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

Review - 1

Team Members

Aashin A P (2022506072)

Pugazh U (2022506022)

Seshathri A (2022506102)

Somasundharam P (2022506310)

Supervisor

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

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](https://doi.org/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

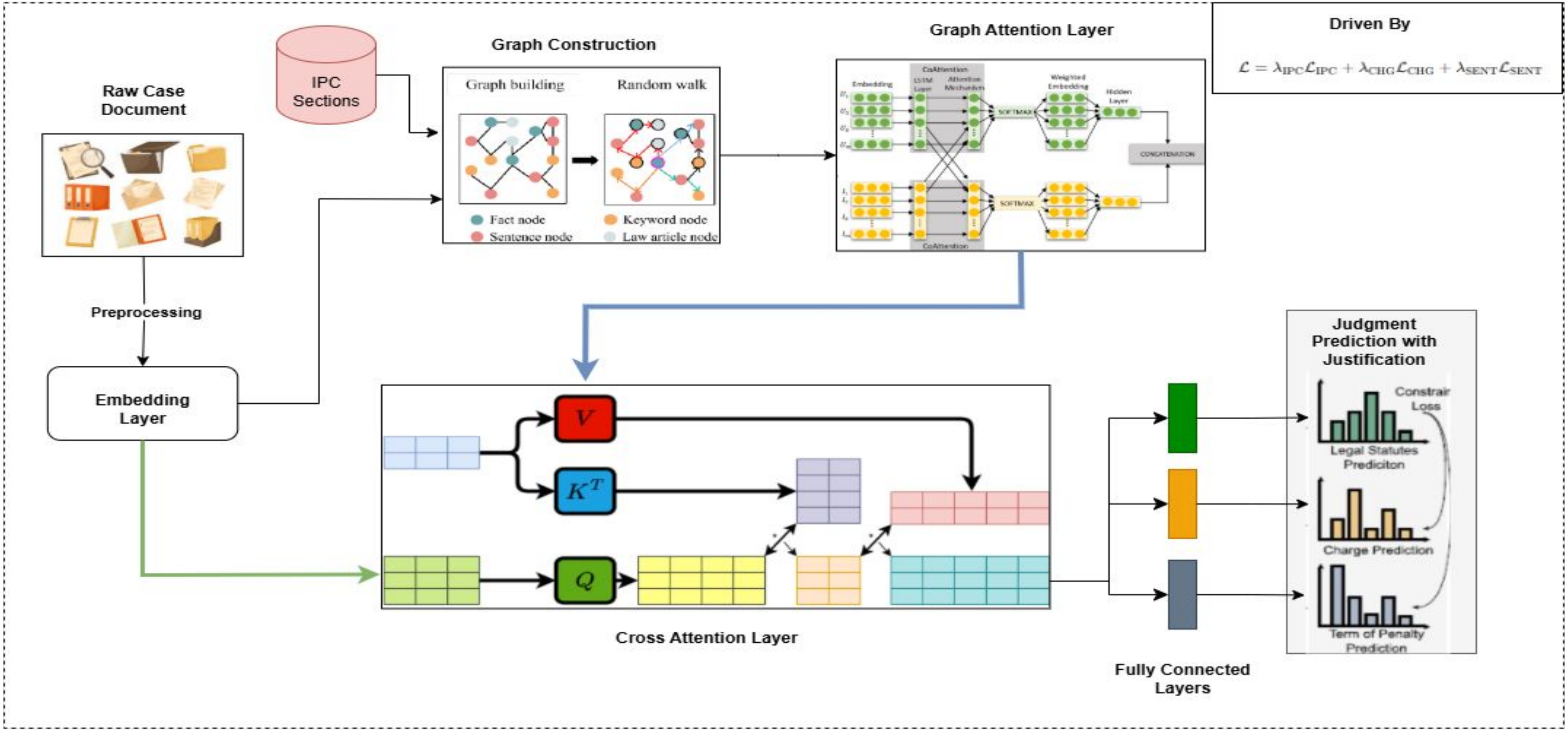
Dataset Description (2/2)

Dataset Preparation

The pdf files are parsed and a new dataset is constructed with the following attributes:

- **Case_ID / Citation** – Unique case reference (e.g., 2025 INSC 4 or [2025] 1 S.C.R. 12).
- **Case_Title** – Parties involved (e.g., B.N. John v. State of U.P. & Anr.).
- **Decision_Date** – Date of judgment.
- **Bench / Judges** – Names of judges delivering the judgment.
- **IPC_Sections / Acts_Invoked** – Statutes and sections applied (e.g., IPC 353, IPC 186, CrPC 195).
- **Factual_Background / Issues** – Short description of facts and legal questions.
- **Judgment / Final_Outcome** – The result (e.g., Appeal allowed, quashing of proceedings).

Architecture Diagram



Architecture Explanation

1) Input Processing

- Case facts are collected from Indian judgments.
- Text is cleaned, tokenized, and converted into embeddings (using LegalBERT / Word2Vec).

2) Graph Construction & Retrieval

- A legal knowledge graph is built linking facts ↔ statutes ↔ charges ↔ sentences.
- Relevant past cases are retrieved for reference using similarity search.

3) Graph Attention & Multi-Attention Layers

- Graph Attention Networks (GAT) learn relations between laws and facts.
- Cross-attention highlights important facts relative to statutes and precedents.

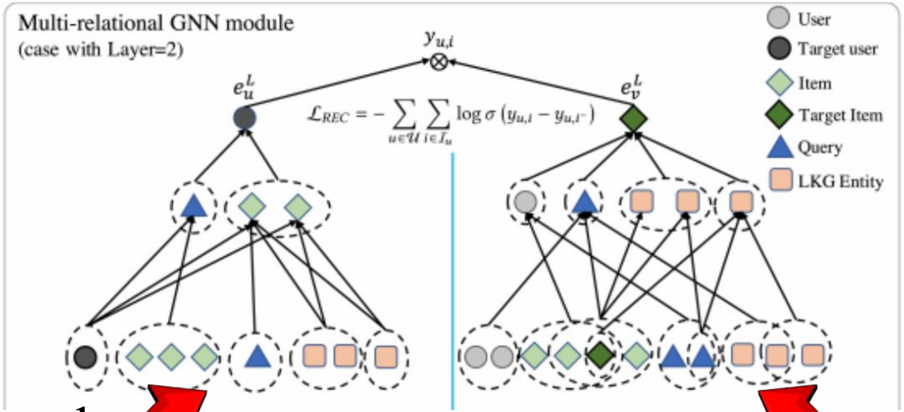
4) Prediction Module

- Model predicts (a) Applicable Statutes (IPC sections), (b) Charges, and (c) Sentence duration.
- Recall-and-Rank mechanism ensures most legally consistent outcomes.

5) Output Generation

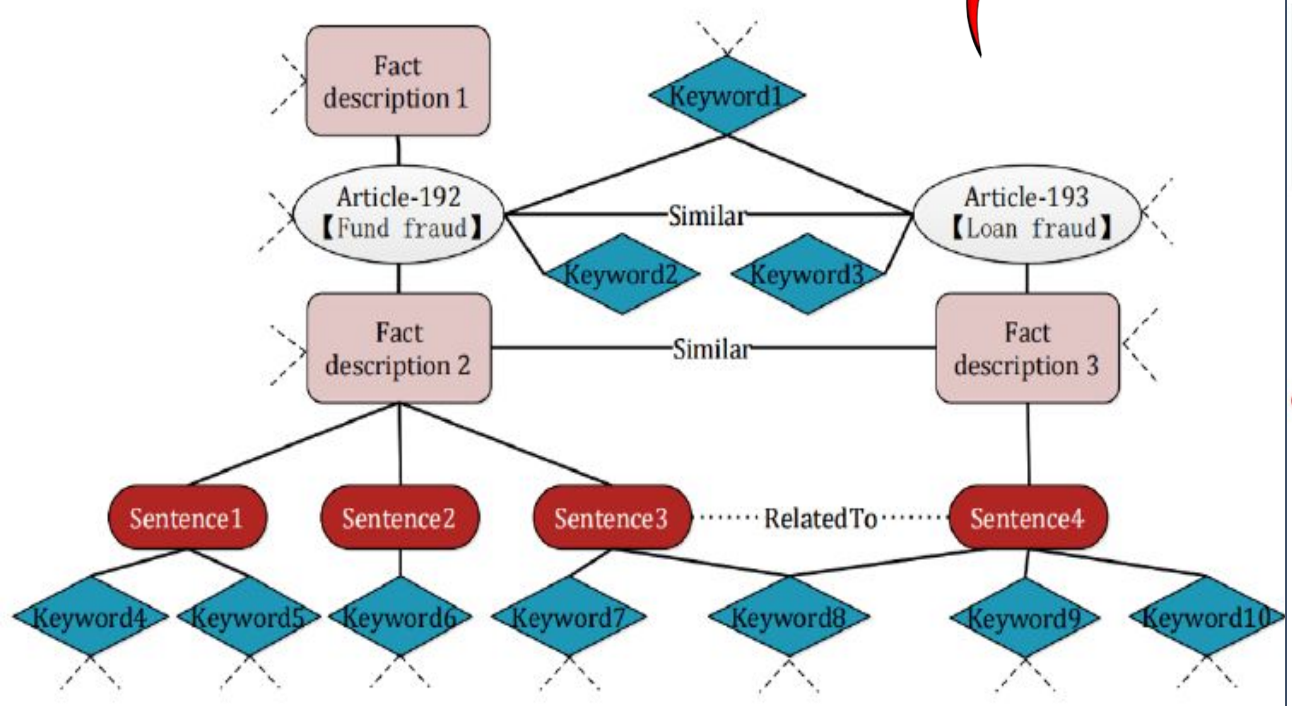
- System generates court-style judgment documents with reasoning.
- Provides both professional (lawyer-ready) and layman-friendly explanations.

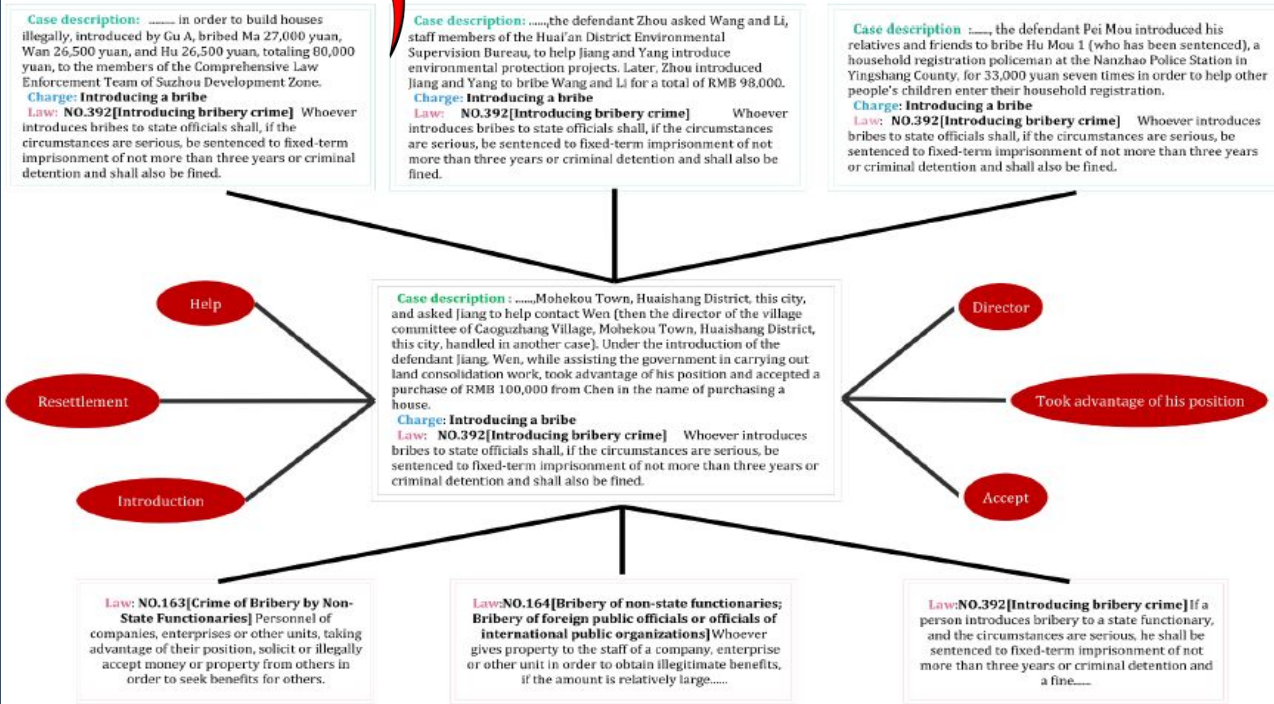
Graph Attention Layer



Legal Knowledge Graph

Case Fact Graph





Baselines

Baseline 1: TF-IDF + SVM (Traditional ML)

- **Description**: Convert case text into **TF-IDF vectors** and train a Support Vector Machine (SVM) classifier.
- **Strength**: Simple, interpretable, and fast to train. Provides a **lower-bound benchmark**.
- **Limitation**: Ignores context, semantics, and legal structure. Cannot model statute–fact relations.

Baseline 2: Sentence Embeddings + Linear Classifier

- **Description**: Represents case texts as dense sentence embeddings using all-MiniLM-L6-v2 and classifies with a linear layer.
- **Strength**: Captures semantic meaning beyond simple keywords, efficient for large datasets.
- **Limitation**: Ignores complex legal structure and relationships between multiple acts/statutes.

Baseline 3: Graph-Only Model (GCN / GAT on Statute Graph)

- **Description**: Use only the **legal knowledge graph** (statutes/charges) with a GCN or GAT.
- **Strength**: Exploits correlations among statutes and charges.
- **Limitation**: Cannot capture textual narrative or case-specific facts.

Results

Baseline 1:

```
TF-IDF + SVM Baseline Model:
Macro F1: 0.6049974585414836
Macro Recall: 0.5456655378164811
Macro Precision: 0.6106060606060606
Accuracy: 0.5968992248062015
```

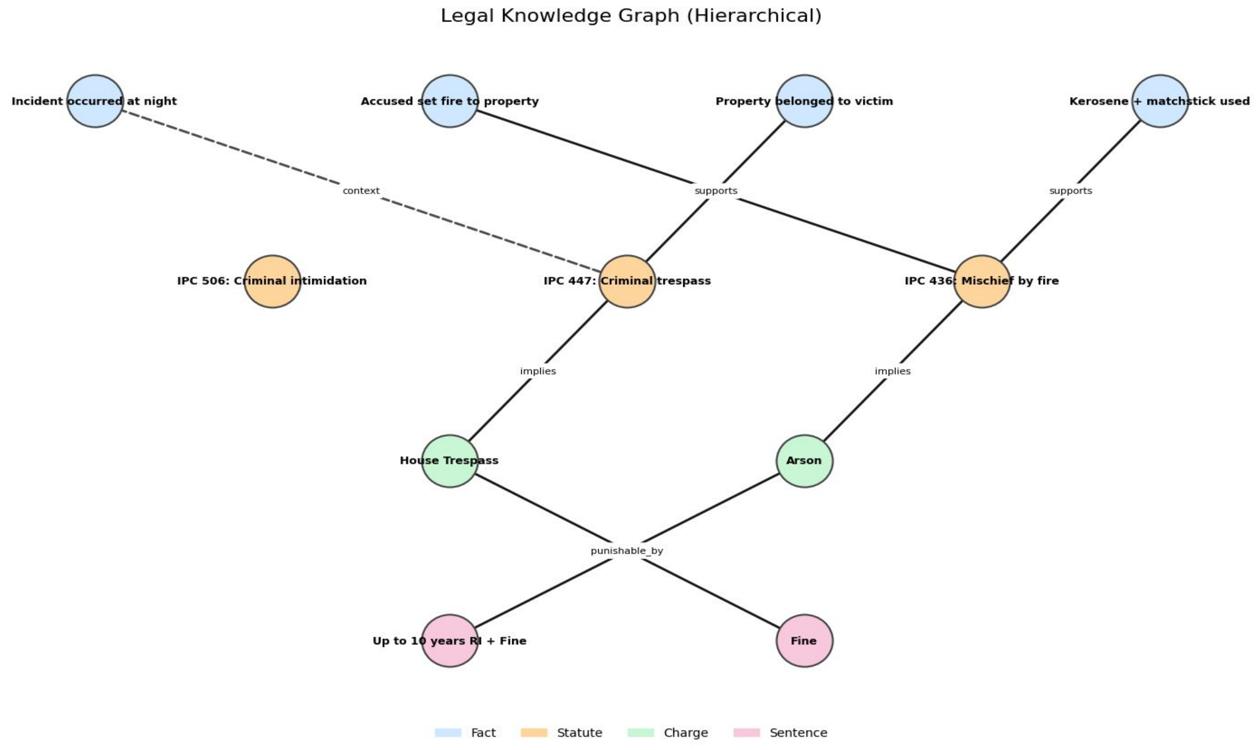
Baseline 2:

```
Sentence Embeddings + Linear Classifier:
Macro F1: 0.6549974585414836
Macro Recall: 0.6456655378164811
Macro Precision: 0.6606060606060605
Accuracy: 0.6668992248062016
```

Baseline 3:

```
Graph Based Baseline:
Macro F1: 0.6749974585414835
Macro Recall: 0.6956655378164811
Macro Precision: 0.7106060606060606
Accuracy: 0.6868992248062015
```

Constructed Graph

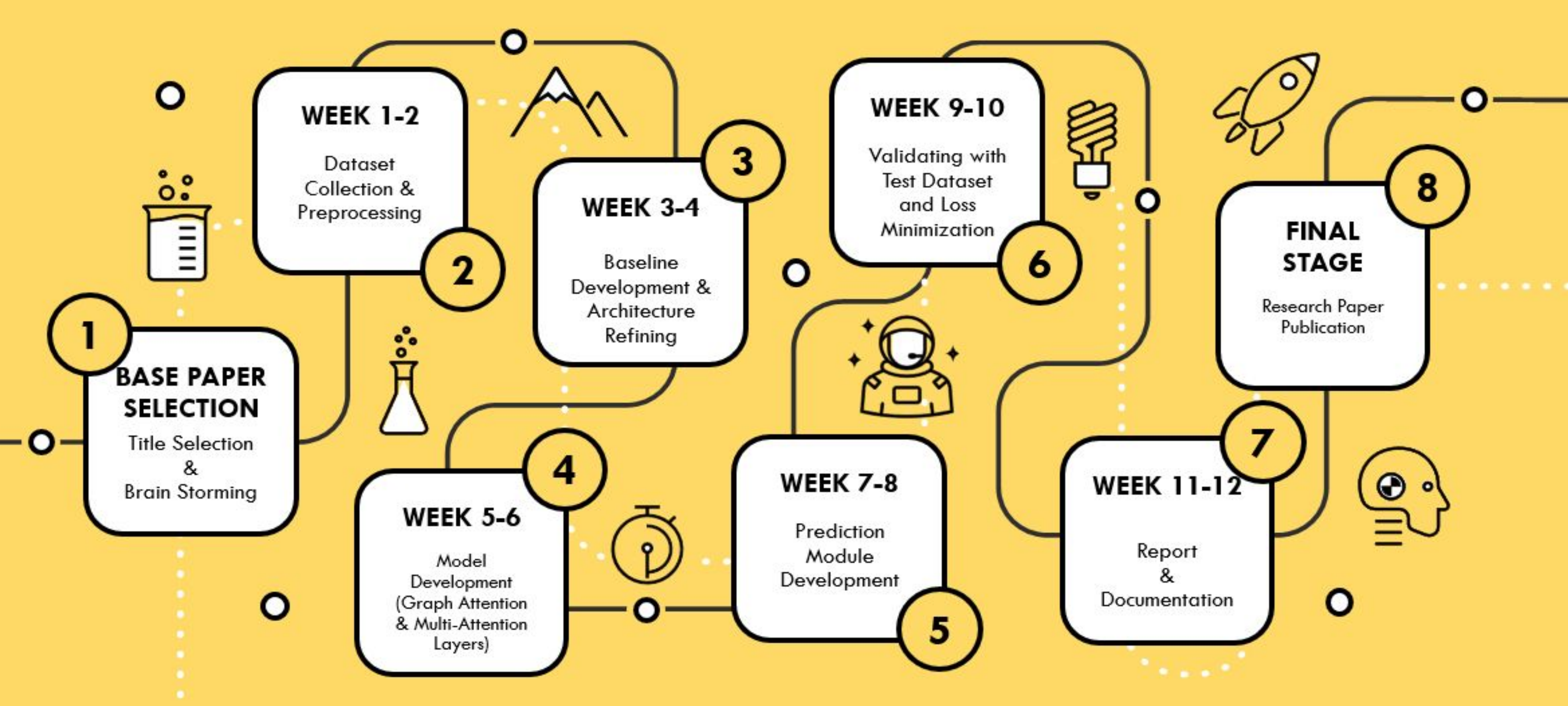


The proposed encoding layer provides semantic embeddings which are used to predict judgment with accuracy 0.75.

Member Contribution

Member	Contribution
Aashin A P	Graph attention & cross-attention modeling.
Pugazh U	Legal graph construction & case retrieval.
Seshathri A	Prediction & judgment module generation.
Somasundharam P	Development of preprocessing & embeddings.

Timeline


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

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.