Knowledge-Infused Legal Wisdom: Navigating LLM Consultation through the Lens of Diagnostics and Positive-Unlabeled Reinforcement Learning

Yang Wu♣* **Chenghao Wang**♡ **Ece Gumusel**◇ **Xiaozhong Liu**♣

♣Worcester Polytechnic Institute ♡Peking University ◇Indiana University at Bloomington

{ywu19, xliu14}@wpi.edu chenghao.wch@gmail.com egumusel@iu.edu

Abstract

The integration of generative Large Language Models (LLMs) into various applications, including the legal domain, has been accelerated by their expansive and versatile nature. However, when facing a legal case, users without a legal background often struggle to formulate professional queries and may inadvertently overlook critical legal factors when presenting their case narrative to LLMs. To address this issue, we propose the Diagnostic Legal Large Language Model (D3LM), which utilizes adaptive lawyer-like diagnostic questions to collect additional case information and then provides high-quality feedback. D3LM incorporates an innovative graph-based Positive-Unlabeled Reinforcement Learning (PURL) algorithm, enabling the generation of critical questions and enhancing user-LLM interactions. Moreover, an integrated LLM-based stopping criterion facilitates precise Court Views Generation (CVG). Our research also introduces a new English-language CVG dataset[1] based on the US case law database, enriching the realm of LLM research and deployment with a vital dimension. D3LM surpasses classical LLMs by delivering outstanding performance and a remarkable user experience in the legal domain.

1 Introduction

Recent years have witnessed a substantial increase in the demand for legal services (Gans-Morse, 2017; Purba and Syahrin, 2019). Despite this growing need, legal resources remain scarce, especially for individuals with modest means (Hadfield, 2010). This scarcity has led to a significant rise in the cost of legal assistance (Brescia et al., 2014), making it difficult for ordinary people to afford the necessary legal support. Furthermore, this situation exacerbates the inherent inequities in legal proceedings between the affluent and the underprivileged. Affluent individuals often benefit from dedicated legal teams to represent them in disputes, whereas the economically disadvantaged lack similar access to legal aid (Horwitz, 2020). This discrepancy not only hinders the ability of the poor to seek justice but also perpetuates a system of latent unfairness in legal cases across socio-economic divides.

With the advent of LLMs, ordinary individuals have gained increased access to legal resources and services. However, the current mode of interaction between users and LLMs predominantly remains user-led (Achiam et al., 2023; Touvron et al., 2023), where users input queries and receive responses from LLMs. This interaction mode presents significant limitations, particularly as many users lack a legal background (Zamfirescu-Pereira et al., 2023). Their queries often lack the essential factual details about their legal issues, which in turn, hampers the ability of LLMs to provide effective assistance. Recognizing this challenge, we have developed D3LM, a novel framework designed to act as a legal assistant, as illustrated in Fig. 1. In this toy example, we introduce the D3LM model through a real-life scenario where a client, facing legal charges from a minor altercation, seeks assistance. Initially, the client provides a vague description, prompting a comparison among three different approaches. Traditional LLMs offer a broad perspective of the court's view, but often overlook critical details due to insufficient information. In contrast, consultations with a lawyer provide a thorough investigation through targeted questions, albeit at a significant cost. D3LM distinguishes itself by autonomously generating specific, case-relevant questions, thereby gaining a deeper understanding of the situation and enhancing the accuracy of legal outcome predictions, all without incurring the high expenses associated with legal counsel. By engaging users in a dialogue to extract detailed information about their legal predicaments, D3LM mirrors the proactive approach of human lawyers in gathering pertinent data to better assist

---

*Corresponding author
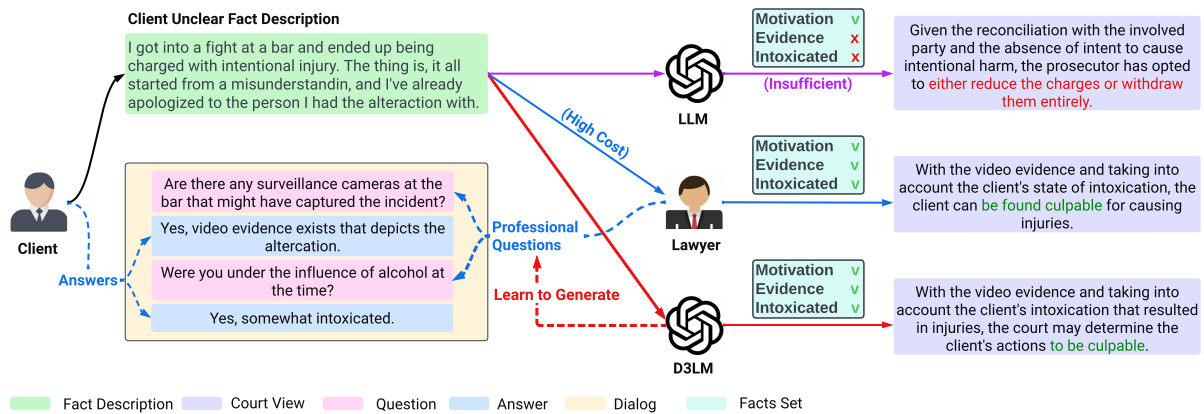[1]https://github.com/YANGWU001/US_CVG_dataset.git

Figure 1: Comparison of legal service methodologies, highlighting traditional LLMs, lawyer consultations, and the D3LM model. D3LM innovatively generates professional questions, mirroring the actions of a lawyer, to improve legal outcome accuracy without high costs, demonstrating a cost-effective, precise approach to legal assistance.

their clients. This innovative engagement method ensures D3LM delivers highly accurate and cost-effective legal insights, demonstrating its potential to transform the provision of legal services.

Moreover, we introduce a novel approach that leverages LLMs to efficiently process and analyze extensive U.S. legal cases. Utilizing LLMs, we generate structured fact-rule graphs and the 'Issue, Rule, Analysis, Conclusion' (IRAC[2]) (Gensler, 1985) to transform complex case narratives into concise representations. Our method integrates these graphs into a comprehensive knowledge base, designed to identify and fill in missing factual information in legal scenarios. By employing a reinforcement learning approach, we refine the selection of relevant fact nodes, enhancing the model's ability to guide users through legal inquiries effectively. This system is rigorously evaluated to ensure its accuracy and practicality in real-world legal settings, showcasing its potential to innovate legal case analysis.

Our contributions in this paper are threefold and can be summarized as follows:

- Novel User-LLM Interaction Paradigm: We introduce a new user-LLM conversion framework in legal contexts by shifting from user-led enquiries to LLM-navigated diagnostics. Our approach actively seeks detailed case information from user through targeted questioning, resulting in unparalleled accuracy and efficacy in legal advice.

- Pioneering Positive-Unlabeled Reinforcement Learning (PURL) Algorithm: Leveraging the

fusion of LLM and legal graph, PURL identifies crucial factors dynamically for adaptive question generation, which enabling LLM's information solicitation capabilities in legal contexts.

- Introducing a new English legal dataset: a new focus on US legal cases. Departing from prior studies relying on Chinese data, this dataset establishes a new benchmark for legal AI experiments in the US case law system.

2 Related Work

2.1 Legal Assistant

Recent advancements in computational methods have significantly impacted judicial decision-making, leveraging the text-centric nature of legal data for natural language processing (NLP) applications. Early research by Lin et al. (2012), Zhong et al. (2018), and Hu et al. (2018) treated charge prediction as a classification problem within legal judgments, laying the groundwork for NLP in legal contexts. This foundation expanded to include case retrieval (Chen et al., 2013) and law article recommendation (Chen et al., 2022), enhancing NLP's scope in legal informatics. The interpretability of legal decisions, explored by Ye et al. (2018) and (Wu et al., 2020), introduced court view generation (CVG) as a means to predict judicial outcomes based on case descriptions, emphasizing the need for nuanced legal text understanding and generation. The introduction of LLMs like Lawformer (Xiao et al., 2021) and Chatlaw (Cui et al., 2023) further advanced this field, demonstrating LLMs' potential in generating accurate legal texts. Un-

---

[2]The detailed explanation of IRAC is in Appendix A.2

like previous models, our D3LM approach mimics a lawyer's consultative strategy, engaging in dynamic interaction through targeted questioning to gather comprehensive case details, marking a significant evolution in the application of AI for legal informatics.

2.2 Enhanced Learning through PU Learning and Reinforcement Learning

In the landscape of computational learning, Positive-Unlabeled (PU) Learning has emerged as a key strategy for binary classification, leveraging a mix of a small number of positive instances and a large volume of unlabeled data. This methodology has been applied in diverse areas including chronic disease prediction (Chen et al., 2020; Wu et al., 2023), fake news detection (Liu and Wu, 2020), and recommendation systems (Zhou et al., 2021), utilizing advanced techniques like uPU and nnPU (Kiryo et al., 2017) alongside unbiased risk estimators (Du Plessis et al., 2015) to mitigate data bias. Innovations such as predictive adversarial learning (Hu et al., 2021; Wei and Li, 2018) and Dist-PU (Zhao et al., 2022) further refine this approach by enhancing data generation and label distribution consistency. Additionally, the adaptation of reinforcement learning, particularly bandit learning, has introduced novel training signals and strategies, including context-free (Chapelle and Li, 2011; Auer et al., 2002) and contextual bandits (Chu et al., 2011; Zhou et al., 2020), to improve model outcomes. Our work introduces the PURL framework, integrating the advanced capabilities of LLMs with domain-specific PU models to achieve decision-making precision tailored to unique application contexts, signifying a notable advancement in computational learning methodologies.

2.3 Large Model Dataset Creation

In the NLP field, LLMs have transformed dataset generation, producing extensive datasets that capture diverse linguistic phenomena, enhancing versatility for various NLP tasks. Examples include WANLI (Liu et al., 2022), utilizing GPT-3 for generation and dataset cartography (Swayamdipta et al., 2020) for refining NLI examples, and DISCO (Chen et al., 2023), which creates counterfactual examples through local perturbations. However, deploying LLMs in legal domains, especially for U.S. cases, faces challenges like narrative length and complexity, which exceed LLMs' token limits, hindering dataset creation and model training

(Jin et al., 2024). Our approach leverages LLMs to streamline U.S. legal narratives, facilitating manageable CVG dataset creation and bridging English-language resource gaps with legal expert collaboration. This method addresses English legal analysis dataset scarcity, adding significant resources to the domain.

3 Methodology

In this section, we begin by introducing the definition of the CVG problem in Section 3.1. We then proceed to present the overall D3LM framework in Fig. 2 in Section 3.2. Finally, we demonstrate the PURL algorithm in Alg. 1 in detail in Section 3.3. To facilitate understanding and clarity, we provide a summary of the notations used in this paper in Tab. 1.

Table 1: Notations.

Notations	Descriptions
D	Cases fact descriptions set.
C	Cases court views set .
$\overline{D}$	Reconstructed fact descriptions set without masked facts.
Q	Generated questions based on golden court views set.
F	Candidate fact nodes set.
$\hat{G}$	Fact-rule graphs set for cases.
G'	N-hop graphs set of masked fact-rule graphs.
$\overline{G}$	Masked fact-rule graphs set.
G	Merged fact-rule graph.

3.1 Problem Definition

In this study, we address the challenge of generating court views in US criminal cases. Our focus is to formalize this problem through a clear delineation of fact descriptions and court views.

Fact Description (D) is the comprehensive narrative of a given legal case, encompassing all pertinent facts and events. We represent fact descriptions as $D = (D_1, D_2, ..., D_{|D|})$, where each D_i corresponds to the fact description of a unique case.

Court Views (C) represent the judicial rationale underlying the court's verdict for a case. We define court views as $C = (C_1, C_2, ..., C_{|C|})$. The primary task in court views generation is to accurately derive C given D.

3.2 D3LM Framework

We elaborate on the D3LM framework as illustrated in Fig. 2. Our framework activates upon receiving a legal query from a client, which is then processed by an advanced Language Model (LLM) to formulate an initial court opinion. However, this opinion is not immediately disclosed; it is contingent on the following procedure. Leveraging the methodology proposed by Schick et al. (2023), we deploy a specialized token to evaluate the completeness of case information provided. This binary token 'Yes' or 'No' guides the subsequent actions: a 'No' activates our PURL algorithm to craft new, case-specific question for client to gather more case information, whereas a 'Yes' moves the process toward the conclusion and delivery of the court opinion. This dynamic cycle of evaluation and information enhancement continues until a 'Yes' token signifies the attainment of a fully informed court opinion.

Utilizing a Case-View Generation (CVG) dataset comprising a set of case descriptions $D = (D_1, D_2, ..., D_{|D|})$ and corresponding court opinions $C = (C_1, C_2, ..., C_{|C|})$, for each pair ($D_i$, C_i), we randomly mask certain facts from both D_i and C_i to derive $\overline{D_i}$ and $\overline{C_i}$. Subsequently, we append the special tokens e_i at the end of each court view, as:

$$e = <\text{T}> \rightarrow \ token \ </\text{T}> \qquad (1)$$

where "$<\text{T}>$", "$\rightarrow$", and "$</\text{T}>$" are fixed special tokens[3]."$token$" is "Yes" for original C_i and "No" for masked $\overline{C_i}$. This augmented dataset serves to fine-tune the LLM. During the inference phase, regular decoding proceeds until the fine-tuned LLM outputs the "$\rightarrow$", which dictates the decision to either continue or conclude the information gathering process. By prompting users to provide detailed responses through case-specific questions, the D3LM framework iteratively refines the court view, ensuring that the final advice is accurate and tailored to the client's needs. This innovative method enhances the model's ability to collect pertinent information, leading to more precise legal guidance.
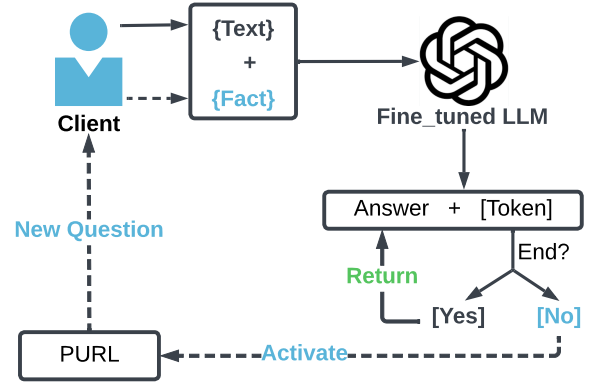

Figure 2: D3LM Model Framework Overview: Illustrates D3LM's engagement through context-driven questions, guided by the PURL algorithm from continuous and historical dialogues. Aims to collect comprehensive case details until a fine-tuned LLM token signals adequate information acquisition.

3.3 Positive-unlabeled Reinforcement Learning (PURL) Question Generation

In this section, we showcase our PURL framework for next question generation through Fig. 3 and detail its training process in Alg. 1.

As illustrated in Fig. 3, for each extensive legal caselaw document, denoted as $Case_i$, we utilize an LLM to extract a fact-rule graph G_i representing the case[4]. Concurrently, we apply the legal IRAC strategy to summarize the case's fact description D_i and court views C_i. Subsequently, the LLM generates a set of questions Q_i based on the text of C_i. For model training, we randomly mask some fact nodes in G_i to create a masked graph $\overline{G_i}$, and then use the LLM to reconstruct the fact description and corresponding court view as $\overline{D_i}$ and $\overline{C_i}$, respectively. Based on the nodes in $\overline{G_i}$, we extract an N-hop subgraph G' from the knowledge graph G, which is a merged graph comprising $\{G_1, G_2, \ldots, G_{|G|}\}$. We then define the candidate fact set as $F_i = \left\{ v \mid v \in V_{G'_i} - V_{\overline{G_i}} \right\}$, where $V_{G'_i}$, $V_{\overline{G_i}}$ represent the nodes in G'_i and $\overline{G_i}$, respectively. The task involves selecting the most relevant node from F_i for the masked case description $\overline{D_i}$. Following node selection, we use a prompt-based approach to transform the selected node into a case-specific question to client. The detailed process for CVG data generation is formalized in Alg. 2.

---

[3]In practice, we use the token sequences "[", "]" and "->" to represent "$<\text{T}>$", "$</\text{T}>$" and "$\rightarrow$"

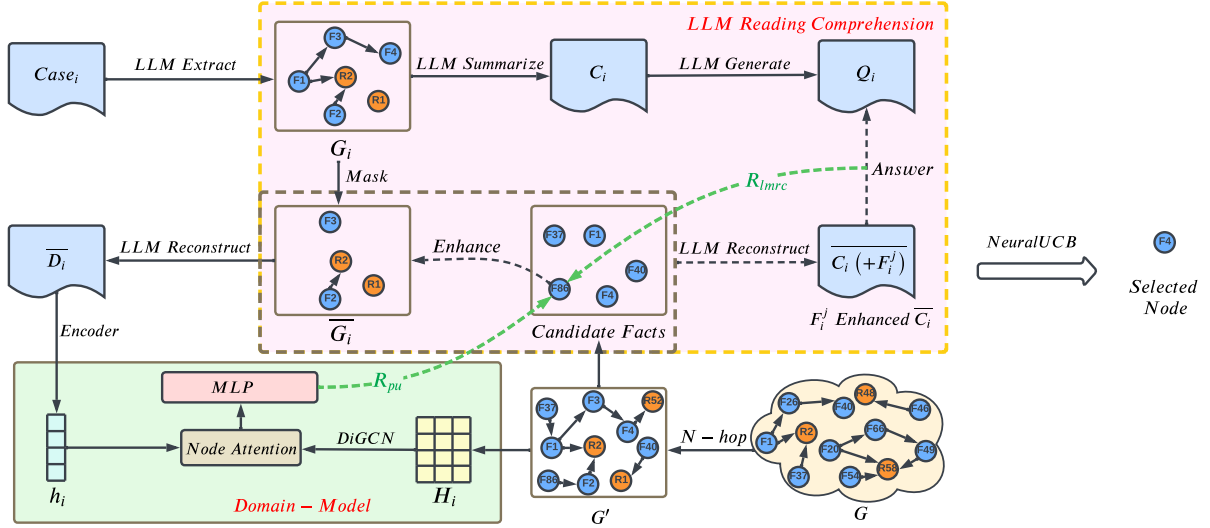[4]The detailed fact-rule graph generation prompt is in Appendix A.2

Figure 3: Illustrative Representation of the PURL Network in Action. This diagram showcases the PURL algorithm's training process using $Case_i$ as an example. It visually delineates the sequential steps of extracting, summarizing, and reconstructing case facts, followed by question generation. Specifically, the collaboration of the LLM with the domain-specific PU model facilitates the final adaptive node selection through NeuralUCB.

3.3.1 Domain Positive-Unlabeled Model Training

Given a masked case description $\overline{D_i}$ and its corresponding subgraph G', we initialize the text embedding h_i using RoBERTa (Liu et al., 2019) and randomly initialize the node embeddings as a matrix $H_i \in \mathbb{R}^{n \times d}$, where n denotes the number of nodes and d denotes the embedding dimension. After processing the node embeddings through several DiGCN layers (Tong et al., 2020), a node attention[5] network is applied to capture semantic relationships between case description text and candidate fact nodes. Subsequently, a Multilayer Perceptron (MLP) is utilized to compute the probability scores.

For each $\{h_i, H_i^j\}$ pair, we designate the set of masked nodes as positive instances and the remaining nodes as unlabeled instances. During the training process, we employ the nnPU (Kiryo et al., 2017) risk estimator, which is defined as follows:

$$
\begin{aligned}
\text{Risk}_{pu}(g) = & \frac{\pi_p}{n_p} \sum_{i=1}^{n_p} l(g(x_i^p), +1) + \\
& max(0, \frac{1}{n_u} \sum_{i=1}^{n_u} l(g(x_i^u), -1) - \\
& \frac{\pi_p}{n_p} \sum_{i=1}^{n_p} l(g(x_i^p), -1))
\end{aligned} \quad (2)
$$

where π is the positive class prior, $l(\cdot, \cdot)$ is any

---

[5]The detailed information of node attention design is in Appendix A.1

surrogate loss of zero-one loss (Du Plessis et al., 2015), n_p represents the number of labeled positive instances, n_u represents the number of unlabeled instances, x_i^p and x_i^u denote instances in the labeled positive set and the unlabeled set, respectively. All cases shared the same domain model layer in training process. The details are formalized in Alg. 3.

Algorithm 1 PURL Training

Input: Training cases fact descriptions set $\overline{D}$; N-hop subgraphs set G'; positive class priors set π; training epochs E; step size η; discount factor γ; LLM
Output: Domain_pu model parameters θ and case-specific NeuralUCB parameters set $\{\phi_1, \phi_2, \ldots, \phi_{|X|}\}$
1: Randomly initialize θ and $\{\phi_1, \phi_2, \ldots, \phi_{|X|}\}$
2: $\theta = \text{Train\_Domain\_PU}(\overline{D}, G', \pi, E, \eta, \gamma)$ using Alg. 3
3: Calculate reward R based on Eq. 4
4: **for** $i = 1, 2, \ldots, |X|$ **do**
5:     **for** $t = 1, 2, \ldots, T$ **do**
6:         Observe $\left\{x_{t,F_i^j}\right\}_{F_i^j \in F_i}$
7:         Select action F_i^j with the largest upper confidence bound using NeuralUCB (Zhou et al., 2020)
8:         Play action F_i^j and observe reward $R(F_i^j)$
9:         Update NeuralUCB parameters ϕ_i using $\left\{x_{z,(F_i)_z}\right\}_{z=1}^{t}$ and $\left\{R_{z,(F_i)_z}\right\}_{z=1}^{t}$
10:     **end for**
11: **end for**
12: **Return** Domain_pu model with updated parameters θ and case-specific NeuralUCB with updated parameters set $\{\phi_1, \phi_2, \ldots, \phi_{|X|}\}$

3.3.2 LLM Reading Comprehension

For each masked case description $\overline{D_i}$ and its corresponding candidate facts set F_i, we utilize the

reading comprehension capabilities of the LLM to assess the significance of each candidate fact $F_i^j \in F_i$. Incorporating F_i^j into the masked description aims to enrich the case narrative by reconstructing the fact description and generating an enhanced court view $\overline{C_i(+F_i^j)}$. This enhanced view, $\overline{C_i(+F_i^j)}$, is then used to answer the questions set Q_i, derived from the original court views text C_i. We evaluate the reading comprehension score of F_i^j for case $\overline{D_i}$ as:

$$RC_{Score}\left(\overline{D_i}, F_i^j ; LLM\right) = \frac{\left| S\left(\overline{C_i + F_i^j}, Q_i\right)\right|}{|Q_i|} \quad (3)$$

where $S\left(\overline{C_i + F_i^j}, Q_i\right)$ is a subset of Q_i that consists of questions correctly answered using $\overline{C_i(+F_i^j)}$.

3.3.3 Model Training

For each case description text $\overline{D_i}$ and candidate fact nodes F_i, we denote their corresponding embeddings are h_i and H_i (H_i is matrix of all fact nodes embeddings). We calculate two types of rewards: 1)R_{pu}, the probability score from the domain-specific PU model, calculated as $P\left(h_i, H_i^j ; \theta\right)$, where θ is the domain positive-unlabeled model; 2) R_{lmrc}, the language model reading comprehension score calculated as Eq. 3. To optimize fact node selection, we employ reinforcement learning (bandit), which allows the model to refine its choice dynamically. The selected fact node is then formulated as the next query for the legal client. The overall reward function is:

$$R_{total} = R_{pu} + \lambda R_{lmrc} \quad (4)$$

where λ adjusts the contribution of the LMRC score to the total reward, and we set 0.5 in our model. By leveraging bandit learning for its flexibility, our model surpasses traditional methods by better handling uncertainty in selecting the most relevant fact node for next question generation.

3.3.4 Model Inference

During inference, the process involves mapping the provided case description text to its corresponding fact-rule graph. From this graph, a subgraph is extracted from the overarching knowledge graph, which yields a set of candidate facts. For each of these facts, the corresponding bandit algorithm is employed to facilitate the selection process.

4 Experiments

4.1 Data Construction

Addressing the notable absence of English-language datasets for court views generation (CVG), our research introduces the first US-CVG dataset, derived from criminal legal documents (Caselaw Access Project, 2024). Unlike the structured format of Chinese legal documents, U.S. case law's complexity and length (averaging over 7694 tokens) pose significant challenges for CVG tasks, often surpassing the token limits of advanced LLMs.

Our dataset leverages the IRAC framework to process US caselaw documents, using GPT-4.0 to summarize and extract fact descriptions and court views, which enables the reduction of the total token length to a structured IRAC format. Additionally, we create fact-rule graphs for each case to support complex analysis. The dataset's integrity is ensured through meticulous review by three legal professionals, after which it is segmented into training, validation, and test sets in an 8:1:1 ratio.

This pioneering dataset fills a critical void for English-language CVG research, offering a structured approach to legal document analysis that paves the way for future computational legal studies.

Table 2: Statistics of US legal dataset.

Type	US
#Train	19491
#Dev	2436
#Test	2437
Avg.#tokens in fact desc.	264
Avg.#tokens in court views	138

4.2 Experimental Settings

The experimental setup for our D3LM model utilized the Llama2-13B as the base architecture. The sampling parameters were set with a temperature of 0.8 and a top_p value for nucleus sampling at 0.9, alongside a maximum token count of 4096. We initialized word embeddings with RoBERTa, while node embeddings were randomly generated. The model's Positive-Unlabeled learning component consisted of six Multi-Layer Perceptron (MLP) layers. The learning rate for the PURL was established at $1e^{-4}$, paired with a batch size of 2000. Conducted on an NVIDIA A100 GPU, our experiments spanned over 100 epochs to ensure comprehensive training.

Table 3: Results on US court views generation. **Best** and second values are both highlighted.

Method	ROUGE(%)			BLEU(%)		
	R-1	R-2	R-L	B-1	B-2	B-N
BM25	30.8	16.1	28.2	20.1	13.9	10.9
AC-NLG	35.8	18.5	32.7	23.6	16.2	12.7
Llama2-13B-ST	37.5	24.3	37.5	20.8	14.6	11.4
FT-Llama2-13B-ST	39.7	26.5	41.8	24.3	17.4	13.2
GPT4.0-ST	60.3	42.7	57.1	27.2	18.2	14.1
D3LMw/oPU	51.5	31.3	51.5	35.1	25.8	21.5
D3LMw/oLMRC	50.0	27.9	44.1	35.6	27.4	22.7
D3LMw/oRF	54.1	31.9	54.1	39.6	29.9	25.1
D3LM	63.3	53.1	59.2	38.7	31.7	26.9

4.3 Comparisons with Baselines

Evaluation Metrics. To evaluate our model, we implemented both automatic and human assessment methods. For automatic evaluation, we used Rouge (**Rouge-1**, **Rouge-2**, and **Rouge-L** for LCS-based statistics) and BLEU scores (**BLEU-1**, **BLEU-2**, and an average of BLEU-1 to BLEU-4, denoted as **BLEU-N**), which are benchmarks in natural language generation tasks. Furthermore, human judgment focused on three aspects: 1) **fluency**, 2) **accuracy**, and 3) **adoptable** of rationales in generating court views, rated on a five-point scale (with 5 being the highest). This evaluation was conducted by three legal professionals, ensuring a comprehensive and expert assessment of the model's performance in producing contextually accurate and legally coherent text. This streamlined approach balances thoroughness with brevity, maintaining the integrity of the evaluation process while ensuring clarity and conciseness.

Table 4: Results of human judgement

MODEL	FLUENT	ACC.	ADOPT
Llama2-13B-ST	4.68	4.32	4.03
GPT4.0-ST	4.74	4.39	4.06
D3LM	4.71	4.48	4.19

Baselines. We implement the following baselines for comparison:

• **BM25** (Robertson and Walker, 1994) is a classic retrieval baseline that identifies the most closely related text content between case fact descriptions and court views' rationale sections. In our study, we treat all court views data (train, dev, test) as a database, retrieving the court views with the highest relevance score to each case's fact description using BM25 as our baseline model.

• **AC-NLG** (Wu et al., 2020)is an attention and counterfactual-based generative model, designed to transform fact descriptions into court views by training with these elements as input and output, respectively.

• **Llama2-13B-ST** (Touvron et al., 2023) is a LLM developed by Meta, which utilizes prompts and fact descriptions as inputs to generate court views responses. (ST means single-turn interaction between user and LLM.)

• **FT-Llama2-13B-ST** is a fine-tuned version of Llama2-13B-ST. We use the training data to fine-tune the model, employing fact descriptions as input and court views as output in the fine-tune process.

• **GPT4.0-ST** (Achiam et al., 2023)is an LLM created by OpenAI, employing prompts and fact descriptions to interactively produce court views responses.

• **D3LMw/oPU** is a variant of our proposed D3LM model that does not incorporate positive-unlabeled learning in its framework.

• **D3LMw/oLMRC** is a variant of our proposed D3LM model excluding the use of language model reading comprehension.

• **D3LMw/oRF** is a variant of our proposed D3LM model that operates without the integration of reinforcement learning.

Experimental Result. In the automatic evaluation detailed in Tab. 3, the D3LM model showcases superior performance in generating U.S. court views, achieving the highest ROUGE and BLEU scores. Notably, GPT4.0-ST stands out as a particularly strong competitor, especially in terms of ROUGE metrics. Furthermore, the variant D3LMw/oRF registers impressively high BLEU scores, highlighting its adeptness at capturing the subtle linguistic qualities of the generated court views. The comparison between D3LM and its ablated versions underlines the significance of each individual module, con-

Original Fact Description	During a robbery on the evening of January 15, 2015, defendant A allegedly used a firearm and shot a pursuing security guard. Surveillance footage captured part of the robbery, but A's presence at the scene has been disputed due to a timestamped hospital record indicating her hospitalization at the time of the incident.
Golden Court view	Given the substantial evidence supporting defendant A's alibi, the court concludes that there is reasonable about regarding defendant A's involvement in alleged crime. Therefore, defendant A is found not guilty of the charges related to the use of a firearm and shooting a pursuing security guard during the robbery on January 15, 2015.
Masked Fact Description	During a robbery on the evening of January 15, 2015, defendant A allegedly used a firearm and shot a pursuing security guard. Surveillance footage captured part of the robbery.
GPT4.0 Court View	Defendant A faces serious charges, including armed robbery and aggravated assault, due to the use of a firearm in the commission of a crime and the subsequent shooting of a security guard.
PURL Node Selection	Alibi
Generated Question	Can you provide any evidence or details that confirm your whereabouts during the time of the incident?
Recovered Fact	A timestamped hospital record.
D3LM Court View	Based on the credible alibi provided, the court finds defendant A not guilty of the charges of armed robbery and aggravated.

Figure 4: Case Study.

firming that every element plays a vital role in the overall success of the model in the complex task of legal text generation.

In the human evaluation, as illustrated by Tab. 4, our D3LM model demonstrates superior performance in accuracy and adaptability, achieving scores of 4.48 and 4.19, respectively, which marginally exceed those of GPT4.0-ST by about 0.1 points in each metric. GPT4.0-ST, however, takes the lead in fluency with a score of 4.74, proving to be a formidable competitor in semantic coherence. This could be attributed, in part, to our model's foundation on the Llama2-13B architecture. Despite this, the results emphasize D3LM's capability to generate precise and contextually adaptable court views, underlining its efficacy in comparison to a notably strong rival in legal text generation.

Table 5: Results of usability testing

Model	Reli.	Sati.	Pref.(%)
GPT4.0-ST	3.03	3.17	37.7
D3LM	**3.58**	**3.65**	**62.3**

4.4 Human Acceptability

To ensure fairness when comparing our model's metrics with those of traditional LLMs, we conducted a usability testing to gauge user experience and perceptions regarding our approach, which involves greater interaction with legal clients and thus, allows our model to gather more case-specific information. We invited 69 participants to evaluate their experiences and opinions on the state-of-the-art LLM, GPT-4.0, versus our model, D3LM, across three key metrics: 1) **Reliability** (Reli.), the perceived reliability of the models; 2) **Satisfaction**

(Sati.), the level of satisfaction with the models; and 3) **Preference** (Pref.), their preference for either model. We employed a five-point scale for both reliability and satisfaction assessments (where 5 represents the highest rating), and participants were asked to indicate a preference between GPT-4.0 and D3LM.

The findings, detailed in Table 5, revealed that on legal tasks, users reported approximately 0.5 points higher reliability and satisfaction scores for D3LM compared to GPT-4.0. Moreover, 62.3% of participants expressed a preference for using D3LM over GPT-4.0. These results underscore the acceptance and appreciation for our model's interactive questioning approach, highlighting D3LM's enhanced accuracy and reliability in legal applications.

4.5 Case Study

Fig. 4 presents a comparison of court views generated by GPT4.0 and the proposed D3LM method for a masked fact description in a specific US criminal case. Upon comparison with the actual court's view, it is evident that our PURL feature within D3LM can precisely recover the fact "Alibi" and systematically generate questions to elicit further information about the case. Consequently, D3LM produces a more accurate court view than GPT4.0.

5 Conclusion and Future Work

Our study introduces D3LM, an innovative approach in legal assistant systems for court view generation, diverging from traditional methods by simulating real lawyer-client interactions. By innovatively employing Diagnostic Analysis through continuous user engagement, we enable the generation of more accurate and case-specific court

views. The introduction of our PURL algorithm significantly enhances the precision of information extraction. Additionally, we contribute a novel English-language dataset for American legal cases, meticulously validated by legal professionals, addressing the scarcity in legal analysis datasets and setting new benchmarks in legal informatics. Future work will explore extending this model to other domains, such as medical and consultancy, leveraging its capabilities for broader application.

Limitations

In our study, we explored the potential of Large Language Models (LLMs) to improve legal assistance through active engagement with clients, prompting for case-specific details. However, our findings underscore several limitations: the PURL algorithm's effectiveness is confined to the criminal cases domain due to its reliance on a specialized knowledge graph, which is not only resource-intensive to create but also limits cross-domain applicability. Moreover, our evaluation was restricted to English language cases, overlooking the model's performance in diverse linguistic contexts-a critical consideration for global legal system applicability. Furthermore, the model demands significant computational and human annotation resources, with its operational speed lagging behind that of existing large models, potentially hindering its practicality in time-sensitive settings. These challenges underscore the need for future research to focus on optimizing computational efficiency, broadening linguistic and domain applicability, and reducing resource dependencies, aiming to enhance the practicality and accessibility of LLMs in legal assistance across a wider range of contexts.

Acknowledgements

References

Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer. 2002. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47:235–256.

Raymond H Brescia, Walter McCarthy, Ashley McDonald, Kellan Potts, and Cassandra Rivais. 2014. Embracing disruption: How technological change in the delivery of legal services can improve access to justice. *Alb. L. Rev.*, 78:553.

Caselaw Access Project. 2024. Caselaw access project data. Online. Accessed: 2024.

Olivier Chapelle and Lihong Li. 2011. An empirical evaluation of thompson sampling. *Advances in neural information processing systems*, 24.

Xuxi Chen, Wuyang Chen, Tianlong Chen, Ye Yuan, Chen Gong, Kewei Chen, and Zhangyang Wang. 2020. Self-pu: Self boosted and calibrated positive-unlabeled training. In *International Conference on Machine Learning*, pages 1510–1519. PMLR.

Yen-Liang Chen, Yi-Hung Liu, and Wu-Liang Ho. 2013. A text mining approach to assist the general public in the retrieval of legal documents. *Journal of the American Society for Information Science and Technology*, 64(2):280–290.

Yuh-Shyan Chen, Shin-Wei Chiang, and Meng-Luen Wu. 2022. A few-shot transfer learning approach using text-label embedding with legal attributes for law article prediction. *Applied Intelligence*, 52(3):2884–2902.

Zeming Chen, Qiyue Gao, Antoine Bosselut, Ashish Sabharwal, and Kyle Richardson. 2023. Disco: distilling counterfactuals with large language models. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5514–5528.

Wei Chu, Lihong Li, Lev Reyzin, and Robert Schapire. 2011. Contextual bandits with linear payoff functions. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pages 208–214. JMLR Workshop and Conference Proceedings.

Jiaxi Cui, Zongjian Li, Yang Yan, Bohua Chen, and Li Yuan. 2023. Chatlaw: Open-source legal large language model with integrated external knowledge bases. *arXiv preprint arXiv:2306.16092*.

Marthinus Du Plessis, Gang Niu, and Masashi Sugiyama. 2015. Convex formulation for learning from positive and unlabeled data. In *International conference on machine learning*, pages 1386–1394. PMLR.

Jordan Gans-Morse. 2017. Demand for law and the security of property rights: The case of post-soviet russia. *American Political Science Review*, 111(2):338–359.

Howard Gensler. 1985. Irac: One more time. *Duq. L. Rev.*, 24:243.

Gillian K Hadfield. 2010. Higher demand, lower supply-a comparative assessment of the legal resource landscape for ordinary americans. *Fordham Urb. LJ*, 37:129.

Steven Horwitz. 2020. 4 growth, inequality, and unfairness. *Capitalism and Inequality: The Role of State and Market*, page 61.

Wenpeng Hu, Ran Le, Bing Liu, Feng Ji, Jinwen Ma, Dongyan Zhao, and Rui Yan. 2021. Predictive adversarial learning from positive and unlabeled data. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 7806–7814.

Zikun Hu, Xiang Li, Cunchao Tu, Zhiyuan Liu, and Maosong Sun. 2018. Few-shot charge prediction with discriminative legal attributes. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 487–498.

Hongye Jin, Xiaotian Han, Jingfeng Yang, Zhimeng Jiang, Zirui Liu, Chia-Yuan Chang, Huiyuan Chen, and Xia Hu. 2024. Llm maybe longlm: Self-extend llm context window without tuning. *arXiv preprint arXiv:2401.01325*.

Ryuichi Kiryo, Gang Niu, Marthinus C Du Plessis, and Masashi Sugiyama. 2017. Positive-unlabeled learning with non-negative risk estimator. *Advances in neural information processing systems*, 30.

Wan-Chen Lin, Tsung-Ting Kuo, Tung-Jia Chang, Chueh-An Yen, Chao-Ju Chen, and Shou-de Lin. 2012. Exploiting machine learning models for chinese legal documents labeling, case classification, and sentencing prediction. *Processdings of RO-CLING*, 17(4):140.

Alisa Liu, Swabha Swayamdipta, Noah A Smith, and Yejin Choi. 2022. Wanli: Worker and ai collaboration for natural language inference dataset creation. *arXiv preprint arXiv:2201.05955*.

Yang Liu and Yi-Fang Brook Wu. 2020. Fned: a deep network for fake news early detection on social media. *ACM Transactions on Information Systems (TOIS)*, 38(3):1–33.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Indra Gunawan Purba and Alvi Syahrin. 2019. Demand against law and using authority in corruption criminal action. *Budapest International Research and Critics Institute-Journal (BIRCI-Journal) Vol*, 2(4):194–206.

Stephen E Robertson and Steve Walker. 1994. Some simple effective approximations to the 2-poisson model for probabilistic weighted retrieval. In *SIGIR'94: Proceedings of the Seventeenth Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval, organised by Dublin City University*, pages 232–241. Springer.

Timo Schick, Jane Dwivedi-Yu, Roberto Dessì, Roberta Raileanu, Maria Lomeli, Luke Zettlemoyer, Nicola Cancedda, and Thomas Scialom. 2023. Toolformer: Language models can teach themselves to use tools. *arXiv preprint arXiv:2302.04761*.

Swabha Swayamdipta, Roy Schwartz, Nicholas Lourie, Yizhong Wang, Hannaneh Hajishirzi, Noah A Smith, and Yejin Choi. 2020. Dataset cartography: Mapping and diagnosing datasets with training dynamics. *arXiv preprint arXiv:2009.10795*.

Zekun Tong, Yuxuan Liang, Changsheng Sun, Xinke Li, David Rosenblum, and Andrew Lim. 2020. Digraph inception convolutional networks. *Advances in neural information processing systems*, 33:17907–17918.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. 2023. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*.

Huihui Wei and Ming Li. 2018. Positive and unlabeled learning for detecting software functional clones with adversarial training. In *IJCAI*, pages 2840–2846.

Yang Wu, Xurui Li, Xuhong Zhang, Yangyang Kang, Changlong Sun, and Xiaozhong Liu. 2023. Community-based hierarchical positive-unlabeled (pu) model fusion for chronic disease prediction. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, pages 2747–2756.

Yiquan Wu, Kun Kuang, Yating Zhang, Xiaozhong Liu, Changlong Sun, Jun Xiao, Yueting Zhuang, Luo Si, and Fei Wu. 2020. De-biased court's view generation with causality. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 763–780.

Chaojun Xiao, Xueyu Hu, Zhiyuan Liu, Cunchao Tu, and Maosong Sun. 2021. Lawformer: A pre-trained language model for chinese legal long documents. *AI Open*, 2:79–84.

Hai Ye, Xin Jiang, Zhunchen Luo, and Wenhan Chao. 2018. Interpretable charge predictions for criminal cases: Learning to generate court views from fact descriptions. *arXiv preprint arXiv:1802.08504*.

JD Zamfirescu-Pereira, Richmond Y Wong, Bjoern Hartmann, and Qian Yang. 2023. Why johnny can't prompt: how non-ai experts try (and fail) to design llm prompts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–21.

Yunrui Zhao, Qianqian Xu, Yangbangyan Jiang, Peisong Wen, and Qingming Huang. 2022. Dist-pu: Positive-unlabeled learning from a label distribution perspective. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14461–14470.

Haoxi Zhong, Zhipeng Guo, Cunchao Tu, Chaojun Xiao, Zhiyuan Liu, and Maosong Sun. 2018. Legal judgment prediction via topological learning. In *Proceedings of the 2018 conference on empirical methods in natural language processing*, pages 3540–3549.

Dongruo Zhou, Lihong Li, and Quanquan Gu. 2020. Neural contextual bandits with ucb-based exploration. In *International Conference on Machine Learning*, pages 11492–11502. PMLR.

Yao Zhou, Jianpeng Xu, Jun Wu, Zeinab Taghavi, Evren Korpeoglu, Kannan Achan, and Jingrui He. 2021. Pure: Positive-unlabeled recommendation with generative adversarial network. In *Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*, pages 2409–2419.

A   Appendix

A.1   node attention

Given a masked case description embedding h_i and its corresponding subgraph node embeddings H_i. We conduct node embedding as following:

$$e_{ij} = \sigma \left(W \cdot \left[h_i, H_i^j \right] + b \right) \tag{5}$$

$$\alpha_{ij} = \frac{\exp \left(\sigma \left(W \cdot \left[h_i, H_i^j \right] + b \right) \right)}{\sum_{k \in H_i} \exp \left(\sigma \left(W \cdot \left[h_i, H_i^k \right] + b \right) \right)} \tag{6}$$

$$Z_i = \sum_{j \in H_i} \alpha_{ij} h_{ij} \tag{7}$$

where σ denotes $Relu$ activation function, $[.,.]$ represents embedding concatenation. Furthermore, we concatenate the Z_i with each fact node embedding H_i^j, and pass through a MLP to obtain output probability as:

$$P \left(h_i, H_i^j \right) = \phi \left(\left[Z_i, H_i^j \right] \right) \tag{8}$$

A.2   Prompt details

A.2.1   Case type classification

"Please select the most appropriate category from the following options: Contract Law, Criminal Law, Property Law, Intellectual Property Law, Business Law, Tax Law, Estate and Trust Law, Family Law, Administrative Law, Civil Law, Tort Law, Bankruptcy Law, Environmental Law. Provide the category name that best fits the case described below."

A.2.2   IRAC summarization

"Based on the case description, please generate legal case IRAC (Issue, Rule, Analysis, Conclusion) in following pattern:
Issue:
Identifying the legal question or questions that need to be addressed in the legal case.
Rule:

Stating the applicable legal rules, statutes, regulations, or case law that pertain to the identified issue.
Analysis:
Applying the relevant legal rules to the facts or circumstances of the case. This involves explaining how the rules should be interpreted and applied.
Conclusion:
Summarizing the court judgement to the legal issue."

A.2.3   Fact-Rule graph extraction

"I'm developing a project to analyze legal cases deeply by distilling case texts into concise formats for graphical representation. Help me extract key facts influencing court decisions, the applicable rules, and the relationships between them without considering the decision's outcome. Facts should be generalized (e.g., "Contract", "Damages") and rules brief (e.g., "Civil Law", "Contract Law"). Edges should either show a fact's dependency on another fact ("Depends On") or its compliance with or violation of a rule ("Complies With"/"Violates"). Provide three Python lists for networkx: facts, rules, and edges, with edges formatted as ('Fact', 'Rule', 'relation': 'Relation Type'). Ensure edges accurately reflect the connections without adding nonexistent links. Respond in English with only the lists in a single Python chunk, without extra text or explanation."

A.2.4   Case reconstruction

"I have a detailed description of a legal case, including the Issue, Rule, and Analysis of the case. I would like you to help me rewrite this case's IRA, but please disregard the specific aspects that I will mention below."

A.2.5   LLM reading comprehension

Question generation:
"I will provide a conclusion of a legal case, please generate 10 unique questions and corresponding answers based on the conclusion information. Please do not give me explanation of the question, greeting words and any annotation words, just give me generated questions."

  Scoring:
"I will provide a conclusion of a legal case, and 10 unique questions and golden answers, please use the given conclusion to answer the 10 questions, if the conclusion can correctly answer one question, it will get one score, the maximum score is 10, please tell me the final score of the conclusion:"

A.3 Comparison between PU and PN

To verify the effectiveness of our positive-unlabel module, we conducted a comparative experiment between the PU (Positive-Unlabeled) approach and the PN (Positive-Negative) approach. In the PN setting, all unlabeled candidate nodes are treated as negative. The results of the experiment are presented in Table 6. The accuracy, recall, precision, F1, F2, and AUC scores for PU are significantly higher than those for PN.

Table 6: Positive-Unalbeled vs Positive-Negative

Model	Acc.	Rec.	Pre.	F1	F2	AUC
PN	65.4	54.9	78.8	42.3	51.1	90.9
PU	**87.3**	**86.1**	**86.8**	**86.4**	**86.3**	**92.8**

A.4 CVG dataset generation algorithm

Algorithm 2 CVG Dataset Generation via LLM

Input: Legal cases set X
Output: CVG dataset $\{D, C, \overline{D}, Q, F, \hat{G}, G', \overline{G}, G\}$
1: Initialize $D, C, \overline{D}, Q, F, \hat{G}, G', \overline{G}, G$ are all as empty lists
2: **for** $i = 1, 2, \ldots, |X|$ **do**
3:   $G_i = LLM_{\text{extract}}(X_i)$
4:   D_i , $C_i = LLM_{\text{summarize}}(X_i)$
5:   $Q_i = LLM_{\text{generate}}(C_i)$
6:   $\overline{G_i} = Masked(G_i)$
7:   $\overline{D_i} = LLM_{\text{reconstruct}}(D_i)$
8:   $D \leftarrow D + D_i$ $C \leftarrow C + C_i$
   $\overline{D} \leftarrow \overline{D_i} + D_i$ $Q \leftarrow Q + Q_i$
   $\hat{G} \leftarrow \hat{G} + G_i$ $\overline{G_i} \leftarrow \overline{G} + \overline{G_i}$
9: **end for**
10: $G = Merge(\hat{G})$
11: **for** $i = 1, 2, \ldots, |X|$ **do**
12:   $G'_i = G_{\text{N-hop}}(\overline{G_i})$
13:   $F_i = Extract(G'_i)$
14:   $G' \leftarrow G' + G'_i$ $F \leftarrow F + F_i$
15: **end for**
16: Human_Check $\{D, C, \overline{D}, Q, F, \hat{G}, G', \overline{G}, G\}$
17: **Return** LLM-Generated CVG dataset $\{D, C, \overline{D}, Q, F, \hat{G}, G', \overline{G}, G\}$

A.5 Domain PU model training

Algorithm 3 Train_Domain_PU

Input: Training cases fact descriptions set $\overline{D}$; N-hop subgraphs set G'; positive class priors set π; training epochs E; step size η; discount factor γ
Output: Domain_pu model parameters θ
1: Randomly initialize θ
2: Let A be an external SGD-like stochastic optimization algorithm such as Adam (reference)
3: **for** $e = 1, 2, \ldots, E$ **do**
4:   **for** $i = 1, 2, \ldots, |\overline{D}|$ **do**
5:    $h_i = Roberta(\overline{D_i})$
6:    $H_i = DiGCN(G'_i)$
7:    Denote positive instances in H_i as x_p^i, unlabeled instances in H_i as x_u^i
8:    **if** $\widehat{R}_u^-\left(g;\ x_u^i\right) - \pi_p \widehat{R}_p^-\left(g;\ x_p^i\right) \geq 0$ **then**
9:     Set gradient $\nabla_\theta \widehat{R}_{pu}\left(g;\ x_p^i, x_u^i\right)$
10:     Update θ by A with current step size η
11:    **else**
12:     Set gradient $\nabla_\theta(\pi_p \widehat{R}_p^-\left(g;\ x_p^i\right) - \widehat{R}_u^-\left(g;\ x_u^i\right))$
13:     Update θ by A with discounted step size $\gamma\eta$
14:    **end if**
15:   **end for**
16: **end for**
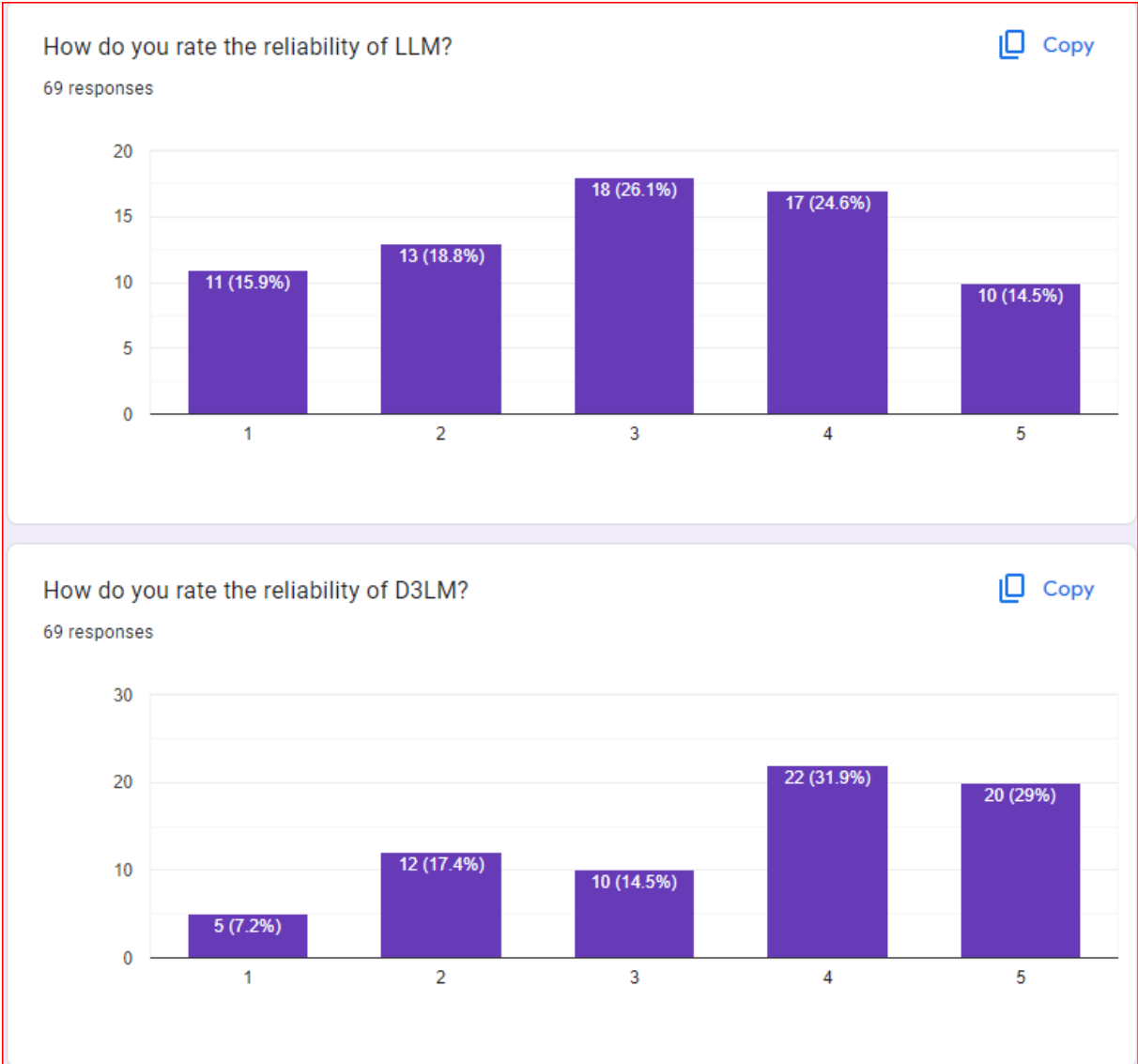17: **Return** Domain_pu model with updated parameters θ
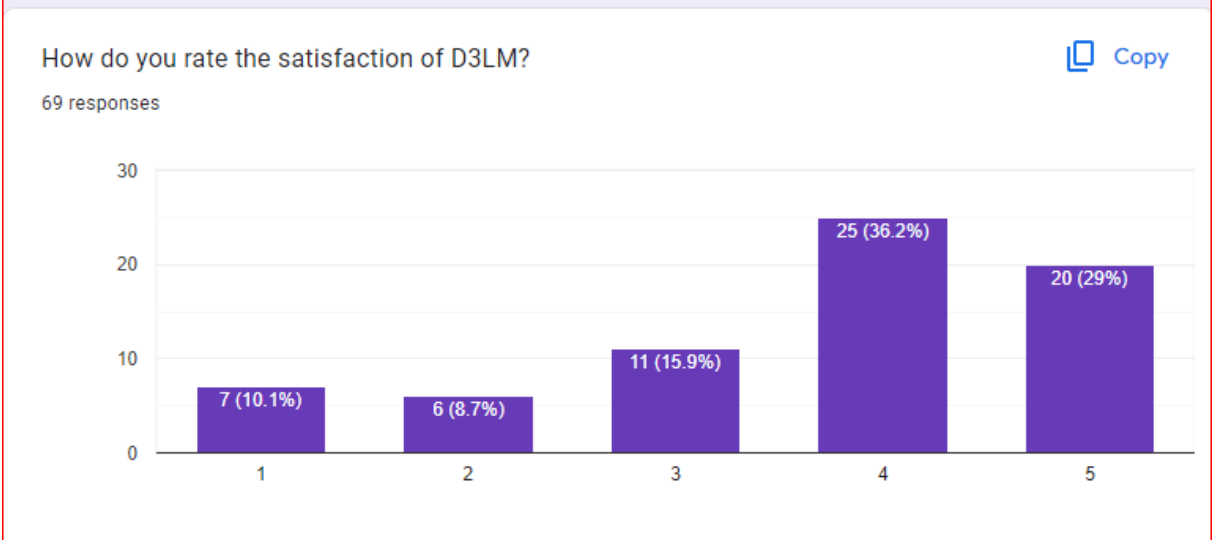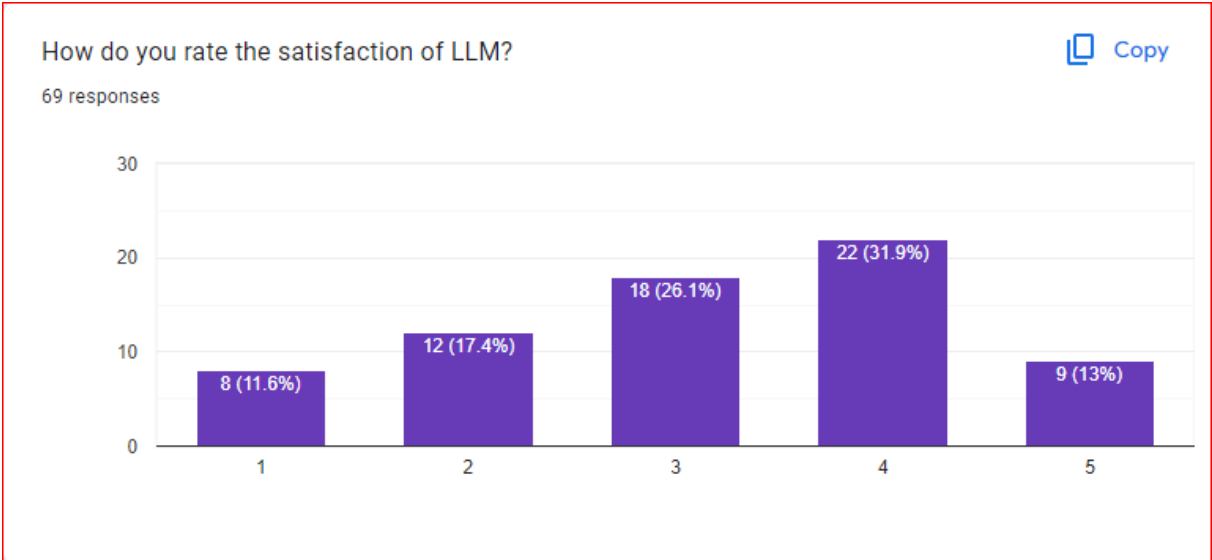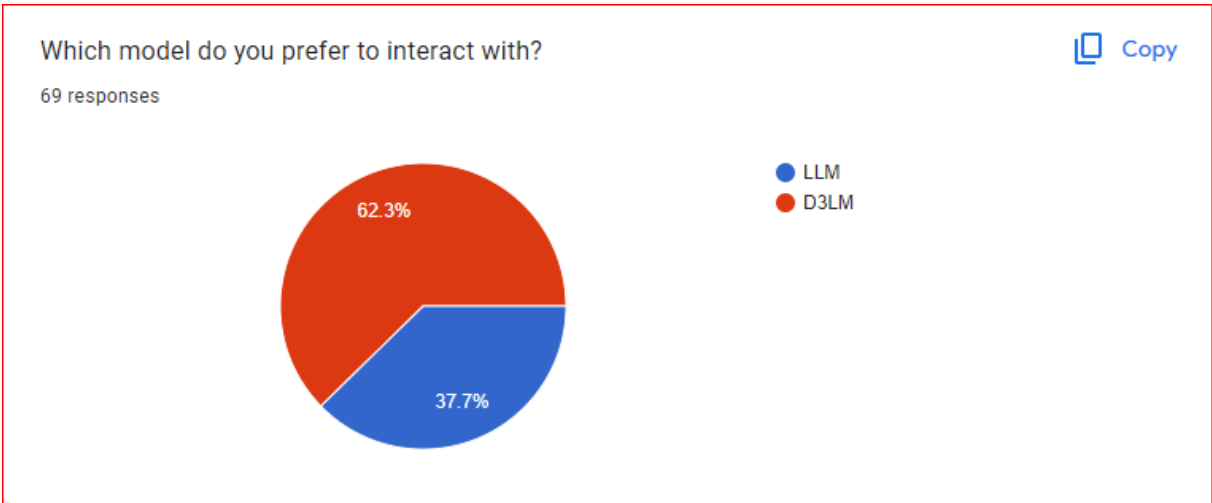

15553

Figure 5: Reliability result

Figure 6: Satisfaction result

Figure 7: Preference result