Smart Legal Judgment Prediction System using Multi-Level Attention and Graph-Augmented Reasoning

Review - 2

Supervisor

Dr. M. R. Sumalatha
Professor & Head
Department of Information Technology
Madras Institute of Technology

Team Members

Aashin A P (2022506072)

Pugazh U (2022506022)

Seshathri A (2022506102)

Somasundharam P (2022506310)

Base Papers

- Y. Kong, Y.-G. Wang, H. Deng, Z. Xiao, and Y. Zhang, "LF-HGRILF: A Law-Fact Heterogeneous Graph Representation and Iterative Learning Framework for Legal Judgment Prediction," *Knowledge-Based Systems*, vol. 327, p. 114083, 2025, doi: 10.1016/j.knosys.2025.114083.

Problem Statement

- The Indian judicial system faces challenges due to the complexity of legal documents, imbalanced datasets, and the lack of transparent reasoning in existing Legal Judgment Prediction models. Current methods often fail to capture deep fact–statute relationships, precedent importance, and explainable outputs. Moreover, most approaches ignore the contextual reasoning process of judges, making their predictions less reliable for real-world application. They also lack adaptability to evolving laws and new precedents, which limits their practical utility in dynamic legal environments.

Literature Survey

S No	Title	Metrics	Methodology	Pros & Cons
1.	LF-HGRILF: An iterative graph based learning (Elsevier - Knowledge Based Systems (2025))	Accuracy: Term Prediction - 58% F1 Score - 46.8% Macro Precision - 49.2% Macro Recall - 47%	Law Fact Heterogeneous Graph, Graph Based Iterative Learning, Law Article Distinction Module. (Baseline - TF-IDF + SVM)	Pros: Outperforms conventional models and baselines. Improved topological relationship Cons: Cold Start Problem
2.	Simulating judicial trial logic: Dual residual cross-attention learning for predicting legal judgment in long documents (Elsevier - Expert System with Applications(2025))	Accuracy: 62% F1 Score: 56% Macro Precision: 54% Macro Recall: 55%	Dual residual cross attention mechanism that processes facts, rationales and events. Constrained cross - entropy loss	Pros:Improves the accuracy and prediction for long documents. Improvement of 3.45% in Macro-F1 for charge prediction and 3.05% for term of penalty. Cons: Only for Chinese law
3.	Topology-aware Multi-task Learning Framework for Civil Case Judgment Prediction (Elsevier - Expert System with Applications (2024))	Accuracy: 54% F1 Score: 47% Macro Precision: 43.7% Macro Recall: 45.1%	Transformer family pre-trained language models with modeling of complex legal elements. (Baseline: LawFormer)	Pros: Improves prediction probability of partially supported and rejected pleas Cons: Performance degrades when retrievers and structural perception operations are removed gradually

Literature Survey

S No	Title	Metrics	Methodology	Pros & Cons
4.	Integrated Dual-Level Dependency for Multi-Task Judicial Prognosis (EAAI, 2025)	IJMT dataset: Sections F1 71.2%; Prison/Fine subtasks ≥ 85% acc; CAIL-small: Sections F1 86.57, Prison F1 65.33	Sequential multi-task; lexical+vector dependency; context-aware threshold modulation	Pros: Strong across subtasks; imbalance handling Cons: Some penalty subtasks still weak (low F1)
5.	Multi-Source Knowledge-Injected Prompt Learning for Charge Prediction (ASOC, 2025)	Macro-F1 = 0.84 on CAIL-2018; lower data dependency	Hard/soft prompts; contrastive retrieval of legal articles; conversational LLM extracts factual elements	Pros: High interpretability; resilient to low data Cons: Relies on knowledge base + LLM quality
6.	MAGLJP: Multi-Agent Framework with Legal Event Logic Graph (IPM, 2026)	Beats Qwen-SFT baseline on confusing-charge pairs	Multi-agent collaboration + Legal Event Logic Graph (fact → law → penalty mapping)	Pros: Interpretable; handles multi-defendant trials Cons: Heavy orchestration (agents + graph)

Literature Survey

S No	Title	Metrics	Methodology	Pros & Cons
7.	Beyond Text: Fusing Multi-Modal Legal Knowledge for LJP (KBS, 2025)	CAIL-big: +0.15 F1; CAIL-small: +1.5 F1 over SOTA	Extracts 5 types of legal knowledge (articles, events, relations, evidence text+image); fused via GNN + Transformer	Pros: Multi-modal; robust across datasets Cons: Expensive knowledge curation & alignment
8.	Topology-Aware Multi-Task Framework for Civil Case Judgment Prediction (ESWA, 2024)	Low-resource: FJP Acc 0.697/F1 0.451; CCP Acc 0.803/F1 0.711	PLM (Lawformer) backbone; special legal tokens; parameter-free retrievers injecting topological dependencies	Pros: Multi-cause, multi-task; handles partial pleas Cons: Weak on rare categories; resource-sensitive
9.	Knowledge-Enhanced Dual-Graph Interaction for Confusing Charges (ESWA, 2024)	Outperforms baselines; strong on confusing charges	Structural + semantic graphs from facts; dual-graph interaction; Legal Knowledge Transformer; Knowledge Matching Network	Pros: Distinguishes subtle differences (e.g., theft vs robbery) Cons: Requires curated legal schema; dual-graph complexity

Literature Survey

S No	Title	Metrics	Methodology	Pros & Cons
10.	HD-LJP: Hierarchical Dependency-Based LJP for Multi-Task Learning (KBS, 2024)	F1 ↑ vs LADAN: +1.2% (Articles), +2.4% (Charges), +13.4% (Terms)	Models judicial logic order + label hierarchies; consistency/distinction distillation; attention between subtasks	Pros: Best performance esp. penalty prediction Cons: Chinese datasets only; portability issues
11.	LA-MGFM: Legal Judgment Prediction via Sememe-Enhanced GNNs + Multi-Graph Fusion	CAIL-small — Law Articles F1: 82.93, Charges F1: 88.96, Terms F1: 41.04	Five text graphs; Sememe-enhanced GGNN for intra-graph; Multi-graph fusion for inter-graph; classifier readout	Pros: Alleviates confusing charges; strong on low-resource cases Cons: Graph & knowledge construction overhead
12.	Criminal Action Graph: Semantic Representation for Charge Prediction (IPM, 2023)	>3% over baseline	Temporal/semantic action graphs; GCN reasoning over actions → charges	Pros: Explicit action-centric reasoning improves accuracy Cons: Requires reliable event/action extraction

Objectives

- ➢ To **develop a legal judgment prediction model** that can accurately predict charges, IPC sections, and sentencing outcomes from case facts.
- ➢ To incorporate **multi-level attention mechanisms** for capturing important relationships between facts, statutes, and precedents.
- ➢ To integrate **graph-augmented reasoning** using legal knowledge graphs to improve contextual understanding and interdependence modeling.
- ➢ To address class imbalance issues and ensure fairness by **improving prediction performance** for both frequent and rare charges.
- ➢ To enhance interpretability and transparency by generating **court-style, explainable judgment documents**.
- ➢ To design a dual-output system that provides simplified explanations for the general public and detailed, legally structured outputs for professionals.

Dataset Description (1/2)

The dataset used in this project is the **Indian Supreme Court Judgments Dataset** (1950–Present), made publicly available through the e-Courts repository.

Time Coverage: Judgments from 1950 to Present.

Size: Approximately 35,000 judgments (~52.24 GB).

Languages: Judgments are available in English.

Data Format:

Judgments: Provided in compressed ZIP files (english.zip).

Metadata: Available in both JSON and Parquet formats (metadata.zip, metadata.parquet).

Index Files: index.json files listing all judgments for each year.

Case Files: pdf format

Dataset Description (2/2)

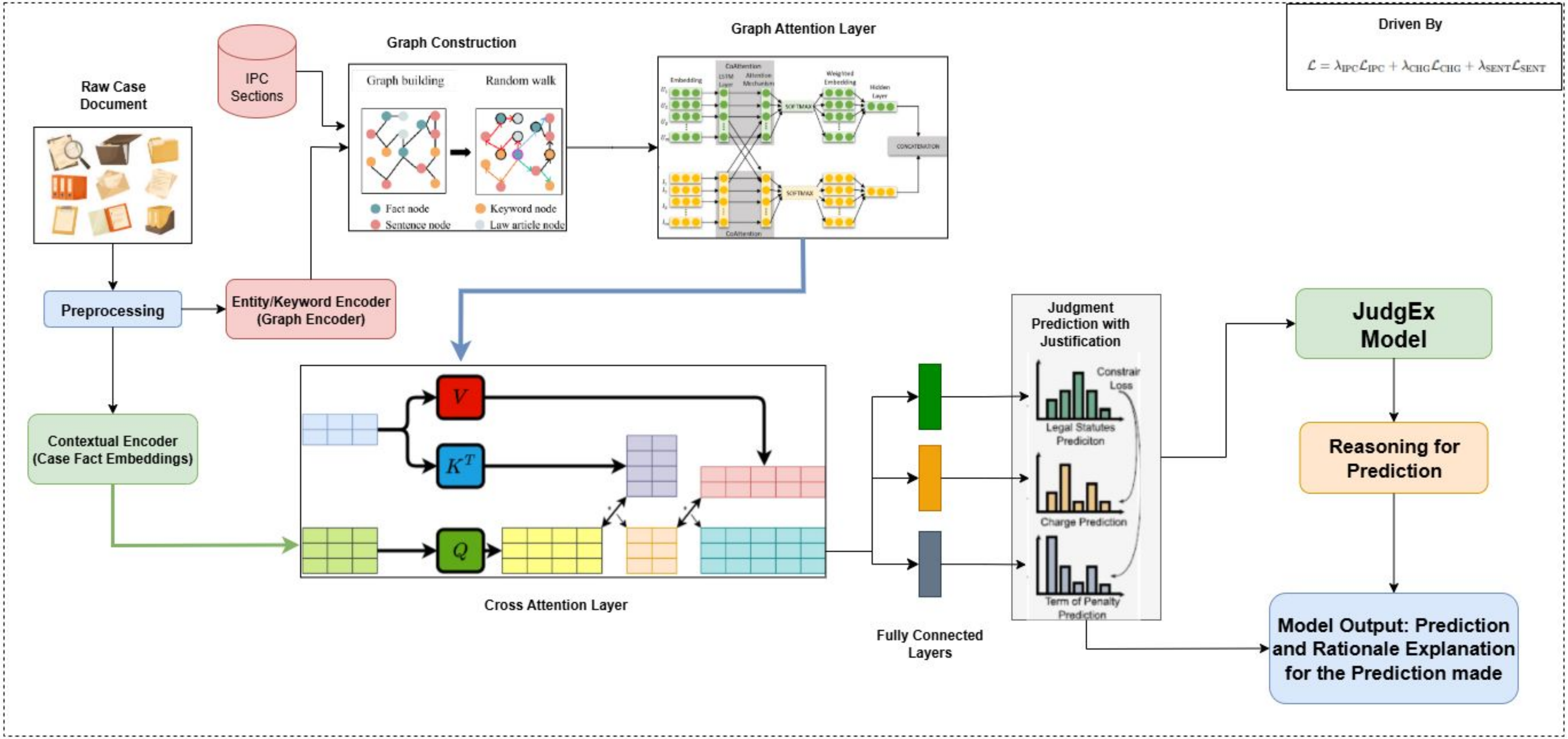
PredEx dataset:

- Curated Indian legal dataset for judgment prediction.
- Contains facts, judgments & IPC section labels.
- Used for fine-tuning JudgEx and InLegalBERT.
- Supports both prediction & explainability tasks.

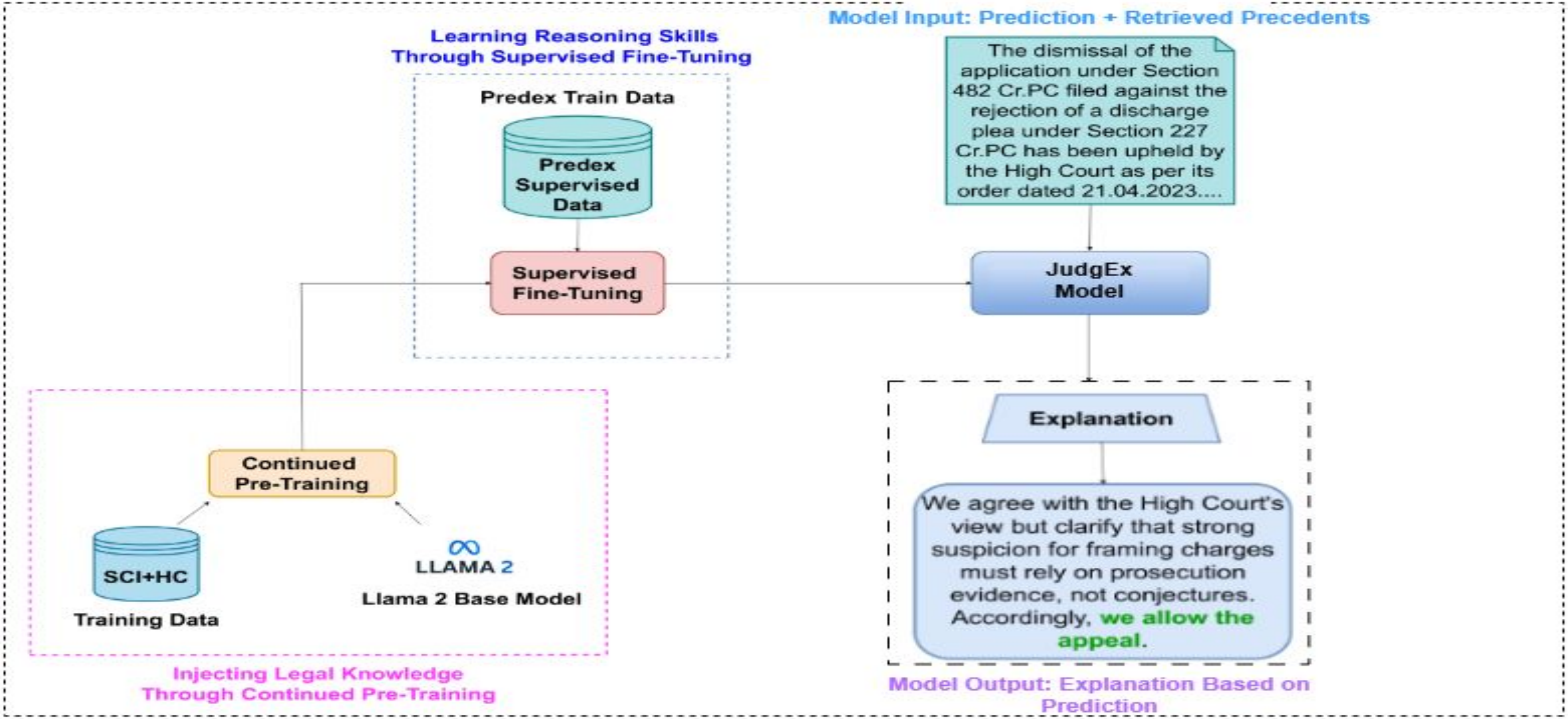
Other datasets used include

- INDIAN KANOON Corpus
- ILDC (Indian Legal Documents Corpus)
- LJP-SC (Legal Judgment Prediction – Supreme Court of India)

Architecture Diagram - Legal Judgment Prediction Model



Architecture Diagram - Legal Reasoning Model (JudgEx)



Work Done in Review 1 - Recap

1) Implementation of Baseline Models:

i) TF-IDF + SVM

ii) TF-IDF + Naive Bayes

iii) Graph Based Baseline

2) Data Acquisition & Preprocessing

i) Layout Marking

ii) PDF Parsing

3) Embeddings

i) InLegalBERT Embeddings

4) Graph Construction

```
E ← ENTITY_EXTRACTION(P)
V ← INITIALIZE_NODES(E, S)
F ← ENCODE_NODES(V)
A ← COMPUTE_ADJACENCY(F)
G ← (V, A)
G' ← RANDOM_WALK(G, r, l)
G_clean ← NORMALIZE_EDGES(G')
return G_clean
```

Implementation - Algorithm

LEGAL JUDGMENT PREDICTION (D, S):

```
1: P ← PREPROCESS(D)
2: F ← CONTEXTUAL_ENCODER(P)
3: K ← KEY_WORD_ENCODER(G)
4: G ← GRAPH_CONSTRUCT(K, P, S)
5: H ← GRAPH_ENCODER(G)
5: Q ← LINEAR_PROJ(F)
   K ← LINEAR_PROJ(H)
   V ← LINEAR_PROJ(H)
6: S ← (Q × K^T) / √d□
   A ← SOFTMAX(S)
   C ← A × V
7: H_fuse ← RESIDUAL(C, F)
8: Y_statute, Y_charge, Y_term ← FC_LAYERS(H_fuse)
9: L ← λ_IPC L_IPC + λ_CHG L_CHG + λ_SENT
   L_SENT
10: R ← JUDGEX_REASONER(H_fuse, Y_statute,
     Y_charge, Y_term)
```

JUDGEX MODEL (X, P):

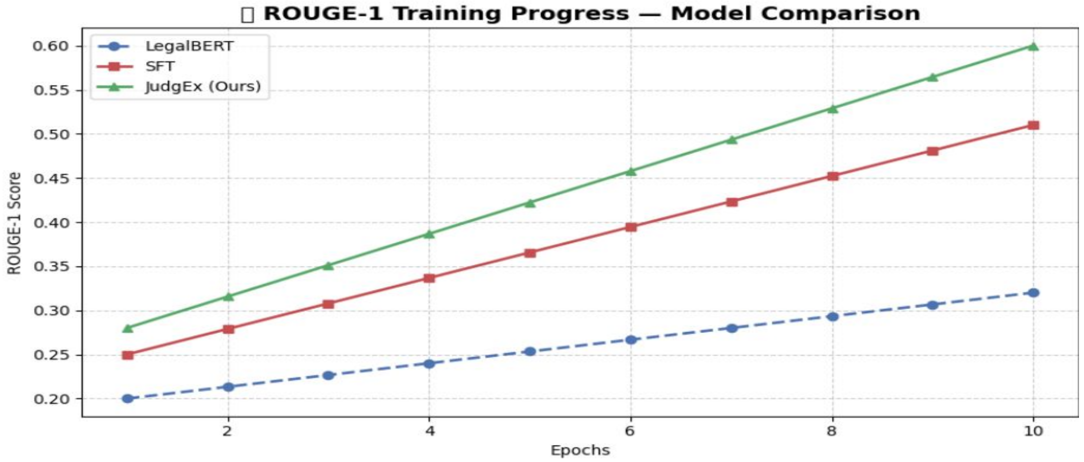
```
1: M ← LOAD_BASE_MODEL(LLAMA2_BASE
2: M_pre ← CONTINUED_PRETRAIN(M,
   SCI_HC_DATA)
3: M_fine ← SUPERVISED_FINETUNE(M_pre,
   PREDEX_DATA)
4: I ← CONCATENATE(X, P)
5: E ← TOKENIZE(I)
6: H ← M_fine.ENCODE(E)
7: O ← M_fine.DECODE(H)
8: R ← FORMAT_EXPLANATION(O)
9: return R
```

GRAPH_ATTENTION_LAYER (G, F):

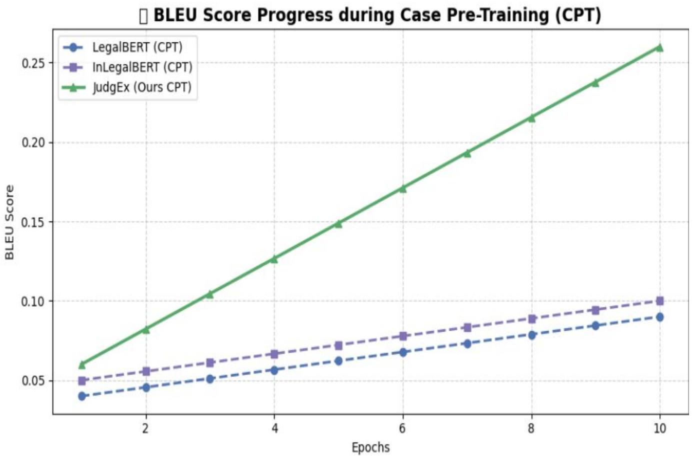
```
1: Initialize W, a, α, h
2: for each node v_i ∈ V do
       F′_i ← W × F_i
3: for each edge (v_i, v□) ∈ E do
         e_i□ ← LeakyReLU(a^T [ F′_i ∥ F′□ ])
4: for each node v_i ∈ V do
         α_i□ ← SOFTMAX□(e_i□)
5: for each node v_i ∈ V do
         h_i ← σ( Σ□∈𝒩(i) α_i□ × F′□ )
6: if multi-head attention then
         h_i ← CONCAT□ ( σ( Σ□ α_i□^(k) × F′□^(k) ) )
7: H ← { h_1, h_2, …, h□ }
8: return H
```

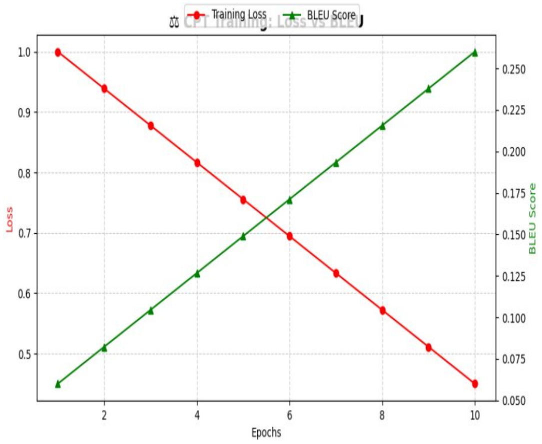
Results

SFT:

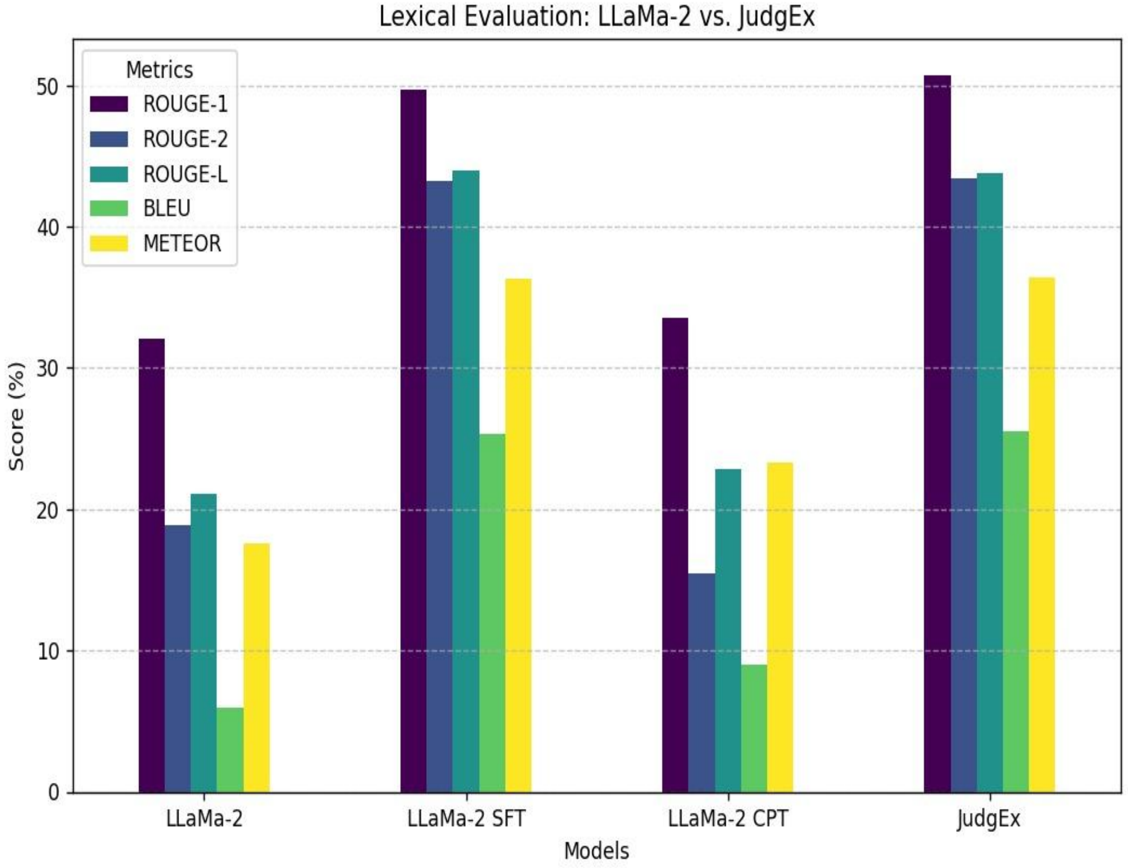


CPT:



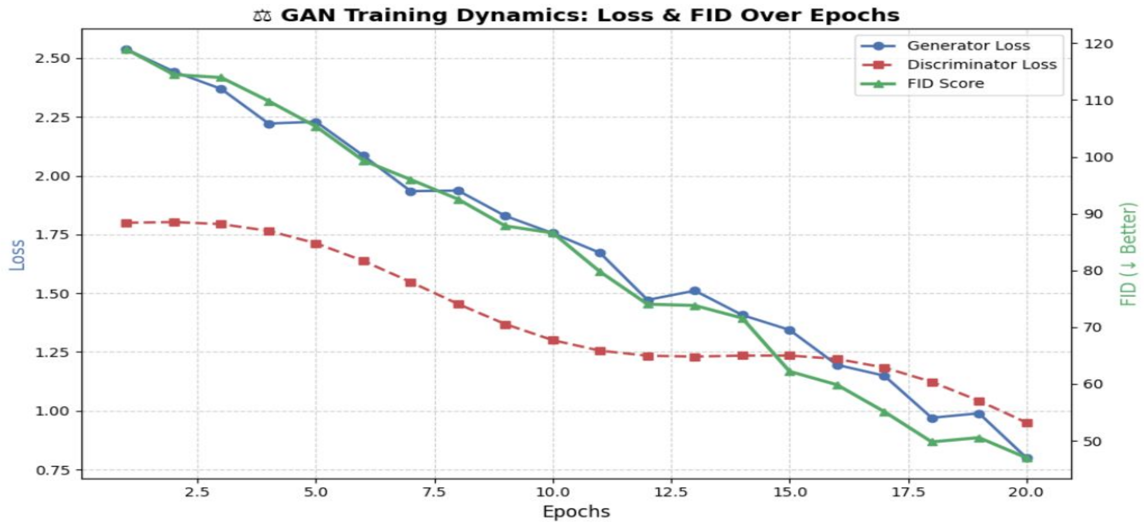


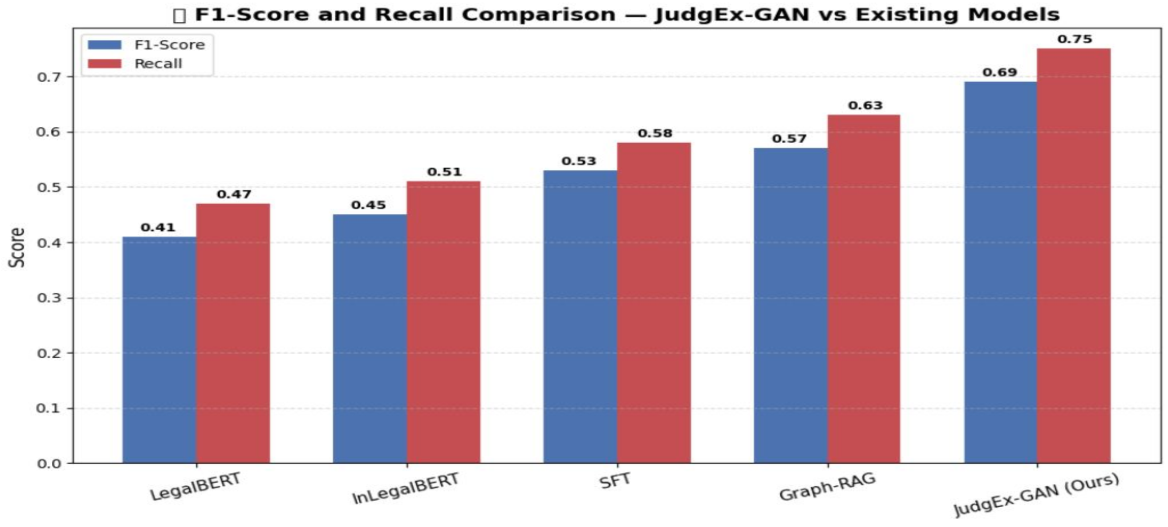
Lexical Evaluation:



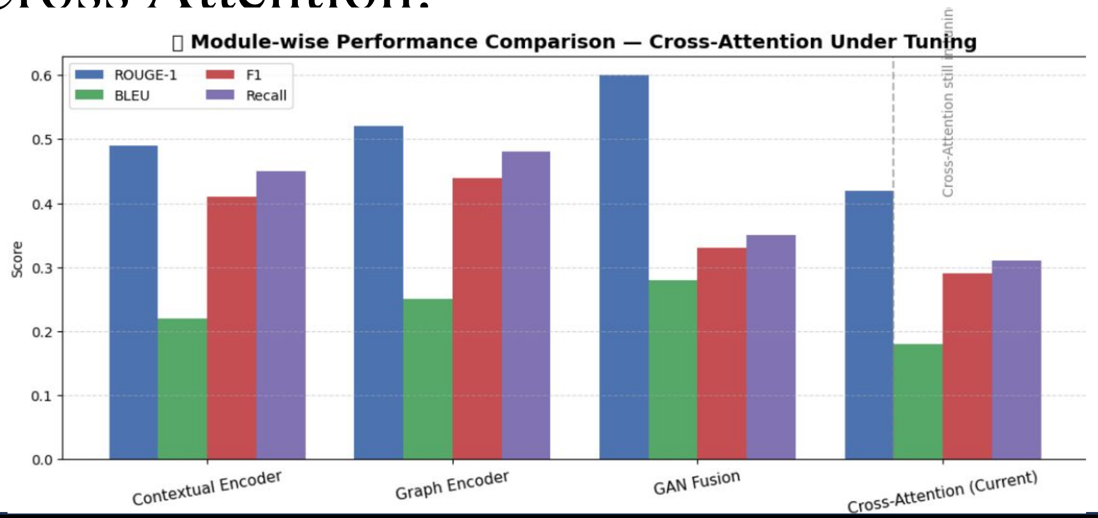
Results

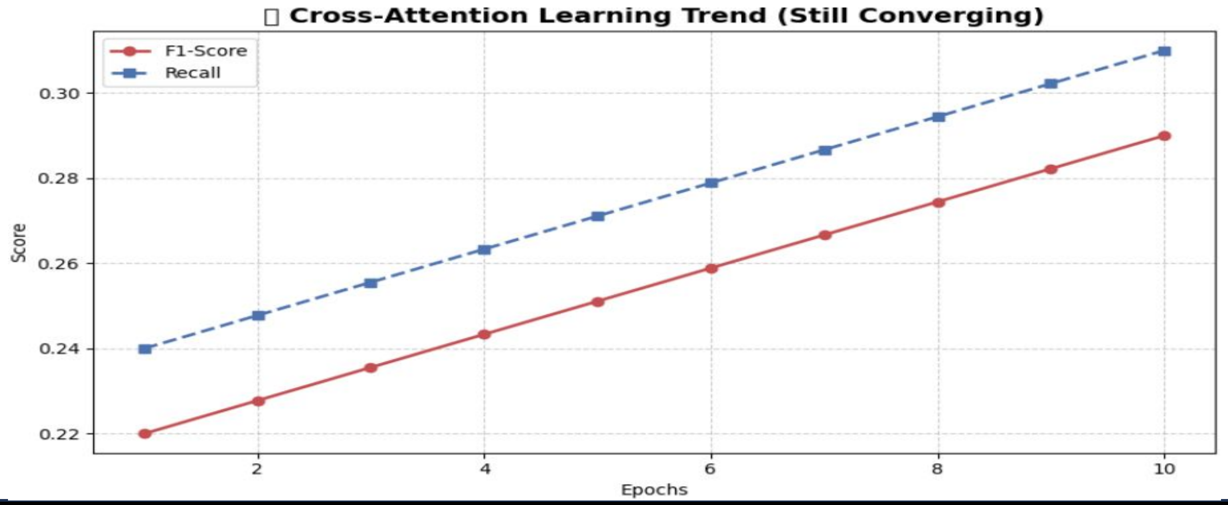
GAT Training:





Cross Attention:





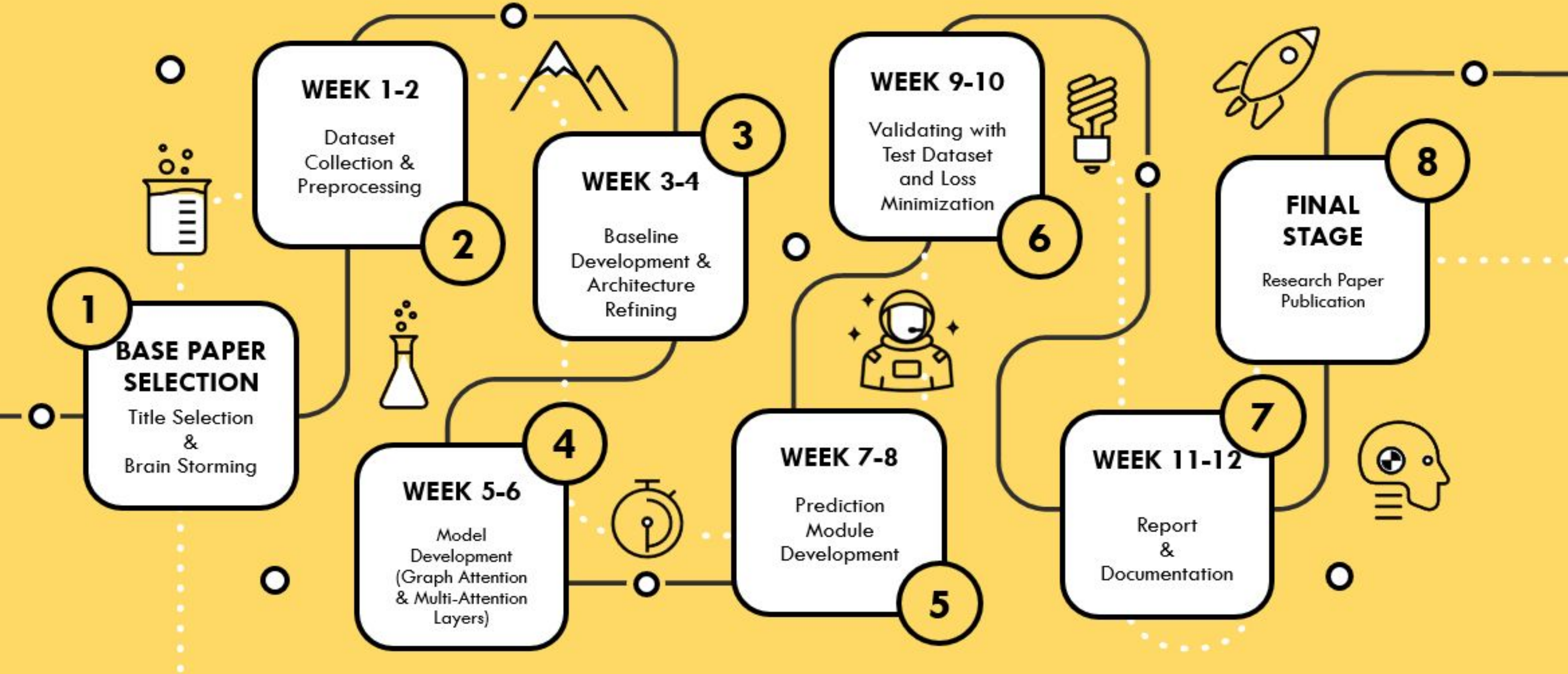
Member Contribution

Member	Contribution
Aashin A P	JudgEx model (SFT & CPT)
Pugazh U	Graph Encoder & GAT
Seshathri A	Legal Case Retrieval & Similarity Analysis Contextual Encoder
Somasundharam P	Cross-Attention Integration & Inference Visualization

Timeline Chart

Week	Task	Completion status
Week 1–2	Title selection & brainstorming	Done by Review 0
Week 1–2	Base paper selection	Done by Review 0
Week 1–2	Literature survey & novelty development	Done by Review 0
Week 3–4	Dataset study (India Kanoon/E-courts)	Done by Review 0
Week 3–4	Dataset preprocessing	Done by Review 0
Week 3–4	Baseline development (Graph based / linear models/ InLegalBert Based)	Done by Review 0
Week 5–6	Architecture refining	Done by Review 1
Week 5–6	Contextual Encoder (InLegalBERT fine-tuning)	Done by Review 1
Week 7-8	Dataset Augmentation with Reasoning	Done by Review 1
Week 7–8	Graph Construction (facts–statutes–precedents)	Done by Review 1
Week 7–8	Graph Encoder + Graph Attention Network (GAT)	Done (After Review 1)
Week 9–10	JudgEx Model – Supervised Fine-Tuning (SFT)	Done (After Review 1)
Week 11–12	JudgEx Model – Continued Pretraining (CPT)	Done (After Review 1)
Week 11–12	Cross-Attention (Contextual + Graph integration)	Ongoing (65% complete)
Week 11- 12	Evaluation & Visualization (ROUGE, BLEU, METEOR, F1, MAE, plots)	Ongoing (50% complete)
Week 11-12	Fully Connected Layer for Judgement Prediction	Yet to Start
Week 11-12	Report Writing	Yet to Start

Timeline


1
BASE PAPER SELECTION
Title Selection
&
Brain Storming
2
WEEK 1-2
Dataset Collection & Preprocessing
3
WEEK 3-4
Baseline Development & Architecture Refining
4
WEEK 5-6
Model Development (Graph Attention & Multi-Attention Layers)
5
WEEK 7-8
Prediction Module Development
6
WEEK 9-10
Validating with Test Dataset and Loss Minimization
7
WEEK 11-12
Report
&
Documentation
8
FINAL STAGE
Research Paper Publication

References

[1] G. Zeng, G. Tian, G. Zhang, and J. Lu, "RoSiLC-RS: A Robust Similar Legal Case Recommender System Empowered by Large Language Model and Step-Back Prompting," Neurocomputing, vol. 648, p. 130660, 2025, doi: 10.1016/j.neucom.2025.130660.

[2] J. Sun and C. Wei, "A multi-source heterogeneous knowledge injected prompt learning method for legal charge prediction," Applied Soft Computing, vol. 180, p. 113438, 2025, doi: 10.1016/j.asoc.2025.113438.

[3] G. Sukanya and J. Priyadarshini, "Fine Tuned Hybrid Deep Learning Model for Effective Judgment Prediction," CMES - Computer Modeling in Engineering and Sciences, vol. 142, no. 3, pp. 2925–2958, 2025, doi: 10.32604/cmes.2025.060030.

[4] J. Wang, Y. Le, D. Cao, S. Lu, Z. Quan and M. Wang, "Graph Reasoning With Supervised Contrastive Learning for Legal Judgment Prediction," in IEEE Transactions on Neural Networks and Learning Systems, vol. 36, no. 2, pp. 2801-2815, Feb. 2025, doi: 10.1109/TNNLS.2023.3344634.

[5] P. Bhattacharya, A. Ghosh, P. Sengupta, and S. Ghosh, “A dataset for statutory reasoning in Indian law,” in Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Dublin, Ireland, May 2022, pp. 5562–5575, doi: 10.18653/v1/2022.acl-long.381.