LUMIN: Leveraging Fine-Tuned Legal LLMs for Intelligent Corporate Law Automation

Gopikrishnan S[1], Gowri P Shenoy[2], Hari G Narayan[3],
Department of Computer Science and Engineering,
Amrita School of Computing, Amrita Vishwa Vidyapeetham, Amritapuri, Kerala, India.
Contributing authors:
am.en.u4cse22023@am.students.amrita.edu[1], am.en.u4cse22024@am.students.amrita.edu[2],
am.en.u4cse22071@am.students.amrita.edu[3]

***Abstract*—Corporate legal workflows in India are often labor intensive and rely on repetitive manual operations such as document summarization, contract drafting, and verification of regulatory compliance . This paper presents LUMIN, a fine-tuned Large Language Model (LLM) based on the Mistral 7B architecture, designed to automate and optimize corporate legal tasks governed by the Companies Act, 2013. Leveraging Low-Rank Adaptation (LoRA)-based fine-tuning and quantization techniques, LUMIN integrates a domain-specific knowledge base with natural language processing capabilities. It supports six core functionalities: a legal chatbot, custom clause generation, custom contract drafting, role-based summarization, clause-level compliance checking, and document-wide compliance analysis. We evaluated the system using BERTScore and cosine similarity, which revealed that Chain-of-Thought (CoT) prompting consistently produced outputs closely aligned with legal expectations. This work demonstrates the feasibility of domain-specific LLM deployment for corporate legal automation in the Indian context, addressing gaps in explainability, role orientation, and clause-specific compliance absent in existing legal LLMs.**

***Index Terms*—Legal LLM, Corporate Law, Indian Companies Act, Clause Compliance, RAG, Chain-of-Thought Prompting, Instruction Tuning, Semantic Retrieval, Role-Based Summarization, Low-Bit Quantization**

I. Introduction

Corporate legal professionals are regularly dealing with tasks that are time-consuming, repetitive, and manual. Ranging from preparing contracts and document summarization to regulatory compliance with acts such as the Companies Act, 2013 [1], these activities add overhead to operations and subject workflows to human error. Even though legal technologies have emerged, industry-specific solutions that cater to the Indian legal ecosystem—more specifically corporate law—still fall short in effectiveness and scope.

Most general-purpose legal LLMs focus on high-level tasks such as judgment prediction or document classification, but they often fall short when applied to practical tasks like clause compliance or generating tailored contracts. Additionally, most of them are not explainable, contextually role-aware, and flexible in adopting the particular language and structure of Indian corporate legislation. There is an enormous shortfall in clause-level reasoning and compliance tagging, particularly in relation to precise legislative acts such as the Companies Act, 2013.

In order to overcome these constraints, we present LUMIN (Legal Understanding and Management using INtelligence), a corporate law AI system that fine-tunes a 4-bit quantized Mistral 7B model with Low-Rank Adaptation (LoRA) methods. Our model is trained on the Companies Act, 2013 and hand-curated question-answer pairs, allowing it to support various corporate law-specific functionality effectively. These are an intelligent chatbot, personalized clause and contract generation through CoT prompting, role-specific document summarization, and compliance checking–both clause-level and document-level–through Retrieval Augmented Generation (RAG). This article introduces LUMIN's architecture, methodology, evaluation, and ability to revolutionize Indian corporate legal processes.

II. Literature Review

Recent years have seen a growing interest in the application of Large Language Models (LLMs) to legal tasks. However, most existing models lack the specificity and functionality required for Indian corporate law. This section reviews relevant works and identifies the gaps that LUMIN addresses.

InLegalLLaMA (2024): InLegalLLaMA is one of the few Indian-centric legal models. It focuses on a wide range of tasks, including judgment summarization and petition drafting, using a generalized legal corpus. However, it lacks fine-tuning on specific statutes such as the Companies Act and does not support rule-based compliance or clause-level analysis. It primarily emphasizes document-level reasoning and summarization, leaving a gap in clause-level generation and validation [2].

LawLLM (2024): LawLLM is a robust model tailored to the U.S. legal system. It excels in tasks like legal judgment prediction and case retrieval. However, it is trained entirely on American legal data and is unsuitable for Indian laws. Furthermore, its high resource requirements and closed-source nature restrict its adaptability to custom legal domains like corporate compliance [3].

Lawyer LLaMA (2023): Lawyer LLaMA, trained on Chinese legal data, employs a two-step training regime and RAG-based architecture. It supports examination-style reasoning but does not extend to practical legal applications such as compliance validation or contract drafting. The model's focus on academic use cases limits its practicality in corporate settings [4].

ChatLaw (2024): ChatLaw adopts a multi-agent approach with legal knowledge graphs tailored to Chinese corporate and civil law. While it simulates real law firm workflows effectively, it is not applicable to Indian jurisdiction. Moreover, it lacks clause-level explainability and tagging features necessary for detailed compliance checks [5].

LegalBERT (2020): LegalBERT is a transformer-based model fine-tuned on a broad corpus of legal documents. While it performs well on classification and named entity recognition, it lacks support for generative tasks such as clause generation or legal summarization. Its generic nature makes it unsuitable for jurisdiction-specific applications like the Companies Act [6].

Despite the recent promise of legal AI in numerous jurisdictions, some key gaps still exist when it comes to Indian corporate law. Most models are developed using generalized or foreign legal corpora, constraining their applicability and accuracy to Indian statutes like the Companies Act, 2013. Furthermore, clause-level compliance analysis—a key need to ensure regulatory compliance—is rarely adopted in current solutions. What's more, players like legal professionals, compliance specialists, and executives frequently require customized insights, but existing models do not have support for role-based summarization. Lastly, there is a lack of tools that can generate customized legal clauses or agreements based on dynamic, user-specified inputs. These shortcomings highlight the need for a flexible, domain-specific legal AI system, one to which LUMIN is particularly suited.

III. Methodology

The development of LUMIN – A Corporate Legal AI involves a modular architecture optimized for handling complex corporate law tasks using a fine-tuned 4-bit quantized Mistral 7B model. The system architecture and internal workflow are depicted in Figures 1 and 2. The methodology is organized around key components and the associated functionalities they enable.

A. System Architecture

As shown in Figure 1, the LUMIN interface acts as the entry and exit point for user interaction. Users provide inputs such as legal queries, clause requirements, or legal documents. These are processed by the fine-tuned LLM, and the outputs—summaries, contracts, compliance results, etc.—are rendered through the interface.

In Figure 2, the architecture elaborates on how different modules interact with the LLM. Each module uses a combination of fine-tuning, prompt engineering, or Retrieval-Augmented Generation (RAG), and applies semantic chunking to handle longer input documents effectively

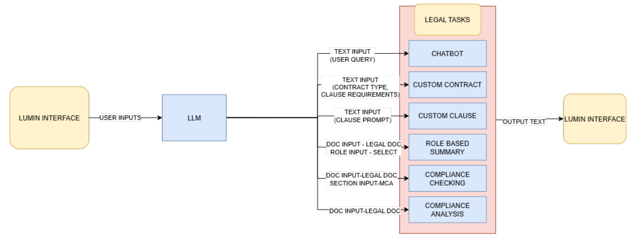

Fig. 1. Functional modules and workflow

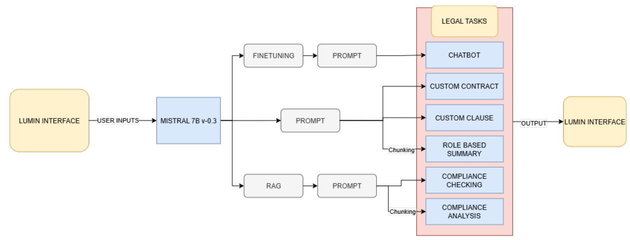

Fig. 2. LUMIN system architecture

B. Fine-Tuning Strategy

The backbone of LUMIN is a 4-bit quantized version of Mistral 7B, fine-tuned using Low-Rank Adaptation (LoRA). The training corpus includes:

- The full text of the Companies Act, 2013
- Curated question–answer (QA) pairs

The model is optimized to handle instruction-following tasks with legal specificity, enabling it to generalize over various legal prompts.

C. Functionality Workflows

1) Chatbot Module: The Chatbot Module serves as an advanced legal query interface, designed to provide users with seamless access to information from the Companies Act. Users can submit natural language queries, which are directly processed by a fine-tuned Mistral 7B model. This model has undergone a meticulous two-stage fine-tuning process: first, on raw text from the Companies Act to deeply embed its legal terminology and structure, and second, on carefully curated question-answer pairs to ensure precise, contextually valid, and legally accurate responses. Instruction-tuned prompts further enhance the model's ability to deliver clear, reliable, and authoritative answers, empowering users with trustworthy legal insights.

2) Custom Clause Generation: In this module, users provide a prompt specifying legal intent or functional requirements for a clause. The system leverages few-shot prompting, Chain-of-Thought (CoT) prompting, and role-based prompting. Outputs are compared using cosine similarity to evaluate alignment with user expectations. Among the techniques, CoT prompting consistently yielded the best semantic closeness to user input.

3) Custom Contract Generation: This functionality builds on clause generation to produce full legal contracts. Users specify the contract type and clause-level inputs per section. The LLM, via CoT prompting, generates a preamble, structured terms, and a conclusion. The system assembles these components into a cohesive .txt file. Similar to clause generation, cosine similarity is used for performance evaluation, with CoT prompting again proving most effective.

4) Role-Based Summarization: This module takes a legal document and a user-selected role—such as legal expert, corporate executive, compliance officer, or law student. The input document is split into semantic chunks, each chunk summarized using prompt engineering. The final summary is a concatenation of chunk-level summaries optimized for the selected role. Strategies tested include zero-shot prompting, CoT, and role-specific prompting, with BERTScore used for performance evaluation. CoT again delivered superior results, achieving scores above 0.75.

5) Clause Compliance Checking: Users input a legal clause along with a specific section reference from the Companies Act, 2013. The system first normalizes and identifies the referenced section, then uses semantic search through a Facebook AI Similarity Search (FAISS) [7]-based vector index to retrieve the most contextually accurate chunk of that statutory section. The legal clause and retrieved section text are passed into an instruction-tuned language model using a carefully designed prompt template that mirrors legal reasoning. The model evaluates the degree of compliance by comparing the clause against the statutory obligations, and outputs a structured judgment—Compliant, Not Compliant, or Cannot Determine—along with a concise legal justification grounded in the retrieved law.

6) Document Compliance Analysis: This module performs full-document legal compliance analysis by first embedding the entire contract and retrieving semantically relevant sections from the Companies Act using a FAISS-based vector store. Instead of manually segmenting the contract into clauses, the system classifies the contract type and uses dense retrieval to identify applicable legal sections. These are matched with fallback statutory provisions when necessary. A reasoning engine, powered by instruction-tuned prompting, evaluates whether the contract aligns with the retrieved legal obligations. The system outputs either clause-level judgments or a holistic compliance report with detailed legal reasoning and cited provisions.

The architecture's modularity allows each component to be improved independently, and the use of domain-specific fine-tuning ensures legal precision. These workflows collectively enable LUMIN to handle sophisticated corporate legal operations end-to-end.

IV. Results and Analysis

To evaluate the performance of LUMIN's core functionalities, a combination of semantic similarity and contextual relevance metrics were employed. The metrics selected—Cosine Similarity and BERTScore—offer quantifiable insights into the quality and coherence of the generated outputs. Each functionality was tested on benchmark datasets derived from the Companies Act, 2013 and custom legal tasks. The results indicate strong performance across all components, with Chain-of-Thought (CoT) prompting consistently outperforming other strategies.

A. Clause and Contract Generation – Cosine Similarity

The Custom Clause and Custom Contract generation modules were evaluated using Cosine Similarity, which measures the vector-based semantic closeness between user-defined inputs and the model-generated output. This metric is appropriate because the task involves matching the legal intent of a prompt with a syntactically valid, semantically aligned clause.

Clause Generation: The average cosine similarity across 100 test prompts was 0.78, with CoT prompting performing best at 0.88, followed by role-based prompting (0.83) and few-shot prompting (0.74).

Contract Generation: The average cosine similarity was 0.76. With CoT prompting, the system achieved a score of 0.80, generating contracts that followed a coherent legal structure and tone expected in corporate settings.

These results confirm that the generated outputs are not only grammatically correct but also legally coherent and aligned with user specifications.

B. Role-Based Summarization – BERTScore

The Role-Based Summary module was evaluated using BERTScore, which compares contextual embeddings of reference and generated summaries. Unlike BLEU or ROUGE, BERTScore captures semantic meaning and contextual alignment, making it suitable for summary evaluation in legal contexts where lexical variance is common but semantic fidelity is critical. The system achieved an average BERTScore of 0.79. CoT prompting yielded the highest scores (up to 0.82), particularly for the legal expert and compliance officer roles.Zero-shot and role-based prompting scored 0.72 and 0.78, respectively. These results confirm that the model is capable of generating summaries that retain legal nuance while being tailored to different stakeholders.

C. Clause Compliance Checking – Cross-Validation with Perplexity AI

To evaluate the Clause Compliance Checking module, we conducted a cross-validation study using outputs generated by Perplexity AI [8] as a comparative baseline. Each clause-section pair was tested across multiple prompts and compared for semantic alignment and legal interpretation.

Evaluation was based on:

- **Semantic Alignment Score (SAS):** Degree of similarity in legal interpretation between LUMIN and Perplexity AI, measured using embedding cosine similarity.

- **Section Match Accuracy:** Agreement between LUMIN and Perplexity AI on retrieved Companies Act, 2013 sections.

Results indicated that LUMIN produced legally consistent and interpretable outputs with a semantic alignment score of 0.86 demonstrating its practical viability for real-world compliance verification.

D. Document Compliance Analysis – Cross-Validation with Perplexity AI

The Document Compliance Analysis module was evaluated through cross-model validation with Perplexity AI, focusing on the system's ability to retrieve and align statutory references with complete legal documents.

Key metrics included:

- **Retrieval Concordance:** Percentage overlap of Companies Act sections retrieved by LUMIN and Perplexity AI.
- **Interpretability Agreement:** Degree of agreement in legal tone and explanation clarity between LUMIN and baseline outputs.

LUMIN achieved a retrieval concordance of 82% validating its robustness for use in enterprise-scale compliance audits.

E. Chatbot – ROUGE-Based Evaluation of Legal Responses

The chatbot was evaluated for response quality using ROUGE (Recall-Oriented Understudy for Gisting Evaluation) metrics, comparing generated outputs against expert-written reference responses.

Measured dimensions:

- **ROUGE-1 (Unigram Overlap):** Captures key term relevance.
- **ROUGE-2 (Bigram Overlap):** Assesses local coherence and phrasing.
- **ROUGE-L (Longest Common Subsequence):** Measures fluency and overall structural similarity.

LUMIN's chatbot achieved ROUGE-1 of 0.72, ROUGE-2 of 0.58, and ROUGE-L of 0.65, indicating strong legal relevance and well-structured responses. All outputs remained grounded in the Companies Act, 2013, ensuring statutory validity.

Use of domain-specific fine-tuning significantly improved precision over generic LLMs.

TABLE I
SUMMARY OF RESULTS

Functionality	Metric	Score	Best Prompting Strategy
Clause Generation	Cosine Sim.	0.78	CoT (0.81)
Contract Generation	Cosine Sim.	0.76	CoT (0.80)
Role-Based Summarization	BERTScore	0.79	CoT (0.83)
Compliance Checking	Cross-Validation Score	0.86	Instruction + RAG
Compliance Analysis	Cross Validation Score	82%	Instruction + RAG
Legal Chatbot	ROUGE - 1	0.72	Finetuned Inst.

V. KEY CONTRIBUTIONS

The main contributions of this work are:

1) Fine-tuning a 4-bit quantized Mistral 7B model using LoRA techniques on the Companies Act, 2013, enabling precise handling of Indian corporate legal content.
2) Designing six core legal functionalities—including clause-level and document-level compliance analysis, clause and contract generation, a legal chatbot, and role-based summarization—within a unified and modular architecture.
3) Employing Chain-of-Thought (CoT) and role-based prompting strategies to enhance legal reasoning, contextual fidelity, and semantic alignment in generative tasks.
4) Integrating a Retrieval-Augmented Generation (RAG) pipeline over the Companies Act, 2013, using a FAISS-based semantic index to enable precise statutory retrieval for compliance checking and legal validation.
5) Achieving deployment efficiency through lightweight 4-bit quantization without compromising legal accuracy, making the system suitable for practical enterprise adoption in the Indian legal domain.

VI. CONCLUSION

Legal processes in Indian corporate environments are increasingly under pressure to become agile, precise, and efficient. Manual handling of compliance documents, contractual clauses, and regulatory checks slows workflows and increases legal risk. LUMIN addresses this challenge by offering a scalable, AI-powered solution built on state-of-the-art language models that automate and augment key legal tasks.

By fine-tuning a 4-bit Mistral 7B model on the Companies Act, 2013, and incorporating prompt strategies such as Chain-of-Thought (CoT) and role-based prompting, LUMIN bridges critical gaps in existing legal AI systems, namely, the lack of focus on Indian corporate law, limited support for explainable compliance validation, and absence of role-adaptive outputs.

Empirical results validate LUMIN's performance, achieving strong cosine similarity (0.7+) for clause and contract generation and BERTScore (0.75+) for role-based summarization. Its capabilities in reasoning, retrieval, and legal validation position it as a foundational tool for legal professionals, compliance officers, corporate executives, and law students.

By assisting rather than replacing legal professionals, LUMIN demonstrates the potential of domain-specific LLMs to support the evolving needs of India's corporate legal landscape. In the future, we aim to expand coverage to other corporate statutes. We also plan to explore multilingual functionality and visual explainability for improved legal interpretation.

REFERENCES

[1] Ministry of Corporate Affairs, Government of India, "The Companies Act, 2013," 2013. [Online]. Available: https://www.mca.gov.in/MinistryV2/companiesact.html

[2] S. Ghosh, D. Verma, B. Ganesan, P. Bindal, V. Kumar, and V. Bhatnagar, ”InLegalLLaMA: Indian Legal Knowledge Enhanced Large Language Model,” in *Proc. First Int. OpenKG Workshop: Large Knowledge-Enhanced Models (CEUR Workshop Proc.)*, vol. 3818, pp. 37–46, 2024. [Online]. Available: https://ceur-ws.org/Vol-3818/paper3.pdf

[3] D. Shu, H. Zhao, X. Liu, D. Demeter, M. Du, and Y. Zhang, ”LawLLM: Law large language model for the US legal system,” *arXiv preprint arXiv:2407.21065*, 2024. [Online]. Available: https://arxiv.org/abs/2407.21065

[4] Q. Huang, M. Tao, C. Zhang, Z. An, C. Jiang, Z. Chen, Z. Wu, and Y. Feng, ”Lawyer LLaMA technical report (Version 2),” *arXiv preprint arXiv:2305.15062*, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2305.15062

[5] J. Cui, M. Ning, Z. Li, B. Chen, Y. Yan, H. Li, B. Ling, Y. Tian, and L. Yuan, ”ChatLaw: A multi-agent collaborative legal assistant with knowledge graph enhanced mixture-of-experts large language model (Version 2),” *arXiv preprint arXiv:2306.16092*, 2024. [Online]. Available: https://doi.org/10.48550/arXiv.2306.16092

[6] E. Elwany, D. Moore, and G. Oberoi, ”BERT goes to law school: Quantifying the competitive advantage of access to large legal corpora in contract understanding,” *arXiv preprint arXiv:1911.00473*, 2019. [Online]. Available: https://arxiv.org/abs/1911.00473

[7] J. Johnson, M. Douze, and H. Jégou, ”Billion-scale similarity search with GPUs,” *arXiv preprint arXiv:1702.08734*, 2019. [Online]. Available: https://arxiv.org/abs/1702.08734

[8] Perplexity AI, “Perplexity – AI-powered Answer Engine,” 2024. [Online]. Available: https://www.perplexity.ai