility*** – Provide a user-friendly digital platform that simplifies access to family and Islamic law in Pakistan. This includes enabling Urdu speech-based search for case laws, legal texts, and judgments, making legal information accessible to the general public, law students, and professionals. The platform will also serve as a valuable educational and research tool, promoting legal awareness and ensuring compliance with Pakistan's legal framework.

- ***Facilitate Efficient Legal Drafting*** – Empower lawyers with AI-assisted tools to draft petitions, contracts, and other legal documents in English efficiently. By leveraging AI, the platform will streamline legal research, eliminate language barrier, improve accuracy in legal documentation, and save time for legal professionals.

AI Project Canvas

Designed for:	Date:	Version:
LLM Course Project	March 09, 2025	V1

Data
What data do you need?

- Legal documents & case laws (family law, inheritance, property disputes)
- Court rulings & precedents
- Legal FAQs from government websites & legal aid organisations
- User queries & chatbot interaction logs (for improving responses)

Skills
Which skills do you need for development?

- ML & NLP (fine-tuning, retrieval models)
- Data Annotation & Text Processing
- Legal Domain Expertise (lawyers, law students)
- Software Development (backend, API integration)
- UI/UX Design for web-based chatbot interface)

Output
Which key metrics are you optimizing for?

- **Response Accuracy:** Precision of legal answers compared to actual case laws and expert validation.
- **Latency & Speed:** Response time for text and voice queries, ensuring real-time interaction.
- **Language Understanding:** Effectiveness in handling Urdu and English legal queries.
- **Retrieval Efficiency:** in fetching relevant case laws and legal documents.
- **Error Rate:** Minimising incorrect or misleading legal advice.

Value Propositions
What is the value added by your project?

- Access to legal knowledge for underserved populations in Urdu
- Reduced burden on lawyers & legal aid organizations
- Accurate & personalized legal guidance in Urdu & English
- 24/7 availability of legal assistance
- Provision to use urdu input to draft legal documents in English

Integration
How will the project be integrated?

- **Streamlit UI** for user interaction, allowing text/voice input and displaying legal responses.
- **Fine-Tuned LLM** for answering general legal questions directly via API or local model.
- **RAG System** with Vector Database (FAISS or Pinecone) to retrieve case laws and legal documents for complex queries.

Customer
Who are the end customers?

- **General Public:** low-literate population seeking legal guidance, especially underserved communities.
- **Legal Aid Organizations & NGOs:** Groups providing legal assistance and advocacy, looking for automated legal assistance for citizens.
- **Lawyers & Law Firms:** Professionals using the chatbot for quick case references and precedents and drafting.
- **Students & Researchers:** Law students and academics needing structured legal data.

Cost *What costs will the porject incur?*

- Compute & Hosting: Free (Google Colab), or Colab Pro ($10) if needed for fine-tuning
- Vector Storage: Free (FAISS) Data Annotation: Free (Team effort)
- Software Development: Free (Handled by project team)
- Legal Consultation: Free (LUMS law students)
- Deployment: Locally, or optional $20–40 if hosted externally.
- Total Estimated Cost: PKR 5,000–10,000 ($20–40), mainly for optional hosting

Revenue Streams *How will the project generate revenue?*

- Freemium model (1 legal advice free/day, premium for unlimited advices)
- Subscription-based access for businesses & NGOs
- Government & NGO funding for legal aid services
- Referral fees for connecting users with lawyers

Wakeel

Technical Feature Space**

1. Data Sources & Processing

- Input Documents: Legal texts, court rulings, FAQs, and case laws in Urdu & English.
- Document Chunking: Using ***langchain.text_splitter*** to divide large legal documents into smaller, retrievable chunks.
- Embedding Model: ***mxbai-embed-large*** generates vector representations of the text.
- Vector Database: ***ChromaDB*** stores embeddings and supports efficient similarity-based retrieval.
- Query Embeddings: Incoming user queries are also embedded using ***mxbai-embed-large*** for matching against stored documents.

2. Model Training & Fine-Tuning

- Base LLM: ***deepseek-r1***, a pre-trained language model for general text understanding.
- Fine-Tuning: ***LoRA-based fine-tuning*** on ***deepseek-r1*** using legal instruction-response pairs in Google Colab (***T4 GPU***).
- Data for Fine-Tuning: Includes curated legal conversations, FAQs, and case-based responses from legal professionals.

3. Retrieval-Augmented Generation (RAG) Pipeline

- ***Retrieval Mechanism:***
  - User queries are transformed into embeddings using ***mxbai-embed-large***.
  - Top-K relevant chunks are retrieved from ChromaDB.
  - The retrieved chunks are appended to the user query before being fed into the fine-tuned LLM.
- ***Generation Mechanism:***
  - The fine-tuned ***deepseek-r1_finetuned*** model generates responses based on both query and retrieved legal context.

4. User Interaction Interface

- ***Frontend Framework: Streamlit*** provides a web-based interface for chatbot interaction.
- ***Text Input Processing:***
  - Users can enter legal queries via text.
  - The input is processed and converted into embeddings for retrieval.
- ***Speech Input Processing:***
  - Speech-to-text conversion via ***whisper_base*** enables voice queries.
  - Transcribed text is sent to the retrieval system.
- ***Response Handling:***
  - The chatbot generates personalized legal responses.
  - Users receive responses via text output or optional voice synthesis.

5. Deployment & Scalability

- ***Hosting Options:***
  - Local deployment for free access.
  - Optional cloud hosting (~$20-$40) for wider scalability.
- ***Vector Search Optimization:*** ChromaDB allows efficient scaling for larger legal document retrieval.

**The technical specifications are subject to refinement based on testing and evaluation of available options. This document serves as an initial framework.

Wakeel

SOLUTION OVERVIEW

Wakeel is an AI-powered legal assistant that combines fine-tuning and Retrieval-Augmented Generation (RAG) to provide legal drafting and case law search capabilities in Pakistan's family and Islamic law domain.

- **Fine-Tuned Legal Drafting:** The core LLM (deepseek-r1) is fine-tuned using instruction-response pairs to specialize in drafting court petitions, legal notices, affidavits, and contracts. This ensures that legal documents generated by Wakeel follow the proper structure, tone, and legal requirements.

- **RAG for Case Law Search:** Wakeel integrates Retrieval-Augmented Generation (RAG) to search through legal judgments, case laws, and precedents stored in a vector database (ChromaDB). This allows users to retrieve relevant legal references instantly.

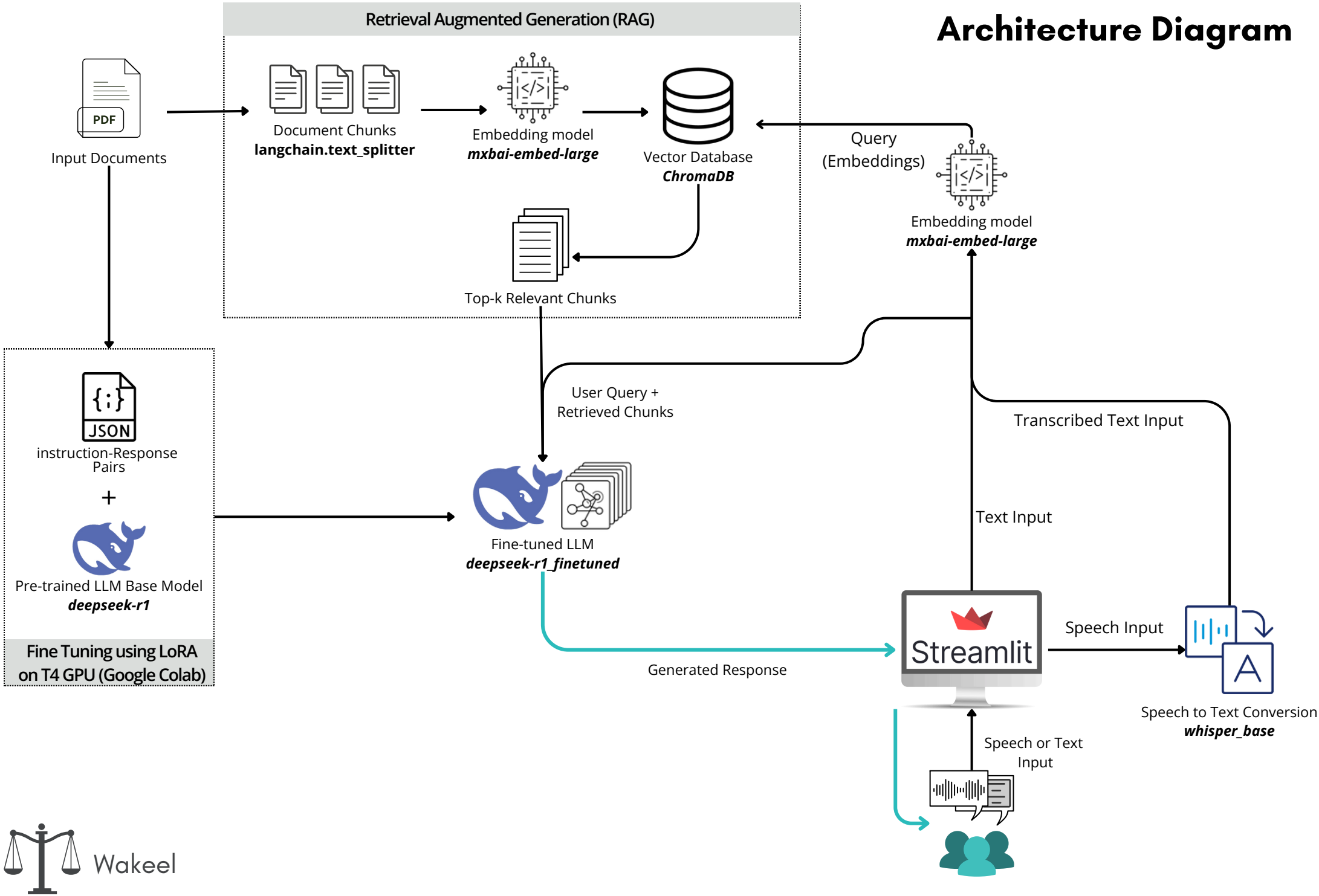
- **Urdu Speech & Text Input:** Users can interact with Wakeel using Urdu speech or text, making legal information more accessible. Whisper-base converts speech to text, enabling seamless voice-based queries.

- **Streamlit-Based User Interface:** The chatbot interface is built on Streamlit, providing a simple and efficient way for users to ask legal questions, search for cases, and generate documents.

- ***In Future*: Lawyer-Specific Kiosk Services:** Wakeel will also support dedicated kiosk-based solutions for lawyers, enhancing workflow automation and document preparation.
-

By combining fine-tuning for legal drafting and RAG for case law retrieval, Wakeel ensures accuracy, efficiency, and accessibility, bridging the legal knowledge gap for professionals, students, and the general public.

Architecture Diagram



TENTATIVE TIMELINE

Phase 1: Problem Understanding (Weeks 1-3)

Tasks:

- Define the scope of legal chatbot (focus on family law, inheritance, property disputes).
- Research underserved communities' legal challenges in Pakistan.
- Identify key stakeholders (legal aid organizations, NGOs, law students).

Phase 2: Product Ideation (Weeks 4-5)

Tasks:

- Finalize chatbot use cases and value propositions.
- Define key features (text/voice input, retrieval-augmented responses).
- Outline UI/UX for the chatbot (basic interface design).
- Identify technical requirements (LLM fine-tuning, RAG with FAISS/Pinecone).

Deliverables:

Team Submission (Due: Feb 23, 2025).
Project Proposal Report (Due: Mar 9, 2025) – includes business problem, data collection, technical features, architecture diagram.

Phase 3: AI Prototyping (Weeks 6-9)

Tasks:

- Collect and preprocess legal data (court rulings, precedents, FAQs).
- Develop LLM-based legal chatbot prototype:
- Fine-tune an LLM for Urdu/English legal queries.
- Implement RAG with a vector database (FAISS/Pinecone).
- Develop basic Streamlit UI for user testing.
- Conduct internal testing (legal accuracy, latency, language understanding).
- Gather initial user feedback from law students/legal professionals.

Deliverables:

Design Review Meeting (Mar 10-23, 2025).
Prototype Demo (Week 9-10) (Apr 1-13, 2025).

Phase 4: Evaluation & Improvements (Weeks 10-14)

Tasks:

- Evaluate chatbot's accuracy, response speed, and retrieval efficiency.
- Improve chatbot based on feedback.
- Optimise retrieval and error handling.
- Enhance UI/UX.
- Deploy chatbot locally or host externally (if budget allows).
- Explore scalability (adding more legal domains).

Deliverables:

- *Prototype Demo (Weeks 10-11).*
- *Final Group Presentation & Report Submission (Last 2 Lectures).*
- *Individual Viva Exams.*

Wakeel

METHODOLOGY FEASIBILITY

Data Collection

- Public Court Archives: Legal documents and case laws will be sourced from publicly accessible archives like Islamabad High Court.
- Expert Annotations: Law students will curate, validate, and refine legal responses to ensure accuracy.
- User Queries & Chat Logs: Real-world queries will be collected to improve the chatbot's contextual understanding and adaptability.

Data Processing

- Text Normalization: Legal texts will be cleaned, structured, and formatted to remove inconsistencies.
- Chunking & Embedding: Large legal documents will be divided into smaller, retrievable sections using langchain.text_splitter, and embeddings will be generated with mxbai-embed-large for efficient retrieval.
- Vector Database Storage: Processed embeddings will be stored in ChromaDB, enabling fast and precise information retrieval.
- Fine-Tuning for Legal Drafting: A dataset of legal instruction-response pairs will be compiled, and the LLM (deepseek-r1) will be fine-tuned using LoRA-based training on Google Colab to enhance legal drafting capabilities.

Evaluation & Feasibility Analysis

- Response Accuracy: LUMS Law Students will validate chatbot-generated responses to ensure reliability.
- Performance & Latency: System efficiency and response time will be tested for real-time interactions.
- Language Processing: The chatbot's ability to handle Urdu and English legal queries will be assessed.
- User Feedback & Iteration: Insights from law students, lawyers, and general users will guide ongoing refinements.

This approach ensures Wakeel delivers accurate, efficient, and accessible legal assistance while continuously improving through real-world feedback.