Interpretable Graph-Language Modeling for Detecting Youth Illicit Drug Use

Anonymous ACL submission

Abstract

Illicit drug use among teenagers and young adults (TYAs) remains a pressing public health concern, with rising prevalence and long-term impacts on health and well-being. To detect illicit drug use among TYAs, researchers analyze large-scale surveys such as the Youth Risk Behavior Survey (YRBS) and the National Survey on Drug Use and Health (NSDUH), which preserve rich demographic, psychological, and environmental factors related to substance use. However, existing modeling methods treat survey variables independently, overlooking latent and interconnected structures among them. To address this limitation, we propose **LAMI** (*LAtent relation Mining with bi-modal Interpretability*), a novel joint graph-language modeling framework for detecting illicit drug use and interpreting behavioral risk factors among TYAs. LAMI represents individual responses as relational graphs, learns latent connections through a specialized graph structure learning layer, and integrates a large language model to generate natural language explanations grounded in both graph structures and survey semantics. Experiments on the YRBS and NSDUH datasets show that LAMI outperforms competitive baselines in predictive accuracy. Interpretability analyses further demonstrate that LAMI reveals meaningful behavioral substructures and psychosocial pathways, such as family dynamics, peer influence, and school-related distress, that align with established risk factors for substance use. Our codebase is available here.

1 Introduction

Illicit drug use (National Center for Health Statistics, 2024) represents a persistent and escalating public health challenge in the United States, contributing to significant social and health burdens. It includes both the consumption of illegal substances, such as cannabis, cocaine, and hallucinogens, and the misuse of prescription medications. Teenagers and young adults (TYAs) are particularly at risk (Simon et al., 2022), exhibiting disproportionately high prevalence rates and associated harms. According to the 2024 *National Survey on Drug Use and Health* (NSDUH), 38.1% of young adults aged 18–25 and 15.1% of adolescents aged 12–17 reported illicit drug use (SAMHSA, 2025a). These alarming statistics highlight an urgent need to focus on TYAs to better understand the behavioral and psychosocial risk factors that contribute to illicit drug use, thereby informing more effective detection and prevention strategies.

Understanding why TYAs are particularly vulnerable is critical for developing effective detection and prevention strategies. Biologically, during adolescence and emerging adulthood, the brain's reward system matures earlier than the executive control system, heightening reward sensitivity while limiting impulse regulation (Steinberg, 2008; Luciana, 2013). This developmental imbalance increases susceptibility to risk-taking behaviors, including illicit drug use (Geier, 2013; Romer et al., 2017). At the same time, TYAs experience major life transitions - such as leaving home, pursuing higher education, or entering the workforce -while coping mechanisms and identity formation are still developing (Arnett, 2000; Patton et al., 2016). Together, these interacting biological, psychological, and social factors contribute to elevated risk and underscore the need to identify the key mechanisms driving substance use vulnerability among TYAs.

To leverage these multifaceted risk factors for detecting illegal drug use among TYAs, researchers have long relied on large-scale surveys, such as the *Youth Risk Behavior Survey* (YRBS) (CDC, 2025) and the *National Survey on Drug Use and Health* (NSDUH) (SAMHSA, 2025b), that capture diverse demographic, psychosocial, behavioral, and environmental variables. Traditional approaches ranging from statistical models to conventional machine learning techniques (Wadekar, 2019; Han et al.,

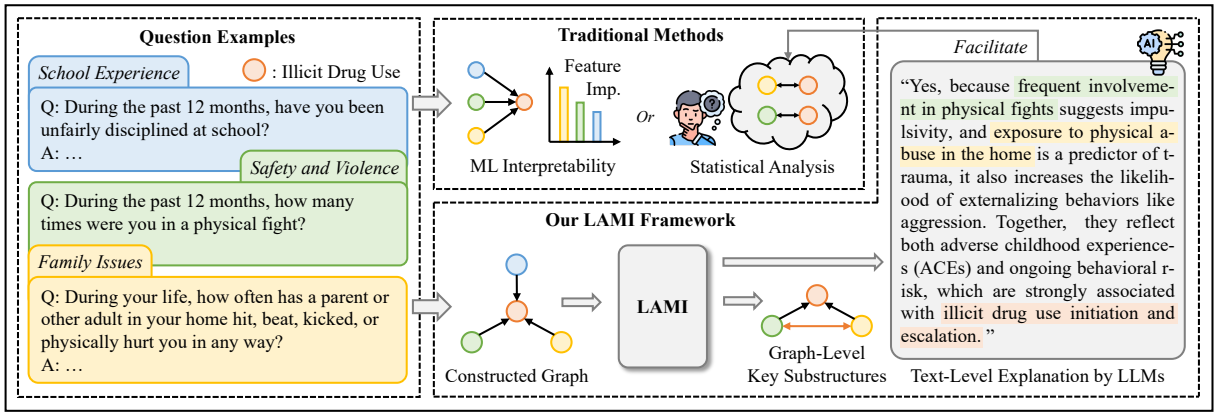

Figure 1: **Comparison between traditional methods and our LAMI for illicit drug use detection.** Traditional approaches typically treat survey variables in isolation, relying on feature importance metrics or expert-selected variables, which overlook the interdependent relationships among behavioral and psychosocial factors. In contrast, LAMI models survey responses as relational graphs, enabling the discovery of latent structural risk factors and facilitating deeper, more interpretable behavioral insights through graph-level analysis and LLM-based explanations.

2020; Thompson et al., 2021; McKinnon, 2024; Kim et al., 2025) have offered valuable insights, but they typically treat survey variables independently. This assumption overlooks the complex and interdependent relationships that shape behavioral outcomes, thereby obscuring latent structures that may reveal deeper mechanisms of risk. As illustrated in Figure 1, uncovering these latent relations can enable a more holistic understanding of how different factors interact and provide new pathways for explaining and detecting TYAs illicit drug use.

Building on this motivation, our work aims to model the relational information embedded within survey data to uncover the underlying behavioral and psychosocial structures associated with illicit drug use. However, capturing these relations presents several key challenges. First, survey data are not inherently relational, making it difficult to represent variable-level dependencies in a suitable form. Second, the latent relations among variables are not explicitly observed and must be inferred through specialized learning methods. Third, predictive frameworks must remain interpretable: able to identify meaningful substructures and generate explanations that support researchers and practitioners in public health and behavioral science.

To address these challenges, we propose **LA**tent relation **M**ining with bi-modal **I**nterpretability (**LAMI**), a novel framework that integrates graph neural networks (GNNs) (Khemani et al., 2024) with large language models (LLMs) (Zhao et al., 2023) for illicit drug use detection among TYAs. Specifically, LAMI (1) innovatively represents each individual's survey responses as a relational graph, (2) learns latent relations among question nodes through a relational graph structure learning layer and a self-supervised pretext task, and (3) enhances interpretability by coupling GNN-derived substructure analysis with LLM-generated natural language explanations. We evaluate LAMI on the YRBS and NSDUH, demonstrating its superior predictive performance and interpretability.

2 Related Works

Illicit Drug Use Prediction. Traditional statistical and machine learning approaches have long dominated the prediction of illicit drug use (Unlu and Subasi, 2025). Valued for their simplicity and interpretability (King et al., 2017), these methods have identified key behavioral and psychosocial predictors, including smoking, alcohol use, mental health indicators (e.g., depression and anxiety), and suicidal ideation (Wadekar, 2019; Han et al., 2020; Islam et al., 2023; Rakovski et al., 2023; Kim et al., 2025). Such findings have contributed substantially to public health research by informing prevention and intervention strategies. More recently, deep learning models have been applied to unstructured data sources such as electronic health records (EHR) (Dong et al., 2021; Kashyap et al., 2023), dietary patterns (Zhang et al., 2024b), and social media (Dou et al., 2021; Ma et al., 2025), demonstrating improved predictive accuracy and potential for digital surveillance. However, their use in large-scale population surveys remains limited, leaving research gap for modeling the complex relational dependencies that underlie behavioral and psychosocial factors in such data.

Graph Neural Networks and Structure Learning. GNNs (Kipf and Welling, 2017; Hamilton et al., 2017; Veličković et al., 2017) have emerged as powerful tools for modeling structured data in which relationships among entities play a central role (Wu et al., 2020). Recent work has extended GNNs to tabular settings, where explicit graph structures are not readily available (Goodge et al., 2022; Du et al., 2022; Telyatnikov and Scardapane, 2023; Yan et al., 2023). However, their effectiveness critically depends on the quality of the manually defined graph structure - poorly constructed graphs can substantially degrade performance (Zhu et al., 2021). To mitigate this issue, Graph Structure Learning (GSL) approaches have been developed to learn or refine graph topologies during training (Franceschi et al., 2019; Jin et al., 2020; Zhao et al., 2021; Liu et al., 2022). Nonetheless, most GSL methods do not explicitly account for the context of node attributes when modeling structural relationships, thereby limiting the ability to capture the nuanced inter-variable dependencies inherent in behavioral and psychosocial survey data.

Enhancing GNN Interpretability with LLMs. Recent research has explored integrating LLMs with GNNs to enhance interpretability, which can be broadly categorized into three paradigms. The first leverages LLMs to augment existing GNN explainers, where the LLM verbalizes, refines, or contextualizes explanations generated by traditional explainer modules (Cedro and Martens, 2024; Zhang et al., 2024a; Pan et al., 2024). The second employs LLMs purely as external interpreters, providing post-hoc textual explanations without participating in model training or gradient propagation (Zhang et al., 2024b; He et al., 2024). The third paradigm achieves tighter coupling by injecting graph embeddings into LLM prompts as prefix tokens, aligning GNN and LLM representations and enabling the LLM to condition its reasoning on graph-encoded structural information (Bagher-shahi et al., 2025). This deeper integration supports joint structural–semantic reasoning, which is particularly well-suited to our setting, where survey variables carry rich semantic meaning and their latent relations are essential for uncovering key psychosocial and behavioral substructures.

3 Methodology

Notations. Let $G = (V, E, \mathcal{T}, \mathcal{R})$ denote a relational graph, where V is the set of nodes, $E \subseteq V \times V$ is the set of edges, $\mathcal{T}$ is the set of node types, and $\mathcal{R}$ is the set of edge types. Let $X \in \mathbb{R}^{|V| \times d}$ be the node feature matrix.

Problem Definition. Given a relational graph G constructed for a single TYA respondent, the objective is to learn a predictive function f that determines whether the individual engages in illicit drug use. To enhance interpretability, f further leverages an LLM to generate natural language explanations grounded in the learned graph representations.

3.1 Relational Graph Construction

Attribute	YRBS	NSDUH
# Graphs	19,931	21,510
# Positives	6,985	7,005
# Negatives	12,964	14,505
# User Nodes Per Graph	1	1
# Question Nodes Per Graph	88	349
# Topic Nodes Per Graph	17	10
# Unique Relations Per Graph	466	1,079

Table 1: Statistics of the constructed datasets.

We utilize two publicly available datasets, the 2023 releases of the YRBS and the NSDUH, to construct individual-level relational graphs based on their official documentation. Each respondent is represented by a graph G comprising a user node, a set of question nodes linked to the user node, and a set of topic nodes connected to their corresponding question nodes, where $\mathcal{T} = \{\text{User}, \text{Question}, \text{Topic}\}$. The edge type between a user node and a question node encodes the respondent's answer to that question. The user node features capture demographic characteristics, while question and topic node features are initialized using BERT-based embeddings (Devlin et al., 2019) of their textual descriptions. To align with our focus on TYAs, we restrict the sample to respondents aged 15–25. Summary statistics of the constructed datasets are provided in Table 1, and additional details are included in Appendix A.

3.2 Graph Latent Relation Learning

Although many survey questions appear unrelated in isolation, they often reflect intertwined psychosocial factors that jointly lead to illicit drug use. For example, *exposure to physical abuse in the home* may increase a respondent's propensity for aggression and, consequently, a higher likelihood of *frequent physical fights*. Taken together, these

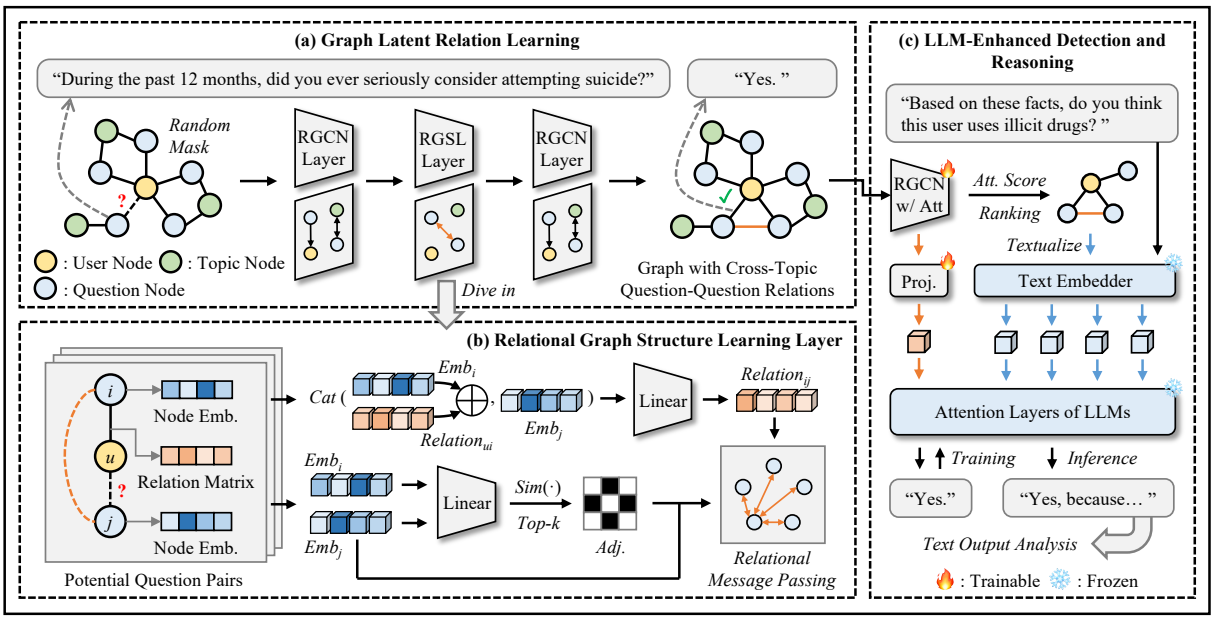

Figure 2: **Overview of the LAMI framework.** (a) Graph latent relation learning for capturing implicit dependencies among survey variables, (b) a relational graph structure learning (RGSL) layer for inferring latent question–question connections, and (c) an LLM-enhanced predictor for detecting illicit drug use and generating text-based explanations. These components enable interpretable illicit drug use detection among TYAs.

items jointly signal adverse childhood experiences (ACEs), which are strongly associated with drug use initiation. Uncovering such latent relations is therefore essential for both accurate prediction and interpretable analysis. Thus, the goal of this module is to refine the graph structure by discovering latent relations among nodes.

To model these hidden dependencies, we design a framework capable of discovering new relational structures while simultaneously performing representation learning. Specifically, our model stacks standard Relational Graph Convolutional Network (RGCN) layers with a *Relational Graph Structure Learning (RGSL)* layer (Fig. 2a). The RGCN layers follow the standard message-passing formulation (Schlichtkrull et al., 2018):

$$h_i^{(l+1)} = \sigma\Big(\sum_{r\in\mathcal{R}}\sum_{j\in\mathcal{N}_i^r}\frac{1}{c_{i,r}}W_r^{(l)}h_j^{(l)} + W_0^{(l)}h_i^{(l)}\Big), \tag{1}$$

where $\sigma(\cdot)$ is the activation function, $W_r^{(l)}$ is the learnable weight matrix for relation r at layer l, $h_i^{(l)}$ denotes the embedding of node i, $\mathcal{N}_i^r$ represents the set of r-type neighbors, and $c_{i,r}$ is a normalization constant. The RGCN layers update node embeddings and learn user–question relations, while the RGSL layer leverages these learned relations as contextual signals to infer additional latent edges.

RGSL Layer. The respondent's answer to question i often influences how question i relates to another question j. Motivated by this observation, the RGSL layer is designed to induce context-dependent latent relations among question nodes. Given a target question node i whose user–question edge (u, j) is masked during training, we consider each other question node j from a different topic. The layer (Fig. 2b) (i) learns an adjacency matrix A among question nodes and (ii) computes *relation vectors* that inject user–question relation into cross-question interactions, before performing message passing over the learned structure.

Adjacency matrix learning. We first project question embeddings $h_i \in \mathbb{R}^d$ with a linear transformation $W_a \in \mathbb{R}^{d\times d}$ and compute pairwise scores $s_{ij} = (W_a h_i)^\top (W_a h_j)$ between any pairs of questions. Then select the top-k_{sim} neighbors for each i to form a sparse adjacency $A \in \{0,1\}^{Q\times Q}$ (with Q questions). A hard mask is used in the forward pass, while gradients flow through s_{ij}. Additionally, we incorporate a penalty into the variance of the nodes' degree to incentivize the model to discover more varied connections.

Relation vectors. Let r denote the known edge type between the user node u and question node i, and let $W_r \in \mathbb{R}^{d\times d}$ be the corresponding relation matrix produced by the preceding RGCN layer. For

each selected node j, we compute a relation vector

$$r_{ij} = \sigma\Big(W_r^{\mathrm{Rel}}\Big[\Big(h_i + \frac{1}{d}\sum_{k=1}^{d}(W_r)_k\Big) \Big\| h_j\Big]\Big), \quad (2)$$

where $W_r^{\mathrm{Rel}} \in \mathbb{R}^{2d\times d}$ is a trainable linear layer and $[\cdot\|\cdot]$ denotes concatenation. Injecting W_r into r_{ij} conditions the cross-question relation on the user–question relation context.

Message passing. We update node i by aggregating relation vectors from its learned neighbors:

$$h_i' = \sigma\Big(\sum_j \hat{A}_{ij}\,(r_{ij} \odot h_j) + W_s h_i\Big), \quad (3)$$

where $\hat{A}$ is the row-normalized A, $\odot$ is element-wise multiplication, and $W_s \in \mathbb{R}^{d\times d}$ is a residual transformation. This equips the model to pass information only along newly discovered cross-topic edges that reflect latent relations.

Self-Supervised Edge Type Prediction. To drive the discovery of latent relations, we adopt a self-supervised pretext task that predicts masked user–question edge types. For each graph, we randomly mask one edge (u, q) (e.g., the user's response to *physical fight*) and train the model to infer its type from the remaining graph $G_{\setminus(u,q)}$, thereby encouraging it to capture dependencies such as the link between *exposure to physical abuse* and *physical fight*. Formally, letting $y^r_{(u,q)} \in \{0, 1\}$ denote the true edge type and $P_r(u, q \mid G_{\setminus(u,q)})$ the predicted probability, we minimize

$$\mathcal{L}_{\text{edge}} = -\sum_{r\in\mathcal{R}} y^r_{(u,q)}\, \log P_r\big(u, q \mid G_{\setminus(u,q)}\big)\,. \quad (4)$$

During this pretext task, we restrict message flow through the user node: information from other questions cannot aggregate into u to inform the prediction. Consequently, the model must rely on the RGSL-induced question–question edges to route information to the masked target, ensuring that improvements only stem from *newly learned* latent structure. After this stage, we update the graph structure according to the learned adjacency matrix A, forming an enriched relational graph.

3.3 LLM-Enhanced Detection and Reasoning

To identify key substructures and gain human-readable insights related to the detection results, we develop a bi-modal module that combines graph-based and language-based reasoning. As illustrated in Figure 2 (c), the integration of GNN and LLM enables both structure-aware prediction and natural language explanation.

We first apply a standard RGCN model to compute node embeddings over the updated graph induced from the previous section: $\mathbf{H} = \mathrm{RGCN}(G) \in \mathbb{R}^{|V|\times d}$. To aggregate the information relevant to the user node, we introduce an attention mechanism on the question nodes, where the attention scores reflect the contribution of each node $\alpha_{i,u} = \mathrm{softmax}_{i\in\mathcal{N}_u}\big(\mathrm{MLP}\big([\mathbf{h}_i \,\|\, \mathbf{h}_u]\big)\big)$, with $\mathbf{h}_i, \mathbf{h}_u \in \mathbb{R}^d$ and a learnable scorer MLP. The user embedding is updated via attention-weighted aggregation:

$$\mathbf{h}_u^{\text{agg}} = \sum_{i\in\mathcal{N}_u} \alpha_{i,u} \cdot \mathbf{h}_i, \quad (5)$$

Textualization of Key Substructures. Based on $\alpha_{i,u}$, we select the top-k_{att} question nodes and convert them to text using the original question wording and the user's response. We also provide cues about learned interdependencies among selected questions. A short prompt then asks the LLM to predict whether the user uses illicit drugs. The resulting tokens are embedded by the LLM to obtain text embeddings $\mathcal{Z} = [\mathbf{z}_{t_1}, \ldots, \mathbf{z}_{t_k}]$.

Bi-Modal Optimization. To align modalities, we project the graph-based user embedding into the LLM embedding space via a learnable projection head $\mathbf{z}_u = P_{\text{proj}}(\mathbf{h}_u^{\text{agg}})$. We prepend this projected graph token to the textual embeddings and feed the sequence into a frozen LLM to generate an output sequence $Y = (y_1, \ldots, y_r)$:

$$p_{\theta,\phi}(Y|\mathcal{Z}, \mathbf{z}_u) = \prod_{i=1}^{r} p_{\theta,\phi}(y_i|y_{<i},\, [\mathbf{z}_u; \mathcal{Z}])\,, \quad (6)$$

where θ denotes frozen LLM parameters and ϕ includes trainable parameters of the graph encoder and P_{proj}. The concatenation $[\mathbf{z}_u; \mathcal{Z}]$ forms the full prompt, with $\mathbf{z}_u$ acting as a soft graph-conditioned prefix. In parallel, we feed $\mathbf{h}_u^{\text{agg}}$ to a lightweight classifier for illicit drug use prediction.

Let $y \in \{0, 1\}$ be the ground-truth label and $\hat{y}$ the classifier output. We supervise the LLM with a verbalized label token (e.g., `Yes`/`No`) at the beginning of Y. The total loss combines a generation loss and a classification loss:

$$\mathcal{L}_{\text{bi}} = \mathcal{L}_{\text{gen}} + \mathcal{L}_{\text{cls}}, \quad \mathcal{L}_{\text{cls}} = \mathrm{CE}\big(\hat{y}, y\big),$$
$$\mathcal{L}_{\text{gen}} = -\log p_{\theta,\phi}(y \mid \mathbf{z}_u, \mathcal{Z})\,. \quad (7)$$

Method	YRBS					NSDUH				
	Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
LightGBM	80.03	78.17	78.79	78.45	86.70	79.95	77.57	**80.62**	78.35	87.91
XGBoost	79.93	78.21	77.50	77.82	85.59	80.20	77.46	79.75	78.24	87.52
MLP	75.05	72.80	71.56	72.05	76.22	76.73	74.13	69.29	70.61	79.61
TabNet	80.07	78.22	78.43	78.32	78.43	75.36	72.51	74.65	73.12	74.65
RGCN	80.29	78.56	78.74	78.65	86.57	79.29	76.44	78.90	77.24	87.14
RGAT	79.42	77.74	79.07	78.21	86.53	78.36	75.94	79.07	76.64	87.32
Qwen3-8B										
- *Zero-shot*	65.08	70.04	50.12	39.70	50.12	68.23	34.11	53.42	40.61	50.03
- *LoRA*	82.20	78.52	67.71	72.68	78.93	78.51	69.74	57.29	62.90	72.80
LAMI (Ours)	**82.21**	**80.39**	**80.78**	**80.57**	**88.45**	**81.47**	**78.58**	79.29	**78.86**	**88.23**

Table 2: Performance comparison of different methods on YRBS and NSDUH datasets.

Gradients from $\mathcal{L}_{\text{bi}}$ update ϕ only. This dual-objective setup provides optimization signals from both textual space and graph space, enabling GNN-based structural reasoning and LLM-based textual interpretation.

Detection and Explanation. At inference time, we use the graph classifier for illicit drug detection, and additionally prompt the LLM to justify the decision in natural language. This output serves as a post-hoc explanation, enabling qualitative analysis and uncovering interpretable patterns associated with illicit drug use.

4 Experiments

4.1 Experiment Setup

Baselines. We evaluate LAMI by comparing it against a diverse set of baseline models, including traditional machine learning methods (LightGBM (Ke et al., 2017) and XGBoost (Chen and Guestrin, 2016)), neural network models for tabular data (MLP and TabNet (Arik and Pfister, 2021)), relational graph neural networks (RGCN (Schlichtkrull et al., 2018) and RGAT (Busbridge et al., 2019)), and large language models (Qwen3-8B (Yang et al., 2025)) evaluated in both zero-shot and LoRA (Hu et al., 2022) settings.

Notably, traditional machine learning algorithms such as LightGBM and XGBoost remain among the most competitive and robust baselines for learning from tabular data (Shwartz-Ziv and Armon, 2022; Borisov et al., 2022), providing a strong benchmark for comparison.

Evaluation Settings. We evaluate our framework on both the YRBS and NSDUH datasets, using a standard data split of 70% for training, 15% for validation, and 15% for testing. To comprehensively assess model performance across multiple dimensions, we report five widely used evaluation metrics: accuracy, precision, recall, macro-averaged F1-score, and ROC-AUC. All results are averaged over 10 runs with different random seeds to ensure robustness. For LAMI, we adopt Qwen3-0.6B as the LLM backbone to balance performance with computational feasibility, ensuring that the framework can be trained efficiently under available hardware resources. The model is optimized using the Adam optimizer with a learning rate of 5×10^{-5} and a weight decay of 5×10^{-4}. Other key hyperparameters include: hidden dimension = 128, $k_{\text{sim}} = 5$, $k_{\text{att}} = 20$. Additional implementation details are provided in Appendix B. Prompts used are provided in Appendix C

4.2 Main Results

Table 2 presents the performance of LAMI compared to several strong baselines across the YRBS and NSDUH datasets. On both datasets, LAMI consistently outperforms baselines across most evaluation metrics. Experiments show that RGCN and RGAT achieve performance comparable to, or in some metrics even surpassing, that of traditional machine learning methods, whereas MLP and TabNet underperform across the board. This highlights the effectiveness of our graph-based modeling approach in incorporating structural information into survey data analysis. In contrast, the relatively poor performance of the LLM-only model suggests that textual analysis alone is insufficient to uncover the underlying patterns embedded in the data. The strong results achieved by LAMI demonstrate that our framework effectively learns latent structures among nodes in the graph and leverages both struc-

Model Variant	YRBS					NSDUH				
	Accuracy	Precision	Recall	F1-Score	AUC	Accuracy	Precision	Recall	F1-Score	AUC
LAMI	**82.21**	**80.39**	**80.78**	**80.57**	**88.45**	**81.47**	**78.58**	**79.29**	78.86	**88.23**
w/o Relation Matrix	81.43	78.26	79.82	78.63	86.79	80.44	76.33	79.10	77.21	87.56
w/o Latent Relation Learning	80.53	79.34	78.63	79.13	87.29	79.73	77.12	79.05	77.80	87.85
w/o LLM	80.06	79.53	80.44	80.13	87.03	79.80	78.22	79.04	**78.91**	87.74

Table 3: Ablation study of **LAMI**. Each variant removes one key component from the full model.

tural and textual modalities to identify key substructures and make accurate predictions.

4.3 Ablation Study

The Impact of Model Components. Table 3 presents the ablation results of LAMI on both YRBS and NSDUH. Each variant removes one key component from the full model to assess its contribution. Specifically, *LAMI (w/o Relation Matrix)* excludes the contextual relation matrices W_r in the RGSL layer. *LAMI (w/o Latent Relation Learning)* disables the latent relation learning entirely, thereby using only the original survey structure. *LAMI (w/o LLM)* removes the LLM-based interpretability module, leaving only the graph-based classifier for prediction. The results show that each component contributes to overall performance improvement, underscoring the effectiveness of combining latent relation learning, contextualized structure learning, and language-based reasoning.

The Impact of Inferred Relations. As shown in Figure 3, we how two key hyperparameters influence the performance of LAMI on YRBS: k_{sim} and k_{att}. k_{sim} controls how many *new* latent relations the RGSL layer adds for each node during latent relation learning; k_{att} controls how many questions are retained both for the graph-side classifier and for textualization to the LLM. For k_{sim}, accuracy initially increases as k_{sim} grows, indicating that learned latent relations help the model capture useful dependencies that are not present in the original structure. However, performance degrades when too many new relations are added to the graph. This is because the inherent relations among question nodes are limited, and the noise induced by over-connecting the graph hinders the model from identifying key substructures. Increasing k_{att} yields clear gains at smaller values since more informative substructures are provided to both the classifier and the LLM. Beyond approximately $k_{\text{att}} = 20$, the improvement plateaus: the model has already captured the most predictive questions, and additional nodes contribute little to illicit-drug prediction while adding redundant information and longer prompts, leading to diminishing returns.

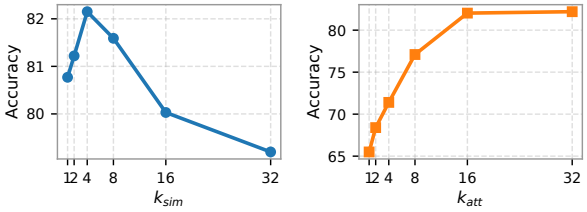

Figure 3: Model performance using different hyperparameter (k_{sim} and k_{att}) settings.

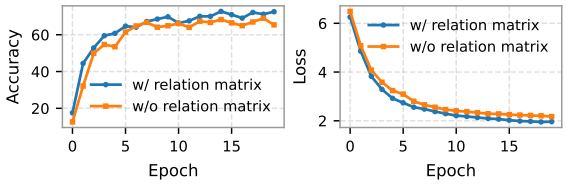

Figure 4: Training loss and model performance on YRBS with and without relation matrix in the latent relation learning stage.

The Impact of Relation Matrix. To further examine the effect of the relation matrix, Figure 4 compares the training dynamics of the latent relation learning stage with and without relation matrices. Incorporating relation matrices leads to faster convergence and higher edge-type prediction accuracy, confirming that contextualized relation modeling provides more stable and informative structural learning.

4.4 Interpretability Analysis

To demonstrate the interpretability of LAMI, we visualize the top 10 questions based on their attention scores toward the user node, as illustrated in Figure 5, and analyze their structural and semantic interrelations. For more detailed case studies on LLM outputs, please refer to Appendix D.

Despite survey design differences between the YRBS and NSDUH, several common themes emerge. Notably, early use of alcohol, nicotine, and electronic vapor products (EVPs) frequently appears among the most influential predictors. This observation is consistent with prior public health literature as well as the "gateway drug effect", showing that early initiation of legal substances is a

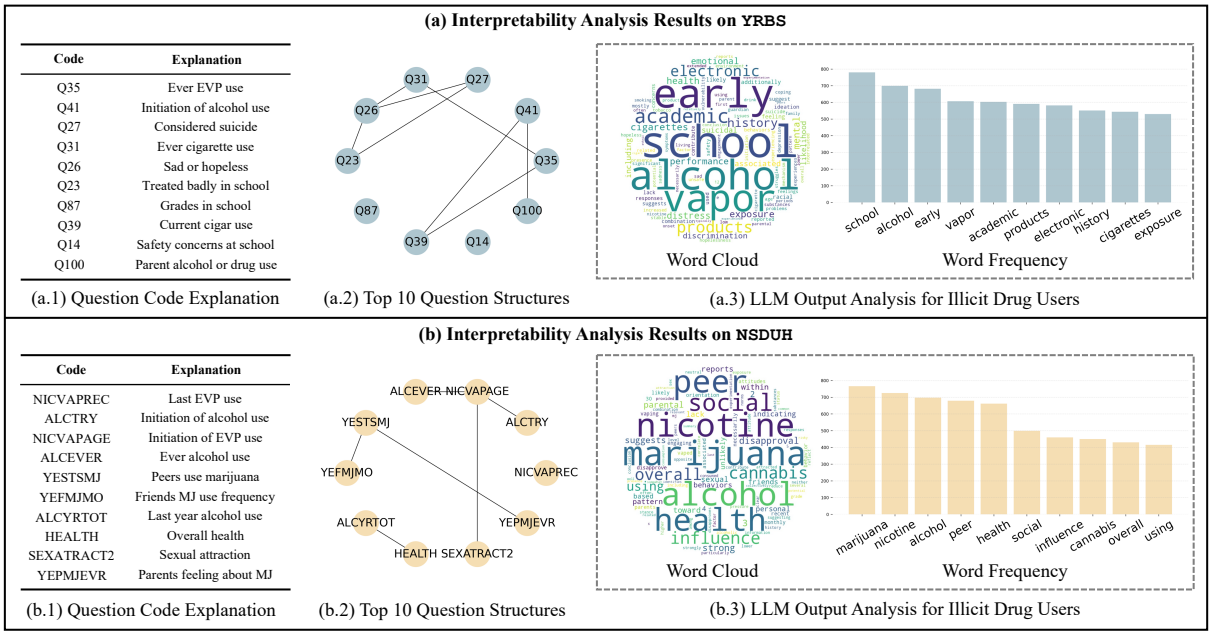

Figure 5: Interpretability analysis of LAMI on YRBS and NSDUH datasets, including (1) the top 10 contributing question nodes, (2) the learned structures among these nodes, and (3) the textual analysis based on the LLM's responses during inference.

significant risk factor for progression to illicit drug use (King and Chassin, 2007; Trenz et al., 2012). Another recurrent theme is the impact of family environment. Parental substance use and negative parenting attitudes are strongly associated with increased likelihood of youth drug use. Previous studies have demonstrated that parental modeling, inconsistent discipline, and lack of emotional support are key risk factors for adolescent substance use (Matejevic et al., 2014; Rusby et al., 2018).

Beyond individual items, LAMI captures meaningful inter-question dependencies. For instance, Q35–Q39 and Q41–Q39 edges reveal connections among different modes of tobacco and alcohol use, underscoring polysubstance patterns (Crummy et al., 2020). More critically, LAMI discovers psychosocial pathways: Q23 (Treated badly at school) links to Q27 (Considered suicide), and both also connect to Q26 (Felt sad or hopeless), forming a cluster that reveals the mental health consequences of school victimization. These psychological states are known precursors to substance use as maladaptive coping strategies (Luk et al., 2010). Q41 (Initiation of alcohol use) connects to Q100 (Parental alcohol or drug use), reflecting the familial modeling of substance behavior (Rusby et al., 2018).

In NSDUH, we observe analogous structures: YESTSMJ (Peer marijuana use) connects to YEPMJEVR (Parental view on marijuana) and YEFMJMO (Friend's marijuana frequency), illustrating the compound social learning influence from both peers and family on youth (Marziali et al., 2022); NICVAPAGE (Initiation of vaping) is linked to SEXATRACT2 (Sexual attraction), possibly reflecting minority stress pathways in LGBTQ+ youth, who exhibit higher substance use due to stigma, isolation, or mental distress (Mann et al., 2021).

In sum, LAMI not only improves predictive performance but also reveals complex interactions between early substance use, family dynamics, school environment, mental health, and peer influence. These insights can inform public health interventions and future research.

5 Conclusion

We introduce LAMI, a novel framework for illicit drug use prediction that models national survey datasets as relational graphs and learns implicit cross-question dependencies to enhance both prediction and interpretability. By combining Graph Neural Networks with Large Language Models, LAMI captures complex, latent structures while producing human-interpretable explanations. Empirical results on the YRBS and NSDUH datasets validate the superiority of our approach over strong baselines. Beyond quantitative performance, LAMI provides qualitative insights into the behavioral and psychosocial factors associated with illicit drug use among teenagers and young adults.

6 Limitations

While LAMI advances illicit drug use detection and interpretability on national surveys, several limitations remain. (i) **Dependence on survey topic design.** Our graph construction leverages survey-defined topic groupings; although these taxonomies are expert-informed, they may bias which cross-item relations are discoverable. Future work will explore data-driven or hybrid topic induction to yield more robust structures. (ii) **LLM capacity constraints.** Due to hardware limits, we employ a relatively small LLM backbone and keep its parameters frozen. Larger models may further improve both accuracy and explanation quality, but at the cost of compute and potential overfitting. (iii) **Temporal scope.** Experiments focus on the 2023 releases of YRBS and NSDUH data. Validating across multiple years and policy environments would better assess robustness to cohort and instrument shifts. (iv) **Faithfulness of explanations.** Our textual rationales are generated by an LLM conditioned on graph embeddings; although we supervise the answer token, the free-form explanations may not always be fully faithful to the decision process. Rigorous faithfulness metrics and expert evaluations are important next steps. (v) **Survey noise and missingness.** Self-reported outcomes can suffer from recall and social-desirability bias; skip logic and nonresponse introduce missingness patterns. Incorporating uncertainty modeling and missing-not-at-random strategies could increase reliability.

7 Ethical Considerations

Data provenance and privacy. All analyses use the public-use microdata releases of NSDUH and YRBS, which are de-identified by the data providers and contain no direct identifiers. We follow the usage terms specified by SAMHSA and CDC, make no attempt at re-identification.

Responsible interpretation of LLM outputs. The natural-language rationales produced by our LLM are research aids rather than authoritative guidance. They may contain errors or omissions and should not be used for individual-level clinical, legal, or policy decisions without expert review. Our system is intended to support hypothesis generation and population-level analysis under human oversight.

References

Sercan Ö Arik and Tomas Pfister. 2021. Tabnet: Attentive interpretable tabular learning. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 6679–6687.

Jeffrey Jensen Arnett. 2000. Emerging adulthood: A theory of development from the late teens through the twenties. *American psychologist*, 55(5):469.

Peyman Baghershahi, Gregoire Fournier, Pranav Nyati, and Sourav Medya. 2025. From nodes to narratives: Explaining graph neural networks with llms and graph context. *arXiv preprint arXiv:2508.07117*.

Vadim Borisov, Tobias Leemann, Kathrin Seßler, Johannes Haug, Martin Pawelczyk, and Gjergji Kasneci. 2022. Deep neural networks and tabular data: A survey. *IEEE transactions on neural networks and learning systems*, 35(6):7499–7519.

Dan Busbridge, Dane Sherburn, Pietro Cavallo, and Nils Y Hammerla. 2019. Relational graph attention networks. *arXiv preprint arXiv:1904.05811*.

CDC. 2025. Youth risk behavior surveillance system (yrbss). https://www.cdc.gov/yrbs/index.html. Accessed: 2025-10-06.

Mateusz Cedro and David Martens. 2024. Graphxain: Narratives to explain graph neural networks. *arXiv preprint arXiv:2411.02540*.

Tianqi Chen and Carlos Guestrin. 2016. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794.

Elizabeth A Crummy, Timothy J O'Neal, Britahny M Baskin, and Susan M Ferguson. 2020. One is not enough: understanding and modeling polysubstance use. *Frontiers in neuroscience*, 14:569.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pages 4171–4186.

Xinyu Dong, Jianyuan Deng, Wei Hou, Sina Rashidian, Richard N Rosenthal, Mary Saltz, Joel H Saltz, and Fusheng Wang. 2021. Predicting opioid overdose risk of patients with opioid prescriptions using electronic health records based on temporal deep learning. *Journal of biomedical informatics*, 116:103725.

Zi-Yi Dou, Anamika Barman-Adhikari, Fei Fang, and Amulya Yadav. 2021. Harnessing social media to identify homeless youth at-risk of substance use. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 14748–14756.

Kounianhua Du, Weinan Zhang, Ruiwen Zhou, Yangkun Wang, Xilong Zhao, Jiarui Jin, Quan Gan, Zheng Zhang, and David P Wipf. 2022. Learning enhanced representation for tabular data via neighborhood propagation. *Advances in neural information processing systems*, 35:16373–16384.

Luca Franceschi, Mathias Niepert, Massimiliano Pontil, and Xiao He. 2019. Learning discrete structures for graph neural networks. In *International conference on machine learning*, pages 1972–1982. PMLR.

Charles F Geier. 2013. Adolescent cognitive control and reward processing: implications for risk taking and substance use. *Hormones and behavior*, 64(2):333–342.

Adam Goodge, Bryan Hooi, See-Kiong Ng, and Wee Siong Ng. 2022. Lunar: Unifying local outlier detection methods via graph neural networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 6737–6745.

Will Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. *Advances in neural information processing systems*, 30.

Dae-Hee Han, Shieun Lee, and Dong-Chul Seo. 2020. Using machine learning to predict opioid misuse among us adolescents. *Preventive medicine*, 130:105886.

Yinhan He, Zaiyi Zheng, Patrick Soga, Yaozhen Zhu, Jundong Li, and 1 others. 2024. Explaining graph neural networks with large language models: A counterfactual perspective for molecular property prediction. *arXiv preprint arXiv:2410.15165*.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, and 1 others. 2022. Lora: Low-rank adaptation of large language models. *ICLR*, 1(2):3.

Uwaise Ibna Islam, Enamul Haque, Dheyaaldin Alsalman, Muhammad Nazrul Islam, Mohammad Ali Moni, and Iqbal H Sarker. 2023. A machine learning model for predicting individual substance abuse with associated risk-factors. *Annals of Data Science*, 10(6):1607–1634.

Wei Jin, Yao Ma, Xiaorui Liu, Xianfeng Tang, Suhang Wang, and Jiliang Tang. 2020. Graph structure learning for robust graph neural networks. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 66–74.

Aditya Kashyap, Chris Callison-Burch, and Mary Regina Boland. 2023. A deep learning method to detect opioid prescription and opioid use disorder from electronic health records. *International journal of medical informatics*, 171:104979.

Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. 2017. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.

Bharti Khemani, Shruti Patil, Ketan Kotecha, and Sudeep Tanwar. 2024. A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions. *Journal of Big Data*, 11(1):18.

Soeun Kim, Hyejun Kim, Seokjun Kim, Hojae Lee, Ahmed Hammoodi, Yujin Choi, Hyeon Jin Kim, Lee Smith, Min Seo Kim, Guillaume Fond, and 1 others. 2025. Machine learning–based prediction of substance use in adolescents in three independent worldwide cohorts: Algorithm development and validation study. *Journal of Medical Internet Research*, 27:e62805.

Adam King, Libo Li, and Yih-Ing Hser. 2017. Common statistical methods for primary and secondary analysis in substance abuse research. In *Research methods in the study of substance abuse*, pages 81–112. Springer.

Kevin M King and Laurie Chassin. 2007. A prospective study of the effects of age of initiation of alcohol and drug use on young adult substance dependence. *Journal of studies on alcohol and drugs*, 68(2):256–265.

Kevin M King and Laurie Chassin. 2008. Adolescent stressors, psychopathology, and young adult substance dependence: A prospective study. *Journal of Studies on Alcohol and Drugs*, 69(5):629–638.

Thomas N. Kipf and Max Welling. 2017. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations*.

Yixin Liu, Yu Zheng, Daokun Zhang, Hongxu Chen, Hao Peng, and Shirui Pan. 2022. Towards unsupervised deep graph structure learning. In *Proceedings of the ACM Web Conference 2022*, pages 1392–1403.

Monica Luciana. 2013. Adolescent brain development in normality and psychopathology. *Development and psychopathology*, 25(4pt2):1325–1345.

Jeremy W Luk, Jing Wang, and Bruce G Simons-Morton. 2010. Bullying victimization and substance use among us adolescents: Mediation by depression. *Prevention Science*, 11(4):355–359.

Tianyi Ma, Yiyue Qian, Zehong Wang, Zheyuan Zhang, Chuxu Zhang, and Yanfang Ye. 2025. Llm-empowered class imbalanced graph prompt learning for online drug trafficking detection. *arXiv preprint arXiv:2503.01900*.

Iris Mann, Tiffany Jenzer, Robert Miranda Jr, and Ethan H Mereish. 2021. Minority stress and nicotine use and dependence among sexual minority youth. *Substance use & misuse*, 56(14):2264–2268.

Megan E Marziali, Natalie S Levy, and Silvia S Martins. 2022. Perceptions of peer and parental attitudes toward substance use and actual adolescent substance use: The impact of adolescent-confidant relationships. *Substance Abuse*, 43(1):1085–1093.

Marina Matejevic, Dragana Jovanovic, and Vesna Lazarevic. 2014. Functionality of family relationships and parenting style in families of adolescents with substance abuse problems. *Procedia-Social and Behavioral Sciences*, 128:281–287.

Izraelle I McKinnon. 2024. Experiences of racism in school and associations with mental health, suicide risk, and substance use among high school students—youth risk behavior survey, united states, 2023. *MMWR supplements*, 73.

National Center for Health Statistics. 2024. Illicit drug use. https://www.cdc.gov/nchs/hus/sources-definitions/illicit-drug-use.htm. Accessed 2025-07-13.

Bo Pan, Zhen Xiong, Guanchen Wu, Zheng Zhang, Yifei Zhang, and Liang Zhao. 2024. Graphnarrator: Generating textual explanations for graph neural networks. *arXiv preprint arXiv:2410.15268*.

George C Patton, Susan M Sawyer, John S Santelli, David A Ross, Rima Afifi, Nicholas B Allen, Monika Arora, Peter Azzopardi, Wendy Baldwin, Christopher Bonell, and 1 others. 2016. Our future: a lancet commission on adolescent health and wellbeing. *The Lancet*, 387(10036):2423–2478.

Coral Rakovski, Pedro L Ballester, Bruno Braga Montezano, Luciano Dias de Mattos Souza, Karen Jansen, Ricardo Azevedo da Silva, Thaise Campos Mondin, Fernanda Pedrotti Moreira, Raquel Brandini De Boni, Benicio N Frey, and 1 others. 2023. Predictors of illicit substance abuse/dependence during young adulthood: A machine learning approach. *Journal of psychiatric research*, 157:168–173.

Daniel Romer, Valerie F Reyna, and Theodore D Satterthwaite. 2017. Beyond stereotypes of adolescent risk taking: Placing the adolescent brain in developmental context. *Developmental cognitive neuroscience*, 27:19–34.

Julie C Rusby, John M Light, Ryann Crowley, and Erika Westling. 2018. Influence of parent–youth relationship, parental monitoring, and parent substance use on adolescent substance use onset. *Journal of family psychology*, 32(3):310.

SAMHSA. 2025a. Key substance use and mental health indicators in the united states: Results from the 2024 national survey on drug use and health. https://www.samhsa.gov/data/sites/default/files/reports/rpt56287/2024-nsduh-annual-national-report.pdf. HHS Publication No. PEP25-07-007, NSDUH Series H-60.

SAMHSA. 2025b. National survey on drug use and health (nsduh). https://nsduhweb.rti.org/respweb/homepage.cfm. Accessed: 2025-10-06.

Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. 2018. Modeling relational data with graph convolutional networks. In *European semantic web conference*, pages 593–607. Springer.

Ravid Shwartz-Ziv and Amitai Armon. 2022. Tabular data: Deep learning is not all you need. *Information Fusion*, 81:84–90.

Kevin M Simon, Sharon J Levy, and Oscar G Bukstein. 2022. Adolescent substance use disorders. *NEJM evidence*, 1(6):EVIDra2200051.

Laurence Steinberg. 2008. A social neuroscience perspective on adolescent risk-taking. *Developmental Review*, 28(1):78–106. Accessed 2025-07-13.

Lev Telyatnikov and Simone Scardapane. 2023. Egg-gae: scalable graph neural networks for tabular data imputation. In *International conference on artificial intelligence and statistics*, pages 2661–2676. PMLR.

Christopher L Thompson, Karl C Alcover, and Sarah W Yip. 2021. Clinical prediction of extra-medical use of prescription pain relievers from a representative united states sample. *Preventive Medicine*, 149:106610.

Rebecca C Trenz, Michael Scherer, Paul Harrell, Julia Zur, Ashish Sinha, and William Latimer. 2012. Early onset of drug and polysubstance use as predictors of injection drug use among adult drug users. *Addictive behaviors*, 37(4):367–372.

Elisa M Trucco. 2020. A review of psychosocial factors linked to adolescent substance use. *Pharmacology biochemistry and behavior*, 196:172969.

Ali Unlu and Abdulhamit Subasi. 2025. Substance use prediction using artificial intelligence techniques. *Journal of Computational Social Science*, 8(1):21.

Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2017. Graph attention networks. *arXiv preprint arXiv:1710.10903*.

Adway Wadekar. 2019. Predicting opioid use disorder (oud) using a random forest. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*, volume 1, pages 960–961. IEEE.

Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S Yu. 2020. A comprehensive survey on graph neural networks. *IEEE transactions on neural networks and learning systems*, 32(1):4–24.

Jiahuan Yan, Jintai Chen, Yixuan Wu, Danny Z Chen, and Jian Wu. 2023. T2g-former: organizing tabular features into relation graphs promotes heterogeneous

feature interaction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 10720–10728.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, and 1 others. 2025. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*.

Jiaxing Zhang, Jiayi Liu, Dongsheng Luo, Jennifer Neville, and Hua Wei. 2024a. Llmexplainer: Large language model based bayesian inference for graph explanation generation. *arXiv preprint arXiv:2407.15351*.

Zheyuan Zhang, Zehong Wang, Shifu Hou, Evan Hall, Landon Bachman, Jasmine White, Vincent Galassi, Nitesh V Chawla, Chuxu Zhang, and Yanfang Ye. 2024b. Diet-odin: A novel framework for opioid misuse detection with interpretable dietary patterns. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 6312–6323.

Jianan Zhao, Xiao Wang, Chuan Shi, Binbin Hu, Guojie Song, and Yanfang Ye. 2021. Heterogeneous graph structure learning for graph neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pages 4697–4705.

Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, and 1 others. 2023. A survey of large language models. *arXiv preprint arXiv:2303.18223*, 1(2).

Yanqiao Zhu, Weizhi Xu, Jinghao Zhang, Yuanqi Du, Jieyu Zhang, Qiang Liu, Carl Yang, and Shu Wu. 2021. A survey on graph structure learning: Progress and opportunities. *arXiv preprint arXiv:2103.03036*.

A Data Preparation Details

In this section, we describe the construction of relational graphs and the labeling strategy for the YRBS and NSDUH survey datasets.

A.1 YRBS

The official YRBS user guide[1] provides detailed information on survey design and codebooks. To define the binary label of illicit drug use, we use variables including `QNILLICT` and `QN46-55`. Respondents with a value of 1 in any of these variables are labeled as positive cases. For graph construction, we include only the raw survey questions (prefixed with `Q`), excluding their dichotomous counterparts (prefixed with `QN`) to avoid redundancy. The user node feature vector is derived from `Q1-7`, where `Q1-Q5` are one-hot encoded, and `Q6` (height) and `Q7` (weight) are normalized. To prevent information leakage, we exclude questions `Q46-55` and `Q92-93`, which directly indicate illicit drug use. Topic assignments are informed by the YRBS Explorer[2] and the official handbook; we also employed an LLM-assisted approach to assign topics consistently.

A.2 NSDUH

The NSDUH user guide and codebook[3] provide detailed documentation on survey methodology and variable design. The binary label of illicit drug use is defined using the variable `ILLYR`, with respondents coded as 1 classified as positive. NSDUH variables fall into three categories: (1) *edited variables*, created after consistency checks; (2) *imputed variables*, with missing values statistically imputed; and (3) *recoded variables*, derived from edited or imputed sources. For graph construction, we use only edited variables to ensure consistency and minimize redundancy. User node features are built from `geographic` and `demographic` questions, one-hot encoded, with normalized height (`HTINCHE2`) and weight (`WTPOUND2`). To prevent information leakage, we exclude variables that directly indicate illicit drug use (e.g., `HPDRGTALK`, "Health professional discussed my drug use with me"). Topic definitions follow the codebook's table of contents.

[1] https://www.cdc.gov/yrbs/data/index.html
[2] https://yrbs-explorer.services.cdc.gov/
[3] https://www.samhsa.gov/data/data-we-collect/nsduh-national-survey-drug-use-and-health/datafiles

B Experimental Settings

In addition to the hyperparameters described in the main paper, we set the batch size to 16 and train for 20 epochs during the Graph Latent Relation Learning stage. For the LLM-Enhanced Detection and Reasoning stage, we use a batch size of 4 and train for 10 epochs. For GNN-based baselines, we employ 3 layers and adopt the same hyperparameters as those used in LAMI wherever applicable.

C Prompt Design

Prompt Templates

System Instruction:
Here are some question-answer pairs provided by a user aged between 15 and 25 years old. Based on these facts, infer whether this user uses illicit drugs.

User Input Format: Here are the question-answer pairs:

```
Question: <question_1>
Answer: <answer_1>
Question: <question_2>
Answer: <answer_2>
...
```

Cues about learned relations: Think about the possible relations between `<question_3>` and `<question_4>` given the user's answers to them.

...

Following Prompt A: Based on these facts, can you infer whether this user uses illicit drugs? Please answer with only "Yes" or "No".

Following Prompt B: Based on these facts, can you infer whether this user uses illicit drugs? Please answer "Yes" or "No" and explain why.

Figure 6: The shared prompt templates we used for experiments.

We use the prompt templates shown in Figure C. When finetuning and evaluating Qwen3-8B with LoRA and evaluating Qwen3-8B under zero-shot setting, we exclude *Cues about learned relations* and adopt *Following Prompt A*. When training LAMI, we adopt *Following Prompt A* to make it easier for computing $\mathcal{L}_{\text{gen}}$. While in inference time,

we adopt *Following Prompt B* to obtain justified textual output from the LLM module to facilitate further analysis.

D Case Study

To qualitatively assess how LAMI integrates behavioral and psychosocial information for prediction and explanation, we present two representative cases. Case A, sampled from the NSDUH dataset, represents a high-risk individual, while Case B, sampled from the YRBS dataset, reflects a low-risk profile. Each case includes the user's survey-derived question–answer pairs, the latent relation cues automatically provided by the model, and the final text-based inference generated by the LLM module.

Case A: High-Risk Profile. Case A describes a respondent exhibiting multiple behavioral and psychosocial risk factors. The user reports recent cigarette and e-cigarette use, low parental supervision, negative attitudes toward school, moderate academic performance, and several depressive symptoms such as fatigue, low energy, and feelings of worthlessness. The individual also lives in poverty and reports having carried a handgun once or twice. These variables collectively depict an environment characterized by psychosocial stress, weak family regulation, and exposure to risk-taking peers—conditions widely recognized in the literature as predictors of substance use.

LAMI's reasoning mirrors these patterns. The learned relation cues emphasize connections between parental monitoring and peer drinking behavior, between poverty and depressive affect, and between low school engagement and vaping. These relations correspond to established mechanisms such as social learning and stress–coping pathways (Trucco, 2020; King and Chassin, 2008). The model's textual output correctly infers illicit drug use, highlighting low parental control, peer influence, and emotional distress as interacting factors contributing to elevated risk. This demonstrates that LAMI can synthesize structural and contextual cues into an interpretable, human-like explanation grounded in behavioral science theory.

Case B: Low-Risk Profile. Case B represents a user with minimal behavioral risk indicators. The individual reports no cigarette, vaping, or alcohol use, no experiences of bullying or violence, and no depressive symptoms. Although the user indicates limited parental awareness, low physical activity, and modest academic performance, these factors alone are insufficient to suggest substance use. Overall, the profile reflects a relatively stable psychosocial environment with low direct exposure to drugs or associated behaviors.

The latent relation cues in this case focus on benign correlations—such as the link between limited parental supervision and physical inactivity, or between sleep duration and academic outcomes—none of which directly imply substance use risk. Consistent with these signals, LAMI predicts a negative outcome (`No`) and produces a concise rationale noting the absence of substance-related behaviors and major psychosocial distress. The model's explanation demonstrates appropriate calibration: it recognizes secondary vulnerabilities but does not overgeneralize them into a positive prediction.

Summary. Together, these cases illustrate that LAMI not only achieves accurate classification but also generates contextually grounded explanations aligned with established psychosocial mechanisms. In high-risk scenarios, the model integrates cross-domain cues—such as peer exposure, parental disengagement, and emotional dysregulation—to justify its prediction. In low-risk cases, it maintains conservative reasoning, acknowledging minor behavioral concerns without inferring unsupported drug use. These results underscore LAMI's interpretability and its potential to aid behavioral scientists in understanding the multifactorial nature of youth substance use.

Case A: input (part I)

System Instruction:
Here are some question-answer pairs provided by a user aged between 15 and 25 years old. Based on these facts, infer whether this user uses illicit drugs.

User Input:

```
Question: EVER SMOKED A CIGARETTE
Answer: 1 - Yes
Question: TIME SINCE LAST SMOKED CIGARETTES
Answer: 1 - Within the past 30 days
Question: AVG # CIGS SMOKED PER DAY/ON DAY SMOKED
Answer: 3 - 2 to 5 cigarettes per day
Question: EVER VAPED NICOTINE WITH E-CIGARETTE OR OTHER VAPING DEVICE
Answer: 1 - Yes
Question: TIME SINCE LAST VAPED NICOTINE WITH E-CIG OR VAPING DEVICE
Answer: 1 - Within the past 30 days
Question: HOW MANY SDNTS YOU KNOW IN GRADE SMOKE CIGARETTES
Answer: 3 - Most of them
Question: HOW MANY SDNTS YOU KNOW GET DRUNK WEEKLY
Answer: 3 - Most of them
Question: HOW YOU FELT OVERALL ABT GOING TO SCHL PAST 12 MOS
Answer: 4 - You hated going to school
Question: HOW OFTEN FELT SCHL WORK MEANINGFUL PST 12 MOS
Answer: 4 - Never
Question: GRADES FOR LAST SEMESTER/GRADING PERIOD COMPLETED
Answer: 3 - A C+ , C or C-minus average
Question: PRNTS CHECK IF YOU'VE DONE HOMEWORK PST 12 MOS
Answer: 4 - Never
Question: PRNTS LIMITED YOUR TIME OUT W/FRNDS PST 12 MOS
Answer: 4 - Never
Question: PRNTS LET YOU KNOW DONE A GOOD JOB PST 12 MOS
Answer: 4 - Never
Question: ARGUED/HAD A FIGHT WITH AT LEAST ONE OF YOUR PRNTS
Answer: 4 - 6 to 9 times
Question: SCORE OF SYMPTOM INDICATORS 1 THRU 9
Answer: 1 - Has 5 or more symptoms
Question: FELT WORTHLESS NEARLY EVERY DAY
Answer: 1 - Has symptom
Question: FELT TIRED/LOW ENERGY NEARLY EVERY DAY
Answer: 1 - Has symptom
Question: INABILITY TO CONCENTRATE OR MAKE DECISIONS
Answer: 1 - Has symptom
Question: RC-POVERTY LEVEL-NEW INC (% OF US CENSUS POVERTY THRESHOLD)
Answer: 1 - Living in Poverty
Question: CARRIED A HANDGUN
Answer: 2 - 1 or 2 times
```

Figure 7: Case A user input showing a high-risk behavioral and psychosocial profile.

Case A: Input (Part II)

Cues about learned relations: Think about the possible relations between PRNTS LIMITED YOUR TIME OUT W/FRNDS PST 12 MOS and HOW MANY SDNTS YOU KNOW GET DRUNK WEEKLY given the user's answers to them.
Think about the possible relations between RC-POVERTY LEVEL-NEW INC (% OF US CENSUS POVERTY THRESHOLD) and FELT WORTHLESS NEARLY EVERY DAY given the user's answers to them.
Think about the possible relations between EVER VAPED NICOTINE WITH E-CIGARETTE OR OTHER VAPING DEVICE and HOW YOU FELT OVERALL ABT GOING TO SCHL PAST 12 MOS given the user's answers to them.
Think about the possible relations between CARRIED A HANDGUN and PRNTS LET YOU KNOW DONE A GOOD JOB PST 12 MOS given the user's answers to them.
Think about the possible relations between HOW MANY SDNTS YOU KNOW IN GRADE SMOKE CIGARETTES and INABILITY TO CONCENTRATE OR MAKE DECISIONS given the user's answers to them.
Following Prompt: Based on these facts, can you infer whether this user uses illicit drugs? Please answer "Yes" or "No" and explain why.

Figure 8: Learned latent relations used as reasoning cues in Case A's LLM prompt.

Case A: Model Output

Model output: Yes. The user reports recent cigarette and vaping use, low parental monitoring, poor school engagement, and several depressive symptoms, which are common indicators linked to higher drug-use risk. The user also has peers who drink frequently, which may reinforce such behaviors through social influence. Considering these patterns together, especially the link between low parental control and strong peer exposure, I think the answer is yes.

Figure 9: Model inference and explanation for the high-risk Case A user.

Case B: Input (Part I)

System Instruction:
Here are some question-answer pairs provided by a user aged between 15 and 25 years old. Based on these facts, infer whether this user uses illicit drugs.

User Input: `Question: Have you ever smoked a cigarette, even one or two puffs? Answer: No Question: During the past 30 days, on how many days did you smoke cigarettes? Answer: 0 days Question: Have you ever used an electronic vapor product? Answer: No Question: During the past 30 days, on how many days did you use an electronic vapor product? Answer: 0 days Question: How old were you when you had your first drink of alcohol other than a few sips? Answer: Never drank alcohol Question: During the past 30 days, on how many days did you have at least one drink of alcohol? Answer: 0 days Question: Did you drink alcohol or use drugs before you had sexual intercourse the last time? Answer: No Question: During the past 12 months, did you ever feel so sad or hopeless almost every day for two weeks or more in a row that you stopped doing some usual activities? Answer: No Question: During the past 12 months, did you ever seriously consider attempting suicide? Answer: No Question: During the past 12 months, have you ever been bullied on school property? Answer: No Question: During the past 12 months, have you ever been electronically bullied? (Count being bullied through texting, Instagram, Facebook, or other social media.) Answer: No Question: During the past 12 months, how many times were you in a physical fight? Answer: 0 times Question: During the past 12 months, how many times did someone threaten or injure you with a weapon such as a gun, knife, or club on school property? Answer: 0 times Question: How often do your parents or other adults in your family know where you are going or with whom you will be? Answer: Never Question: Do you agree or disagree that you feel close to people at your school? Answer: Disagree Question: On an average school night, how many hours of sleep do you get? Answer: 5 hours Question: During the past 7 days, on how many days were you physically active for a total of at least 60 minutes per day? (Add up all the time you spent in any kind of physical activity that increased your heart rate and made you breathe hard some of the time.) Answer: 0 days Question: During the past 12 months, how would you describe your grades in school? Answer: Mostly C's Question: During your life, how often has there been an adult in your household who tried hard to make sure your basic needs were met, such as looking after your safety and making sure you had clean clothes and enough to eat? Answer: Never Question: When was the last time you saw a dentist for a check-up, exam, teeth cleaning, or other dental work? Answer: Never`

Figure 10: Case B user input showing a low-risk behavioral and health profile.

Case B: Input (Part II)

Cues about learned relations: Think about the possible relations between `How often do your parents or other adults in your family know where you are going or with whom you will be?` and `During the past 7 days, on how many days were you physically active for a total of at least 60 minutes per day?` given the user's answers to them.
Think about the possible relations between `On an average school night, how many hours of sleep do you get?` and `During the past 12 months, how would you describe your grades in school?` given the user's answers to them.
Think about the possible relations between `During the past 30 days, on how many days did you have at least one drink of alcohol?` and `Did you drink alcohol or use drugs before you had sexual intercourse the last time?` given the user's answers to them.
Think about the possible relations between `During your life, how often has there been an adult in your household who tried hard to make sure your basic needs were met, such as looking after your safety and making sure you had clean clothes and enough to eat?` and `When was the last time you saw a dentist for a check-up, exam, teeth cleaning, or other dental work?` given the user's answers to them.
Think about the possible relations between `During the past 12 months, have you ever been electronically bullied? (Count being bullied through texting, Instagram, Facebook, or other social media.)` and `During the past 12 months, did you ever feel so sad or hopeless almost every day for two weeks or more in a row that you stopped doing some usual activities?` given the user's answers to them.
Following Prompt: Based on these facts, can you infer whether this user uses illicit drugs? Please answer "Yes" or "No" and explain why.

Figure 11: Learned latent relations used as reasoning cues in Case B's LLM prompt.

Case B: Model Output

Model output: No. They report no cigarette, vaping, or alcohol use, and also no use before sex. No fights or bullying and no recent sadness. There are some risks like less sleep, low activity, weaker grades, and low parental monitoring, but overall there isn't evidence of illicit drug use here.

Figure 12: Model inference and explanation for the low-risk Case B user.