Tittle: Generative AI for Jurisprudential Analysis: Automated Legal Reasoning and Scenario-Specific Guidance in Indian Law with a GPT-3 and LawBERT Hybrid Model.

ABSTRACT

The Indian legal system, rich with diverse case laws, acts, and judgments, poses a significant challenge for efficient legal research and document drafting. This project proposes a hybrid AI-based legal research system that integrates LawBERT for legal document retrieval and GPT-based/Gemma models for intelligent legal text generation. Through an extensive data pre-processing phase involving web scraping, OCR extraction, named entity recognition (NER), and semantic text similarity (STS) analysis, the system ensures high-quality legal data handling. Fine-tuning is performed on Indian-specific legal datasets like InLegalBERT and NyayaAnumana to enhance domain adaptation. Techniques such as Retrieval-Augmented Generation (RAG), model optimization via ONNX, and explainable AI methods like SHAP and LIME are incorporated to build a transparent, fast, and trustworthy system. The proposed system aims to assist legal professionals, students, and citizens by automating case law retrieval, contract drafting, risk prediction, and providing scenario-based legal assistance. Evaluation on Indian legal benchmarks ensures that the solution is effective, accurate, and suitable for real-world legal applications.

Keywords: Indian Legal System, Legal Research, LawBERT, GPT/Gemma, Legal Document Retrieval, Legal Text Generation, Named Entity Recognition (NER), Semantic Text Similarity (STS), Retrieval-Augmented Generation (RAG), InLegalBERT, NyayaAnumana, Optical Character Recognition (OCR), Explainable AI (XAI), SHAP, LIME, Contract Drafting, Case Law Automation, Legal Risk Prediction, Legal Assistance, ONNX Optimization.