



Predicting Judgment decisions using LLM

DS504 Natural Language Processing

Aditya Prakash 12240040

Akshat Kumar 12240080

Ayush Kumar Mishra 12240340

Shivam 12241710

Problem Statement: *Legal systems globally face increasing challenges due to delays caused by high case backlogs. In India, this issue is critical, with over 30 million cases pending and a judge-to-population ratio significantly lower than recommended. Courts like New Delhi's High Court have estimated that the existing backlog could take centuries to resolve without interventions.*

*Advancements in Natural Language Processing (NLP) provide an opportunity to address these challenges. Large Language Models (LLMs) trained on legal data can assist in automating aspects of legal processes, particularly **legal judgment prediction**. This approach not only aids efficiency but also ensures transparency by providing detailed explanations for the predicted outcomes.*

Github Repo : [Predicting-Judgment-decisions-using-LLM](#)

Dataset

The Indian judiciary, characterized by complex legal proceedings and a diverse set of cases, is ideal for deploying NLP solutions. While previous datasets and models have focused on specific tasks like document classification or summarization, they lack comprehensive tools for judgment prediction combined with explanation. The **PredEx dataset**, introduced in the referenced paper, addresses this gap by annotating over 15,000 Indian legal cases with outcomes and rationales <https://doi.org/10.48550/arXiv.2406.04136>.

Objective

This project aims to fine-tune a state-of-the-art LLM for the task of **legal judgment prediction** and explanation in the Indian judicial context. By leveraging datasets like PredEx, the model will:

- Predict legal case outcomes based on case facts and arguments.
- Provide clear, evidence-based explanations for the predictions.

- Give the decision on the basis of facts of case.

Methodology

1. Dataset Preparation

- Using PredEx, which includes case documents and annotations for both outcomes and explanations.
- We preprocessed the Dataset to handle document length, token limits, and noise reduction.

2. Model Training

- Fine-tuning transformer-based LLMs like Llama-2-7B on the PredEx dataset.
- We preprocessed the Dataset to handle document length, token limits, and noise reduction.

3. Evaluation

- Used quantitative metrics such as accuracy, Macro-F1, and BLEU for prediction quality.
- Calculated Semantic similarity scores (BERTScore).

Results

Challenges

Conclusion

References

1. *Legal Judgment Reimagined: PredEx and the Rise of Intelligent AI Interpretation in Indian Courts*. Shubham Kumar Nigam, Anurag Sharma, Danush Khanna, Noel Shallum, Kripabandhu Ghosh, Arnab Bhattacharya.
