



LUMIN

A CORPORATE LEGAL LLM

MOTIVATION

Corporate legal departments and law firms are under increasing pressure to deliver faster, more accurate legal services while managing growing volumes of complex documentation and regulatory changes. However, much of their workflow remains manual, repetitive, and inefficient, leading to delays, higher costs, and greater chances of human error.

Tasks like summarizing legal documents, checking compliance, drafting clauses, and responding to legal queries are time-intensive but follow patterns that are ideal for intelligent automation. With advancements in AI and NLP, there is a strong opportunity to transform how legal professionals work.

PROBLEM STATEMENT

- Corporate law-related tasks rely heavily on manual, time-consuming labour like document summarization, compliance checking, and contract drafting.
- These repetitive processes increase turnaround time, operational costs, and the risk of human error.
- Legal professionals often lack access to tools that streamline clause generation or verify regulatory corporate compliance efficiently.
- There is also no centralized platform for posing legal queries specific to corporate laws and receiving context-aware responses.
- The absence of automation limits scalability and agility in legal operations.

EXISTING SYSTEMS

- InLegalLLaMA (2024)
 - Focuses on Indian legal texts broadly (judgment summarization, petition drafting).
 - Lacks specialization in corporate laws like the Companies Act.
 - No rule-based compliance checking or explainability.
- LawLLM (2024)
 - Trained on U.S. legal data for tasks like case retrieval and legal judgment prediction.
 - Not suitable for Indian jurisdiction or corporate-specific legal reasoning.
 - High resource demand and not open-source.
- Lawyer LLaMA (2023)
 - Two-step training on Chinese legal corpora with retrieval augmentation.
 - Designed for Q&A and exam-type reasoning, not practical contract analysis.
 - Limited domain focus .

FINETUNING

Finetuned weights link -

<https://huggingface.co/its-a-meee/mistral-7b-4bit-merged/tree/main>

- Base Model: unsloth/mistral-7b-instruct-v0.3-bnb-4bit
- Finetuning Strategy: LoRA Adapters
- Fine-tuned on cleaned raw text and QA pairs from the Companies Act, 2013 to ensure deep contextual understanding.

1. CHATBOT

- There is a lack of an engine to answer user queries specific to corporate laws - specifically the Companies Act, 2013, with high accuracy and context awareness.
- Input: User query
- Working: Input → Finetuned Model → RAG(for specific sections) → Output
- Output: Returns contextually grounded, legal-aware text responses.
- Evaluation metrics: BERTScore, Cosine Similarity

2. ROLE BASED SUMMARY

- In the corporate world, different individuals require summaries for different purposes, so providing a summary that aligns with their goal is more useful than a general summary, which may not provide the context required for the task at hand.
- Roles available: Legal expert, Corporate Executive, Law Student, Compliance Officer
- Input: PDF Document
- Working: Input document → chunks → chunk summary using CoT, Zero-shot, Role-based prompting → overall summary
- Output: Summary
- Evaluation metric: BERTScore

3. CUSTOM CONTRACT

- Custom contracts are useful to personalize contracts to cater to user requirements, especially in this dynamic world. Writing a contract from a template and then rectifying it multiple times for particular sections required by the user is time-consuming; hence, a custom clause that generates the sections as per the user's requirements is time-efficient.
- Sections of Contract: Preamble, Terms and Conditions, Conclusion
- Input: Contract type, Clause requirements for sections
- Working: Input → Uses COT Strategy for drafting clauses → Writes to file → Output file
- Output: Contract file
- Evaluation metric: Cosine similarity

4. CUSTOM CLAUSE

- Custom clauses are required by users who have a predefined template, such as for **contract renewal, contract negotiations** in the business world. These do not require a full contract but rather particular clauses that cater to the business goals of the companies to be added to the existing contract.
- Input: Clause requirement
- Working: Input → Uses COT, Few-shot, Role-based prompting for clauses → Outputs clause section
- Output: Clause and subclauses
- Evaluation metric: Cosine similarity

5. CLAUSE COMPLIANCE

- In legal workflows, verifying individual clauses against regulatory frameworks is essential. Clause level compliance checking allows precision level scrutiny instead of relying on full document interpretations, which may miss localized violations.
- Legal teams, startups, and contract drafters can use this tool to validate specific clauses in real time. It helps in drafting compliant clauses from the start or reviewing problematic ones without going through the entire contract.
- Input: A single contract clause entered via text box along with the section number
- Working: Input clause → semantic retrieval of relevant Companies Act sections → LLM-based compliance analysis → verdict + explanation
- Output: Clause-wise compliance verdict
- Legal-style reasoning explaining the result
- Evaluation Metrics: Manual validation against expert-labeled clauses

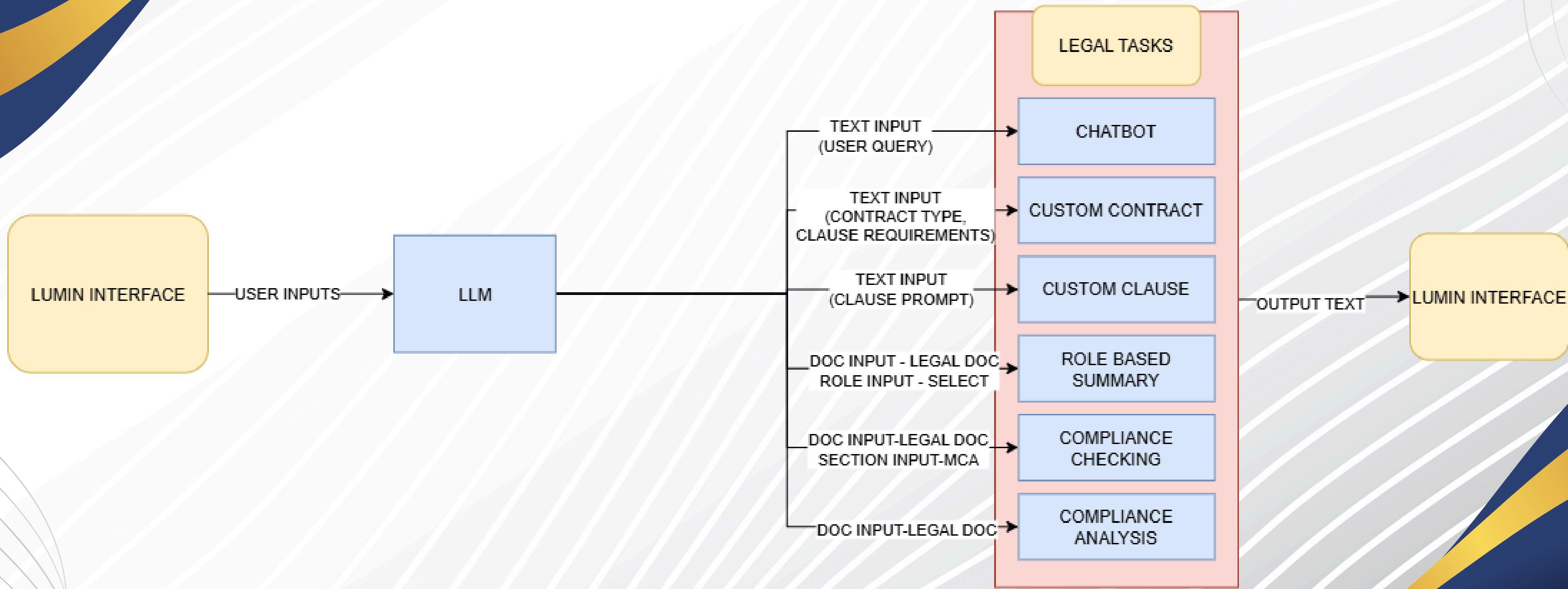
6. DOCUMENT COMPLIANCE

- Contracts are legally binding documents that must reflect compliance not just at the clause level but in their overall structure, language, and intent. A document-wide check ensures that the contract is internally consistent, legally aligned, and does not introduce regulatory risks through omissions or conflicting terms.
- This functionality aids legal professionals and compliance officers in reviewing entire contracts. It supports legal due diligence, contract approvals, and risk assessments with a complete compliance overview.
- Input: Full contract in PDF format
- Working: Input document → chunking into sections →
 - Each chunk: semantic retrieval + LLM analysis →
 - Fallback: keyword-based retrieval if semantic context is weak →
 - Generate chunk-level summaries → Build an overall compliance report
- Output: Full contract compliance report with section wise legal analysis and consolidated summary of risks and alignment

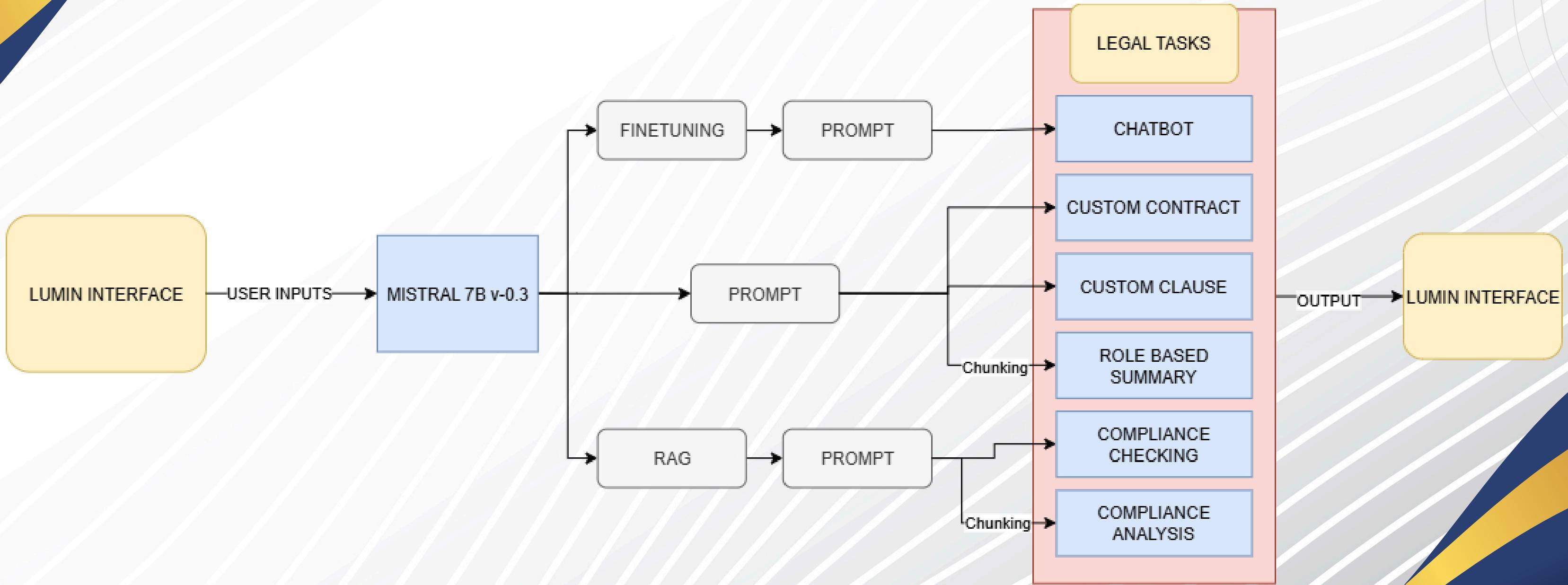
NOVELTY AND UNIQUENESS

- Mistral 7B fine-tuned on the Companies Act, 2013.
- Legal query chatbot
- Role-based Summary
 - Legal experts
 - Corporate executives
 - Compliance officers
 - Law students
- Custom generation
 - Clauses
 - Contracts
- Clause-specific compliance
- Compliance tagging(specific to the Companies Act,2013)

FUNCTIONALITY DIAGRAM



ARCHITECTURE DIAGRAM



LLM TECHNOLOGIES

- Vanilla RAG (Retrieval-Augmented Generation):
 - Used to enhance LLM responses with up-to-date and domain-specific legal information by retrieving relevant context from a custom knowledge base.
- Fine-tuning:
 - Applied to adapt pre-trained models to legal-specific language, improving accuracy and relevance in legal document understanding and generation.
- Prompt Engineering Techniques:
 - Instruction Prompting
 - Role-based Prompting
 - Chain of Thought Prompting
 - Zero-shot Prompting
 - Few-shot Prompting

LLM EVALUATIONS

- Chatbot - ROUGE, BLEU
- Compliance checker and analysis - Manual evaluation by comparison and validation
- Custom clause - Cosine similarity of user input and generated output
- Role-based summary - BERTScore of input document and generated summary

Custom Contract has same evaluation as custom clause

THANK YOU