



# Diet-ODIN: A Novel Framework for Opioid Misuse Detection with Interpretable Dietary Patterns

Zheyuan Zhang\*

University of Notre Dame  
Notre Dame, Indiana, USA  
zzhang42@nd.edu

Zehong Wang\*

University of Notre Dame  
Notre Dame, Indiana, USA  
zwang43@nd.edu

Shifu Hou

University of Notre Dame  
Notre Dame, Indiana, USA  
shou@nd.edu

Evan Hall

University of Notre Dame  
Notre Dame, Indiana, USA  
ehall9@nd.edu

Landon Bachman

University of Notre Dame  
Notre Dame, Indiana, USA  
lbachma2@nd.edu

Jasmine White

Purdue University  
West Lafayette, IN, USA  
white742@purdue.edu

Vincent Galassi

University of Notre Dame  
Notre Dame, Indiana, USA  
vgalassi@nd.edu

Nitesh V. Chawla

University of Notre Dame  
Notre Dame, Indiana, USA  
nchawla@nd.edu

Chuxu Zhang

Brandeis University  
Waltham, MA, USA  
chuxuzhang@brandeis.edu

Yanfang Ye<sup>†</sup>

University of Notre Dame  
Notre Dame, Indiana, USA  
yye7@nd.edu

## ABSTRACT

The opioid crisis has been one of the most critical society concerns in the United States. Although the medication assisted treatment (MAT) is recognized as the most effective treatment for opioid misuse and addiction, the various side effects can trigger opioid relapse. In addition to MAT, the dietary nutrition intervention has been demonstrated its importance in opioid misuse prevention and recovery. However, research on the alarming connections between dietary patterns and opioid misuse remain under-explored. In response to this gap, in this paper, we first establish a large-scale multifaceted dietary benchmark dataset related to opioid users *at the first attempt* and then develop a novel framework - i.e., namely **Opioid Misuse Detection with INterpretable DIetary Patterns (Diet-ODIN)** - to bridge heterogeneous graph (HG) and large language model (LLM) for the identification of users with opioid misuse and the interpretation of their associated dietary patterns. Specifically, in Diet-ODIN, we first construct an HG to comprehensively incorporate both dietary and health-related information, and then we devise a holistic graph learning framework with noise reduction to fully capitalize both users' individual dietary habits and shared dietary patterns for the detection of users with opioid misuse. To further delve into the intricate correlations between dietary patterns and opioid misuse, we exploit an LLM by utilizing the knowledge obtained from the graph learning model for interpretation. The extensive experimental results based on our established benchmark

with quantitative and qualitative measures demonstrate the outstanding performance of Diet-ODIN on exploring the complex interplay between opioid misuse and dietary patterns, by comparison with state-of-the-art baseline methods. Our code, built benchmark and system demo are available at <https://github.com/Diet-ODIN>.

## CCS CONCEPTS

• **Computing methodologies** → **Knowledge representation and reasoning**; **Neural networks**; • **Information systems** → **Social networks**.

## KEYWORDS

Opioid Misuse Detection, Graph Neural Network, Large Language Model, Dietary Pattern

## ACM Reference Format:

Zheyuan Zhang, Zehong Wang, Shifu Hou, Evan Hall, Landon Bachman, Jasmine White, Vincent Galassi, Nitesh V. Chawla, Chuxu Zhang, and Yanfang Ye. 2024. 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 (KDD '24)*, August 25–29, 2024, Barcelona, Spain. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3637528.3671587>

## 1 INTRODUCTION

Opioids are a class of drugs including the illegal drug heroin, synthetic opioids like fentanyl, and prescription pain relievers such as oxycodone [35]. Besides pain relief, opioids could produce euphoria and therefore might be easily misused [8, 39, 40] - e.g., 10.1 million Americans reported misusing opioids in 2019; an estimated 108,000 drug overdose deaths in the United States in 2021 [4], 90% of which involved opioids [47]. Besides the negative impact on public health, opioid misuse and addiction have also caused devastating socioeconomic crises (e.g., domestic violence, child abuse). To battle the deadly opioid epidemic, medication assisted treatment (MAT) [14], which takes the opioid agonists such as methadone or buprenorphine in combination with counseling, has been recognized as the

\*Both authors contributed equally.

<sup>†</sup>Corresponding Author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

KDD '24, August 25–29, 2024, Barcelona, Spain.

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 979-8-4007-0490-1/24/08.

<https://doi.org/10.1145/3637528.3671587>

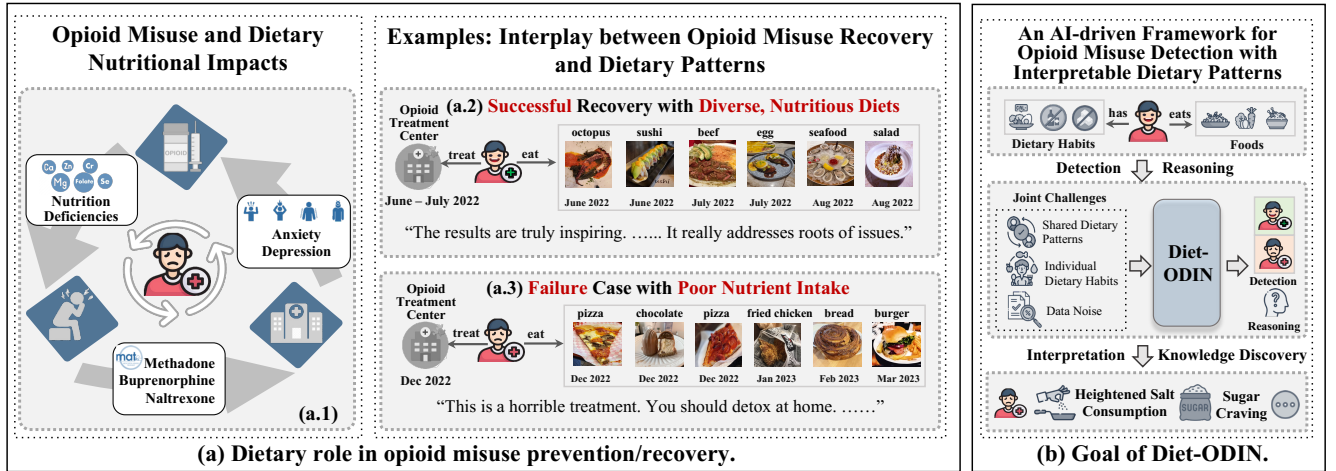


Figure 1: The illustration of (a) dietary role in opioid misuse prevention/recovery, and (b) goal of Diet-ODIN.

most effective treatment for opioid misuse and addiction [30]. Despite its effectiveness, the various induced side-effects can often trigger opioid relapse [8, 30, 34], which calls for novel solutions to promote health resilience for people against the opioid epidemic.

In addition to MAT, it has been demonstrated that dietary nutrition intervention could play an important role in opioid misuse prevention and recovery [6, 8, 23, 25, 34, 52]. As illustrated in Figure 1.(a.1), the research has shown that opioid misuse can often lead to malnutrition, nutrient deficiencies such as reduced protein [28] and dietary fiber intake [24], metabolic disorders that compromise nutrition, and altered body composition, which in turn impede its prevention and recovery; while Figure 1.(a.2) and (a.3) have also exemplified the interplay between opioid misuse recovery and associated dietary patterns. These indicate that the distinct dietary patterns may serve as subtle yet informative indicators for the early detection of opioid misuse, particularly in scenarios such as self-medication or poly-pharmacy, where individuals might not recognize their misuse [29], or in cases where misuse is intentionally concealed. Moreover, the task can not only assist in the timely identification of at-risk individuals for healthcare providers and relevant authorities to intervene, but also pave a new paradigm in opioid misuse and addiction treatment through the understanding of dietary patterns linked to opioid misuse. This naturally brings up the question: *how can we leverage dietary data to identify users with potential opioid misuse, and enhance the understanding of both known and less-known dietary patterns associated with opioid misuse?*

To answer this question, existing works [23, 25, 30, 34, 52] predominantly rely on traditional methods, such as limited surveys and case studies. There exists a notable gap in automating larger scale data analysis to comprehensively explore the alarming connections between dietary patterns and opioid misuse. To bridge this gap, as shown in Figure 1.(b), in this work, we establish a large-scale multifaceted dietary benchmark dataset related to opioid users *at the first attempt* and develop a novel framework - i.e., namely **Opioid Misuse Detection with Interpretable Dietary Patterns (Diet-ODIN)** - to bridge heterogeneous graph (HG) and large language model (LLM)

for the identification of users with opioid misuse and the interpretation of their associated dietary patterns. Specifically, in Diet-ODIN, based on the National Health and Nutrition Examination Survey (NHANES) data [5], which intends to evaluate the health and nutritional status of adults and children in the United States, we first construct the *NHANES Dietary Graph* with heterogeneous graph structure to comprehensively incorporate users’ dietary and health-related information. Then, Diet-ODIN is designed to detect users with potential opioid misuse through their dietary patterns, including food choices and eating habits, and to further provide pertinent interpretations. Nevertheless, distinguishing users with potential opioid misuse from others through their dietary patterns is a non-trivial task, as it involves comprehensively leveraging multifaceted yet complex information, including (1) the common dietary patterns among users, and (2) the specific foods and/or habits adopted by individuals; (3) in addition, noise in real-life dataset, such as users’ faulty dietary recollections or data recording errors, renders the task more challenging. To jointly tackle the three grand challenges above for the detection of users with potential opioid misuse, based on the constructed *NHANES Dietary Graph*, Diet-ODIN presents the **Noise Reducing Heterogeneous Graph Neural Network (NR-HGNN)** by introducing a holistic graph learning framework to fully integrate both users’ individual dietary habits and shared dietary patterns with a refinement component for denoising. Moreover, to interpret the correlation between detected users and associated dietary patterns, Diet-ODIN further bridges LLM and NR-HGNN to provide explicit explanations by utilizing the knowledge learned from the NR-HGNN as prompts for LLM. Lastly, the new findings based on the extensive analysis performed over the outputs from Diet-ODIN unveil new insights for combating the opioid crisis. Overall, the key contributions of this work are summarized as:

- To the best of our knowledge, this is the *first* AI-driven framework that comprehensively explores the correlations between opioid misuse and users’ diets with explainable outputs. In this work, we have also established a large-scale multifaceted dietary

benchmark dataset related to opioid users at the first attempt for other researchers and practitioners towards this line of research.

- We develop an *innovative and holistic* framework named **Diet-ODIN** to detect users with potential opioid misuse with interpretable dietary patterns. Based on the built *NHANES Dietary Graph*, this framework presents a novel **NR-HGNN** model to jointly capitalize shared dietary patterns among users with a micro-level aggregation module and individual dietary behaviors with a macro-level aggregation module, while also emphasizing noise reduction with a refinement component. Moreover, to interpret the relationship between dietary patterns and opioid misuse, Diet-ODIN implements two innovative strategies leveraging the knowledge learned from **NR-HGNN** for an LLM to enhance its effectiveness in analytical reasoning.
- We conduct extensive experiments based on our established benchmark, whose results show that Diet-ODIN outperforms existing baselines in detecting potential opioid misuse and offers reliable interpretations of the correlations between opioid misuse and dietary patterns, aligning with both existing literature and statistical analysis. These interpretations provide *new insights* for combating the opioid crisis, shedding light on prospective avenues for future research.

## 2 RELATED WORK

**Heterogeneous Graph Neural Networks.** Graph Neural Networks (GNNs) have demonstrated effective in a variety of graph learning tasks, outperforming traditional methods [17, 37, 45] by leveraging advanced message-passing frameworks [19, 26, 51]. However, many approaches mainly focus on homogeneous graphs and struggle to interpret the complex semantics inherent in heterogeneous graphs, which consist of diverse node and relation types. To address this, several studies [10, 16, 54–56, 63] have utilized meta-paths to capture high-order semantics from such graphs. Despite their advances, these methods often heavily rely on expert knowledge for meta-path definition and might overlook contextual information. In response, alternative efforts [22, 27, 41] have introduced type-specific transformations to better respect the graphs' heterogeneity on simpler, direct connections, but at the risk of underrepresenting longer-range dependencies. Moreover, both approaches may struggle with noise frequently existing in real-world heterogeneous graphs. In contrast, based on our *NHANES Dietary Graph*, NR-HGNN addresses these issues by balancing the integration of both micro- and macro-level information aggregation for richer node embeddings and introducing a refinement module for noise reduction in practical settings.

**Large Language Models for Reasoning.** Large language models [3, 48, 49, 59, 62] have demonstrated remarkable reasoning capabilities across a range of commonsense tasks [3, 50]. Despite their success, LLMs often encounter challenges in fields that demand specialized knowledge, including healthcare [60], law [2], finance [44], and medicine [46]. Addressing this limitation, researchers have enriched LLMs with domain-specific knowledge through manual interventions, such as tailored prompts [44, 46] and in-context learning [2, 60], by elucidating task-specific factors. Drawing on these advancements, we propose to enhance LLMs' reasoning abilities to discern correlations between dietary patterns and opioid usage.

Specifically, by utilizing the knowledge learned from the NR-HGNN as prompts, we elegantly bridge LLM and NR-HGNN to interpret the complex interplay between opioid misuse and dietary patterns. **Opioid Misuse Detection.** Research in opioid misuse detection encompasses both quantitative and qualitative approaches. Quantitative studies predominantly employ machine learning algorithms for predicting opioid overdose [57, 64] or relapse risk from medical records [15, 20], or for identifying opioid users and traffickers through social media data analysis [13, 21, 38]. Nonetheless, these approaches primarily focus on prediction rather than offering detailed analyses or innovative insights into opioid misuse prevention. On the qualitative side, research within healthcare domain examines the link between dietary patterns and chronic opioid usage through case studies and qualitative analyses [7, 58], highlighting significant correlations [1, 28]. Yet, these studies are often restricted to specific instances and lack a predictive model applicable universally. Addressing this gap, we introduce a comprehensive framework that not only detects but also provides interpretable analyses of opioid misuse based on dietary patterns.

## 3 NHANES DIETARY GRAPH

In this section, we first elucidate the key concepts foundational to our study, followed by an exposition on *NHANES Dietary Graph*.

*Definition 3.1 (Heterogeneous Graphs).* A Heterogeneous Graph (HG) [53] is denoted as  $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{T}, \mathcal{R})$ , comprising sets of nodes  $\mathcal{V}$  and edges  $\mathcal{E}$ .  $\mathcal{T}$  and  $\mathcal{R}$  represent the types of nodes and edges, respectively, with the condition that  $|\mathcal{T}| + |\mathcal{R}| > 2$ . Additionally,  $\tau(\cdot)$  and  $\varphi(\cdot; \cdot)$  serve as mapping functions to identify the types of nodes and edges, where  $\tau(i) \in \mathcal{T}$  for any  $i \in \mathcal{V}$ , and  $\varphi(i, j) \in \mathcal{R}$  for any edge  $(i, j) \in \mathcal{E}$ .

*Definition 3.2 (Meta-Path).* A Meta-path  $m \in \mathcal{M}$  [54] represents a sequence of node and edge types, functioning as a composite of relations:  $\mathcal{P}_m : \mathcal{T}_1 \xrightarrow{\mathcal{R}_1} \mathcal{T}_2 \xrightarrow{\mathcal{R}_2} \dots \xrightarrow{\mathcal{R}_{n-1}} \mathcal{T}_n$ . It encapsulates a particular form of high-order semantics within an HG, with different meta-paths potentially offering complementary insights.

*Definition 3.3 (Meta-Path-based Neighborhoods).* Meta-path-based neighborhoods refer to sets of neighborhoods interconnected through specific meta-paths. We define  $\mathcal{P}_m(i, j)$  as the function measuring the path count connecting nodes  $i$  and  $j$  via meta-path  $m$ , thereby delineating these neighborhoods as  $\mathcal{N}_{\mathcal{P}_m}^k(i) = \{j | \mathcal{P}_m(i, j) \geq k\}$ , typically with  $k = 1$ .

**Constructing NHANES Dietary Graph.** Utilizing NHANES data from 2003 to 2020, we developed NHANES Dietary Graph, the first graph dataset aiming at investigating the relationship between dietary patterns and opioid misuse. This graph is constructed from data on food intake, dietary habits, and prescription drug usage. As depicted in Figure 2, it features five types of nodes: *user*, *food*, *habit*, *ingredient*, and *category*, and four types of relationships: *user-eat-food*, *user-have-habit*, *food-contain-ingredient*, and *food-belong-to-category*. Within this graph, we can define different meta-paths, such as *User-Habit-User* (UHU) and *User-Food-User* (UFU), which represent shared dietary habits and food consumption patterns among users, respectively. Opioid users are identified by either (1) records of illicit opioid drug use, like heroin, within a year,

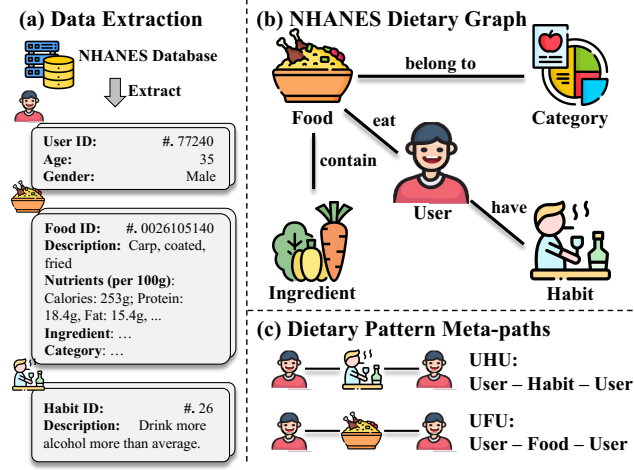


Figure 2: The schema of NHANES Dietary Graph.

or (2) records of prescription opioid medication use for over 90 days, which is a criteria commonly employed in the domain [18]. Thus, we form the opioid user detection as a binary classification task. User features include anonymized demographic information, while food features comprise nutritional values, such as calories, protein, and sodium. For text-described node types such as food, habit, ingredient, and category, we employ pre-trained BERT [9] to encode their descriptions into node features. Details on feature processing and dataset statistics are available in Appendix A.

## 4 THE FRAMEWORK OF DIET-ODIN

Based on the built *NHANES Dietary Graph*, we introduce an innovative framework for Opioid Misuse Detection with INterpretable Dietary Patterns (**Diet-ODIN**), illustrated in Figure 3. The framework incorporates a Noise Reducing Heterogeneous Graph Neural Network (**NR-HGNN**) aimed at identifying users with opioid misuse (detailed in Section 4.1), and a LLM-based interpretation module bridged with NR-HGNN to explore the relationship between dietary patterns and opioid misuse (outlined in Section 4.2).

### 4.1 NR-HGNN for Detecting Users with Opioid Misuse

Detecting users with opioid misuse presents a complex challenge, requiring comprehensive analysis of (1) shared dietary patterns among users, (2) individual food consumption or habits, and (3) reducing noise introduced in data collection. Our NR-HGNN model addresses these issues by jointly aggregating the information at both macro-level (to tackle the first challenge) and micro-level (to address the second challenge) and performing noise reduction through graph refinement strategies (to resolve the third challenge). As illustrated in Figure 3 (a), the NR-HGNN model comprises three principal components: (1) shared dietary pattern learning with marco-level aggregation (i.e., Figure 3 (a.1)), (2) individual dietary habit learning with micro-level aggregation (i.e., Figure 3 (a.2)), and (3) graph refinement with noise reduction for the detection (i.e., Figure 3 (a.3)), with detailed discussions to follow.

**4.1.1 Shared Dietary Pattern Learning with Marco-level Aggregation.** Users with opioid misuse often exhibit unique dietary preferences, notably towards sugar and high-calorie foods, distinguishing them from non-opioid users [31, 32]. Motivated by this observation, we introduce macro-level aggregation to utilize shared dietary patterns between users for the detection of opioid misuse. Specifically, we employ meta-paths *User-Habit-User* (UHU) and *User-Food-User* (UFU) to capture the shared food consumption and habit patterns among users. Following the methodology of [54], these meta-paths facilitate the extraction of subgraphs by linking users with identical dietary patterns, thereby preserving intricate high-order semantics. Nevertheless, direct extraction may lead to the inclusion of noisy subgraphs, as dietary patterns are not solely determined by individual foods or habits [1, 28]. For example, deducing similar dietary habits from a single shared meal, such as fried chicken, is over-simplistic. Addressing this, we introduce a novel subgraph extraction method, termed *connectivity filtering*, which establishes connections between two users only if they share multiple dietary habits or foods:

$$\mathcal{G}^m = (\mathcal{V}^m, \mathcal{E}^m), \mathcal{E}^m = \{(i, j) | j \in \mathcal{N}_{\mathcal{P}_m}^k(i)\}, \quad (1)$$

where  $\mathcal{G}^m$  represents the subgraph generated through meta-path  $\mathcal{P}_m$ , and  $k$  signifies an empirically determined threshold for selection. This methodology ensures the retention of only the most pertinent connections within the subgraphs.

Following the establishment of meta-path-induced subgraphs, we delve into the extraction of semantics from each meta-path, effectively modeled through message-passing mechanisms:

$$\mathbf{h}_i^m = \sum_{j \in \mathcal{N}_{\mathcal{P}_m}^k(i)} \alpha_{ij}^m \cdot \mathbf{W} \mathbf{h}_j^m, \quad (2)$$

where  $\mathbf{h}^m$  represents node embeddings within subgraph  $\mathcal{G}^m$ ,  $\mathbf{W}$  denotes a learnable matrix, and  $\alpha_{ij}^m$  signifies the correlation between nodes  $i$  and  $j$ , which is learnable or fixed. To capture fine-grained knowledge, we further attention mechanisms to dynamically allocate these correlations:

$$\alpha_{ij}^m = \frac{\exp(\sigma(\mathbf{a}^\top \cdot [\mathbf{W} \mathbf{h}_i^m \parallel \mathbf{W} \mathbf{h}_j^m]))}{\sum_{t \in \mathcal{N}_{\mathcal{P}_m}^k(i)} \exp(\sigma(\mathbf{a}^\top \cdot [\mathbf{W} \mathbf{h}_i^m \parallel \mathbf{W} \mathbf{h}_t^m]))}, \quad (3)$$

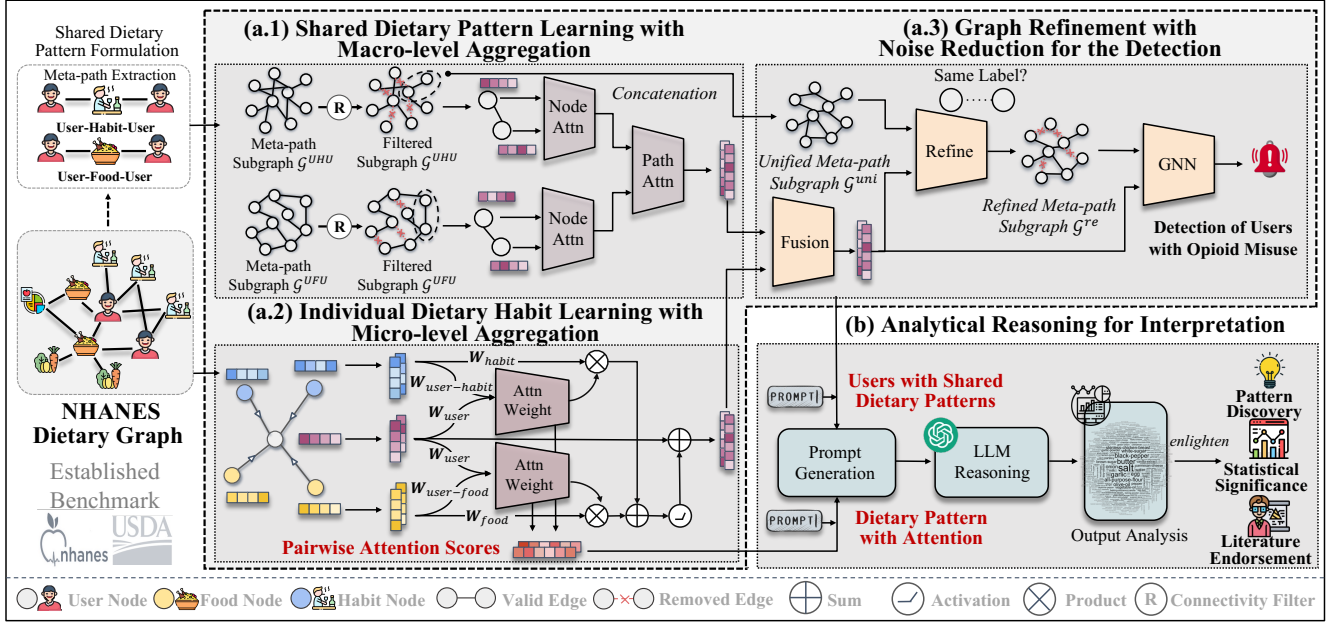
with  $\sigma(\cdot)$  as the activation function and  $\mathbf{a}$  as the learnable attention vector. Recognizing the complex etiology behind opioid misuse, the attention also evaluates the significance of various factors (e.g., food and habits) in prediction. Integration of insights from different meta-paths (UFU and UHU) involves analyzing their importance:

$$\mathbf{h}_i^{Ma} = \sigma(\sum_{m \in \mathcal{M}} \beta^m \cdot \mathbf{h}_i^m), \quad (4)$$

$$\beta^m = \frac{\exp(\sum_{i \in \mathcal{V}^m} \mathbf{q}^\top \cdot \tanh(\mathbf{W} \mathbf{h}_i))}{\sum_{m \in \mathcal{M}} \exp(\sum_{i \in \mathcal{V}^m} \mathbf{q}^\top \cdot \tanh(\mathbf{W} \mathbf{h}_i))}, \quad (5)$$

where  $\mathbf{h}^{Ma}$  encapsulates the macro-level node embeddings, and  $\mathbf{q}$  is an additional attention vector that discerns the relevance of various meta-paths. This hierarchical attention mechanism empowers the model to discern nuanced high-order semantics within the graph.

**4.1.2 Individual Dietary Habit Learning with Micro-level Aggregation.** Macro-level aggregation captures high-order information within HGs, yet may overlook crucial localized knowledge essential for accurate opioid misuse identification [7, 58]. This localized information, embodying the direct relationships between users and their



**Figure 3: The overview framework of Diet-ODIN, which consists of (a) a graph learning framework called NR-HGNN for detecting opioid users, and (b) an LLM-powered reasoning module for interpreting the most important dietary patterns.**

dietary behaviors—specifically, their food intake and habits—varies in its prediction for opioid misuse. To incorporate this low-order information, we introduce a micro-level aggregation module. This module is designed to discern the relevance of particular dietary patterns dynamically, thereby facilitating prediction accuracy:

$$\mathbf{h}_i^{Mi} = \sigma\left(\frac{1}{|\mathcal{N}(i)|} \sum_{j \in \mathcal{N}(i)} \alpha_{ij}^{Mi} \cdot \mathbf{W}_{\tau(j)} \mathbf{h}_j^{Mi}\right), \quad (6)$$

where  $\mathbf{h}^{Mi}$  represents the micro-level node embedding,  $\mathcal{N}(i)$  indicates node  $i$ 's direct neighbors, and  $\mathbf{W}_{\tau(j)}$  applies node-type specific transformations reflecting HG's inherent heterogeneity. The micro-level attention  $\alpha_{ij}^{Mi}$  quantifies each neighbor's significance, capturing the dietary pattern-user relationship:

$$\alpha_{ij}^{Mi} = \text{softmax}\left(\frac{[\mathbf{W}_{\tau(i)} \mathbf{h}_i^{Mi}]^\top \cdot \mathbf{W}_{\varphi(i,j)} \cdot [\mathbf{W}_{\tau(j)} \mathbf{h}_j^{Mi}]}{\sqrt{d}}\right), \quad (7)$$

$\mathbf{W}_{\tau(\cdot)}$  and  $\mathbf{W}_{\varphi(\cdot, \cdot)}$  enable precise modeling of node and relation types, enhancing the framework's ability to navigate HG's complexity. The introduction of a normalization factor  $\sqrt{d}$  in attention scoring mitigates potential gradient vanishing and stabilizes training [50]. Beyond enhancing prediction capabilities, this attention-based method prioritizes specific dietary behaviors in the reasoning process, as further discussed in Section 4.2.

**4.1.3 Graph Refinement with Noise Reduction for the Detection.** The macro-level aggregation captures high-order information while micro-level aggregation processes low-order information, both of which are essential and complementary aspects of our NR-HGNN. To integrate these embeddings and fully leverage the information within the HG, we apply the following fusion operation:

$$\mathbf{h}_i = \text{ReLU}(\mathbf{W} \cdot [\mathbf{h}_i^{Ma} \parallel \mathbf{h}_i^{Mi}] + \mathbf{b}). \quad (8)$$

However, empirical evidence suggests that directly applying the fused embedding  $\mathbf{h}$  for prediction leads to sub-optimal. We attribute this to potential noise from data collection processes, where inaccuracies may arise from users' faulty recollections of dietary intake or frequency of certain habits, as well as possible errors introduced by interviewers. Such inaccuracies can manifest as either spurious or missing connections in the graph, consequently generating noise in the graph structure. This noise can render the embeddings derived from the two aggregation modules unreliable by falsely correlating graph structure with predictions. To address this, we propose initially refining the noisy dietary patterns within the graph structure. Subsequently, we input both the node embeddings and the refined graph into an additional GNN to derive more accurate embeddings.

We consider leveraging user correlations for refinement, positing that users with similar dietary patterns are likely to exhibit similarities. To consolidate various yet complementary dietary patterns, we integrate different meta-paths into a unified graph:

$$\mathcal{G}^{uni} = \bigcup_{m \in \mathcal{M}} \mathcal{G}^m = (\mathcal{V}^{uni}, \mathcal{E}^{uni}), \quad (9)$$

where  $\mathcal{M}$  represents the set of meta-paths, including *User-Food-User* and *User-Habit-User*. To determine the refinement signal, we consider if two connected users share the same label, the connection would be reliable, based on the hypothesis that dietary patterns of opioid and non-opioid users differ significantly. This leads us to model refinement through a link prediction approach:

$$\mathcal{L}_{refine} = -\frac{1}{|\mathcal{E}_{train}^{uni}|} \sum_{(i,j) \in \mathcal{E}_{train}^{uni}} y_{ij} \log(y_{ij}) + (1 - y_{ij}) \log(1 - y_{ij}), \quad (10)$$

where  $y_{ij} = \text{pred}_{refine}([\mathbf{h}_i \parallel \mathbf{h}_j])$  predicts the edge reliability. A subset of edges,  $\mathcal{E}_{train}^{uni}$ , is selected for training the refinement model,

ensuring it only accesses nodes from the training dataset to prevent information leakage from unseen nodes. Subsequently, the refinement model evaluates all edges in  $\mathcal{G}^{uni}$  for reliability, discarding those deemed unreliable:

$$\mathcal{G}^{re} = (\mathcal{V}^{uni}, \mathcal{E}^{re}), \mathcal{E}^{re} = \{(i, j) \in \mathcal{E}^{uni} | \hat{y}_{ij} = 1\}. \quad (11)$$

This process enhances graph homophily, a crucial aspect in analyzing social phenomena between people [12]. The refined graph is then employed to derive reliable node embeddings for predictions:

$$\hat{\mathbf{h}}_i = \phi(\mathbf{h}_i, \mathcal{G}^{re}), \quad (12)$$

$$\hat{y}_i = \text{pred}_{node}(\hat{\mathbf{h}}_i) = \text{softmax}(\mathbf{W} \cdot \hat{\mathbf{h}}_i + \mathbf{b}), \quad (13)$$

where  $\phi(\cdot, \cdot)$  denotes the GNN for refining node embeddings  $\hat{\mathbf{h}}_i$ , and  $\hat{y}_i$  represents the model predictions. Optimization is achieved through minimizing the cross-entropy loss:

$$\mathcal{L} = -\frac{1}{|\mathcal{V}|} \sum_{i \in \mathcal{V}} y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i). \quad (14)$$

As refined graphs maintain more reliable patterns, they also enhance the reasoning process, as elucidated in Section 4.2.

## 4.2 Bridging NR-HGNN and LLM for Interpretation

Beyond detecting users with opioid misuse, our Diet-ODIN framework also excels at interpreting key dietary patterns associated with opioid misuse, thereby aiding in the prevention and treatment of its misuse. To achieve this, Diet-ODIN bridges NR-HGNN and LLM to investigate the nexus between opioid misuse and dietary behaviors, providing individuals with meaningful insights. To this end, our objective is to design prompts that enable LLM to perform trustworthy reasoning. As demonstrated in Figure 4, a prompt consists of three parts: (1) *Instruction*, (2) *Context*, and (3) *Notes*. The *Instruction* conveys task-specific directives to the LLM, clarifying the task at hand. The *Context* provides user-specific information, facilitating tailored reasoning. The *Notes* are employed to guide LLM behavior, such as restricting output length. Notably, *Context* is the most crucial component, presenting user information. A straightforward context design is to directly convert the characteristics of users and their dietary patterns in text, satisfying the input format of LLM; but the method cannot fully exploit the potential of LLM, particularly in domains like healthcare, where domain-specific knowledge is vital [46]. To address this, recent studies [2, 60] have incorporated domain-specific knowledge into prompt generation, thereby enriching the information available for reasoning. Inspired by these works, we introduce two strategies to boost the reasoning power of LLM for our task.

**Strategies for facilitating domain-specific reasoning.** We identify two primary limitations preventing advanced reasoning capability. Firstly, LLMs may not distinguish among various types of dietary patterns (foods and habits) when examining their correlation with opioid misuse. This issue could derive from the lack of domain knowledge of the relationship between specific dietary patterns and opioid misuse, rendering LLMs unable to identify specific foods or habits crucial for detecting opioid misuse.

Secondly, the presence of outliers among users poses another challenge. For instance, certain opioid users might exhibit dietary

**(a)  Domain Knowledge-Enhanced Prompt**

**Instruction:** Act as a skilled nutritionist. Your task is to judge if these dietary patterns are important in terms of predicting opioid misuse.

**Context:** User A is an opioid user. **(Users with Shared Dietary Patterns):** User A shares similar patterns with 10 opioid users and 2 non-opioid users. **(Dietary Pattern with Attention):** The most important factors for predictions are: “Food 1: Chocolate Cake, ...”, and “Habit 1: Drink Alcohol more than average, ...”. The other food and habits items are ...

**Notes:** You should keep your answer concise and accurate.

**(b)  LLM Responses**

Based on the information provided, when deciding whether the User A is an opioid user, the most indicative dietary food items are “salty ham; ...”, the most indicative dietary habit items are “add lots of salt at table; ...”.

**Figure 4: The prompts (highlighted in blue and red) generated from the NR-HGNN for interpreting key dietary patterns that indicate opioid misuse in individuals.**

patterns more commonly associated with non-opioid users, misleading LLMs in interpreting key dietary patterns of opioid misuse.

To address the first limitation, we introduce *Dietary Pattern with Attention*, which aims to help the LLM leverage critical dietary patterns that NR-HGNN learns in detecting users with opioid misuse. Specifically, we utilize attention scores retrieved in Micro-level Aggregation (Section 4.1.2) to rank foods and habits according to their importance. We empirically select the top-10 foods and habits for each user. By highlighting the importance of specific dietary patterns, we reduce the uncertainty of the LLM, thereby enhancing the reliability of their final interpretations. For the second limitation, we introduce the *Users with Shared Dietary Patterns* strategy to assess the typicality of users based on shared dietary patterns. This strategy aids in determining if dietary patterns accurately indicate opioid usage. In particular, we define typical users as individuals who share patterns with a substantial number of peers holding the same labels. We evaluate the user typicality on the refined graph (Equation 11), as it preserves more reliable patterns.

In this way, we avoid the LLM to capture the pseudo-correlation between dietary patterns and interpretations. These two strategies complement each other, which enables the LLM offering more accurate analyses by considering both individual and shared dietary patterns.

## 5 EXPERIMENTS

### 5.1 Experiment Setup

**Baselines.** To demonstrate the superiority of NR-HGNN in detecting users with opioid misuse, we select a range of graph learning baselines. These include vanilla MLP, homogeneous GNNs such as GCN [26], GraphSAGE (SAGE) [19], and GAT [51], alongside heterogeneous GNNs like RGCN [41], HAN [54], HGT [22], and the state-of-the-art Simple-HGN (SHGN) [27] and SeHGNN [61].

**Evaluation Settings.** In this study, we evaluate two settings with distinct training, validation, and testing splits. Specifically, 20% and 40% of nodes are selected randomly to create the training sets, with the remainder being evenly divided into validation and testing

**Table 1: Performance comparison between NR-HGNN and existing methods for identifying users with opioid misuse. Results are reported as mean  $\pm$  std%. The best performance is bolded and runner-ups are underlined.**

| Methods               | Accuracy                         | Train 20% – Valid 40% – Test 40% |                                  |                                  | ROC-AUC                          | Accuracy                         | Train 40% – Valid 30% – Test 30% |                                  |                                  | ROC-AUC                          |
|-----------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|
|                       |                                  | Precision                        | Recall                           | F1-Score                         |                                  |                                  | Precision                        | Recall                           | F1-Score                         |                                  |
| MLP                   | 70.15 $\pm$ 0.18                 | 68.31 $\pm$ 0.69                 | 75.19 $\pm$ 0.17                 | 71.57 $\pm$ 0.45                 | 77.52 $\pm$ 0.08                 | 70.40 $\pm$ 0.33                 | 68.33 $\pm$ 0.94                 | 76.13 $\pm$ 0.20                 | 72.00 $\pm$ 0.40                 | 77.38 $\pm$ 0.04                 |
| GCN[26]               | 73.26 $\pm$ 0.24                 | 69.87 $\pm$ 0.20                 | 81.72 $\pm$ 0.61                 | 75.33 $\pm$ 0.28                 | 78.09 $\pm$ 0.12                 | 72.94 $\pm$ 0.21                 | 68.98 $\pm$ 0.18                 | 83.36 $\pm$ 0.54                 | 75.49 $\pm$ 0.24                 | 78.10 $\pm$ 0.12                 |
| SAGE[19]              | 73.93 $\pm$ 0.25                 | 70.08 $\pm$ 0.27                 | 83.46 $\pm$ 0.56                 | 76.19 $\pm$ 0.25                 | 78.62 $\pm$ 0.14                 | 73.11 $\pm$ 0.27                 | 69.05 $\pm$ 0.17                 | 83.77 $\pm$ 0.56                 | 75.70 $\pm$ 0.30                 | 78.50 $\pm$ 0.11                 |
| GAT[51]               | 75.06 $\pm$ 0.12                 | 69.96 $\pm$ 0.12                 | 87.78 $\pm$ 0.36                 | 77.87 $\pm$ 0.14                 | 80.93 $\pm$ 0.02                 | 74.54 $\pm$ 0.06                 | 69.57 $\pm$ 0.05                 | 87.24 $\pm$ 0.17                 | 77.41 $\pm$ 0.07                 | 80.33 $\pm$ 0.02                 |
| RGCN[41]              | 75.64 $\pm$ 0.09                 | 70.15 $\pm$ 0.21                 | 89.24 $\pm$ 0.37                 | 78.55 $\pm$ 0.03                 | <u>81.27<math>\pm</math>0.02</u> | <u>74.82<math>\pm</math>0.17</u> | 69.82 $\pm$ 0.10                 | 87.43 $\pm$ 0.34                 | 77.64 $\pm$ 0.18                 | 80.93 $\pm$ 0.03                 |
| HGT[22]               | <u>75.72<math>\pm</math>0.06</u> | 69.38 $\pm$ 0.06                 | <u>92.02<math>\pm</math>0.01</u> | <u>79.11<math>\pm</math>0.04</u> | 81.14 $\pm$ 0.01                 | 74.62 $\pm$ 0.05                 | 68.96 $\pm$ 0.04                 | <u>89.54<math>\pm</math>0.07</u> | <u>77.92<math>\pm</math>0.04</u> | 80.69 $\pm$ 0.01                 |
| HAN[54]               | 75.10 $\pm$ 0.02                 | 70.37 $\pm$ 0.01                 | 86.64 $\pm$ 0.03                 | 77.66 $\pm$ 0.02                 | 80.02 $\pm$ 0.01                 | 74.26 $\pm$ 0.03                 | 69.66 $\pm$ 0.02                 | 85.95 $\pm$ 0.07                 | 76.96 $\pm$ 0.04                 | 80.52 $\pm$ 0.01                 |
| SHGN[27]              | 73.92 $\pm$ 1.20                 | <u>70.51<math>\pm</math>1.33</u> | 82.34 $\pm$ 4.23                 | 75.90 $\pm$ 1.59                 | 80.80 $\pm$ 0.85                 | 73.74 $\pm$ 1.13                 | <u>71.20<math>\pm</math>0.89</u> | 79.77 $\pm$ 2.97                 | 75.21 $\pm$ 1.43                 | <u>81.06<math>\pm</math>0.68</u> |
| SeHGNN[61]            | 75.48 $\pm$ 0.17                 | 70.00 $\pm$ 0.20                 | 89.16 $\pm$ 0.60                 | 78.42 $\pm$ 0.20                 | 81.23 $\pm$ 0.18                 | 74.63 $\pm$ 0.22                 | 69.63 $\pm$ 0.23                 | 87.35 $\pm$ 0.59                 | 77.49 $\pm$ 0.23                 | 80.84 $\pm$ 0.19                 |
| <b>NR-HGNN (Ours)</b> | <b>91.23<math>\pm</math>0.63</b> | <b>88.55<math>\pm</math>1.53</b> | <b>94.74<math>\pm</math>1.19</b> | <b>91.52<math>\pm</math>0.54</b> | <b>96.50<math>\pm</math>0.44</b> | <b>91.68<math>\pm</math>0.24</b> | <b>89.65<math>\pm</math>0.52</b> | <b>94.24<math>\pm</math>0.60</b> | <b>91.89<math>\pm</math>0.23</b> | <b>96.62<math>\pm</math>0.34</b> |

**Table 2: Performance comparison with different variants of our proposed NR-HGNN, including (a.1) macro-level aggregation, (a.2) micro-level aggregation, and (a.3) graph refinement with noise reduction.**

| (a.1) | (a.2) | (a.3) | Accuracy                         | Train 20% – Valid 40% – Test 40% |                                  |                                  | ROC-AUC                          | Accuracy                         | Train 40% – Valid 30% – Test 30% |                                  |                                  | ROC-AUC                          |
|-------|-------|-------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|----------------------------------|
|       |       |       |                                  | Precision                        | Recall                           | F1-Score                         |                                  |                                  | Precision                        | Recall                           | F1-Score                         |                                  |
| ✓     | –     | –     | 84.75 $\pm$ 0.72                 | 82.60 $\pm$ 0.56                 | 88.03 $\pm$ 1.11                 | 85.22 $\pm$ 0.74                 | 91.41 $\pm$ 0.42                 | 85.08 $\pm$ 0.62                 | 82.90 $\pm$ 0.58                 | 88.38 $\pm$ 1.04                 | 85.55 $\pm$ 0.64                 | 91.53 $\pm$ 0.35                 |
| –     | ✓     | –     | 73.96 $\pm$ 0.94                 | 72.23 $\pm$ 1.59                 | 78.12 $\pm$ 6.22                 | 74.90 $\pm$ 2.07                 | 81.39 $\pm$ 0.72                 | 75.30 $\pm$ 0.86                 | 73.01 $\pm$ 1.04                 | 80.43 $\pm$ 4.64                 | 76.45 $\pm$ 1.65                 | 82.35 $\pm$ 0.16                 |
| ✓     | ✓     | –     | <u>89.46<math>\pm</math>0.43</u> | <u>87.50<math>\pm</math>1.51</u> | 92.12 $\pm$ 1.89                 | <u>89.72<math>\pm</math>0.42</u> | <u>95.01<math>\pm</math>0.14</u> | <u>88.84<math>\pm</math>0.77</u> | <u>86.74<math>\pm</math>2.20</u> | 91.84 $\pm$ 2.54                 | <u>89.16<math>\pm</math>0.67</u> | <u>95.47<math>\pm</math>0.13</u> |
| ✓     | –     | ✓     | 86.99 $\pm$ 0.70                 | 80.79 $\pm$ 0.95                 | <b>97.06<math>\pm</math>0.45</b> | 88.18 $\pm$ 0.57                 | 92.69 $\pm$ 0.45                 | 88.34 $\pm$ 0.82                 | 83.65 $\pm$ 1.31                 | <b>95.36<math>\pm</math>0.71</b> | 89.11 $\pm$ 0.68                 | 93.35 $\pm$ 0.50                 |
| –     | ✓     | ✓     | 70.91 $\pm$ 10.37                | 65.65 $\pm$ 9.45                 | 94.74 $\pm$ 3.05                 | 77.04 $\pm$ 5.73                 | 76.28 $\pm$ 12.1                 | 60.02 $\pm$ 8.71                 | 56.92 $\pm$ 8.20                 | <u>95.14<math>\pm</math>3.42</u> | 70.77 $\pm$ 4.58                 | 65.01 $\pm$ 9.51                 |
| ✓     | ✓     | ✓     | <b>91.23<math>\pm</math>0.63</b> | <b>88.55<math>\pm</math>1.53</b> | <u>94.74<math>\pm</math>1.19</u> | <b>91.52<math>\pm</math>0.54</b> | <b>96.50<math>\pm</math>0.44</b> | <b>91.68<math>\pm</math>0.24</b> | <b>89.65<math>\pm</math>0.52</b> | 94.24 $\pm$ 0.60                 | <b>91.89<math>\pm</math>0.23</b> | <b>96.62<math>\pm</math>0.34</b> |

sets. Each model is executed on ten consecutive seeds to reduce the influence of randomness. To thoroughly assess the model’s performance across various dimensions, we employ five widely recognized metrics: Accuracy, Precision, Recall, F1-Score, and ROC-AUC score. For a balanced comparison, all methods are configured with a hidden dimension of 256 and utilize the Adam optimizer with a learning rate of 0.001 and a weight decay of 0.001. We fix the total number of epochs at 500. For baseline methods that incorporate an attention mechanism, the number of attention heads is set to four.

## 5.2 Results

**Performance Comparison.** Table 1 demonstrates the performance of baselines and our NR-HGNN in detecting users with opioid misuse on the proposed *NHANES Dietary Graph*, where our NR-HGNN notably surpasses existing methods, including the state-of-the-art SeHGNN. This substantial performance improvement is attributed to the model’s adeptness at concurrently capturing low and high-level information with micro- and macro-level aggregations from the graph, along with graph refinement for noise reduction. Contrary to expectations, increasing training data results in a slight decrease in the performance of most baselines, which might stem from potential over-fitting, where additional data may introduce redundant information. Nevertheless, our model performs better on the 40% training set split, which indicates the robustness of our methods. Notably, conventional methods demonstrate limitations in accurately identifying non-opioid users, as indicated by high recall but low precision values. This discrepancy may arise from these models accentuating patterns common to both opioid and

non-opioid users. In contrast, our NR-HGNN addresses this challenge through its graph refinement process, thereby enhancing the reliability of patterns used for detecting users with opioid misuse.

**Ablation Study.** Our method performs significantly better than other baselines. Therefore, to discern the elements contributing to such performance enhancement, we have conducted a comprehensive ablation study, as depicted in Table 2. Our observations reveal that each component independently enhances the model’s proficiency in detecting users with opioid misuse. When these three components are integrated, the model demonstrates optimal performance, highlighting the synergistic effect of concurrently processing high-order and low-order information alongside noise reduction refinement. Notably, employing either low-order or high-order information in isolation results in a diminished effectiveness of the refinement process. This outcome underscores the importance of synergistically aggregating information at both micro- and macro- levels to capture diverse facets of the graph, thereby ensuring the reliability of the refined patterns.

**Visualization.** To elucidate the efficacy of our NR-HGNN model, we execute a visualization analysis using t-SNE, which elucidates the decision boundary between opioid and non-opioid users as depicted in Figure 5. Relative to established methods such as GCN, RGCN, and the advanced SeHGNN, NR-HGNN exhibits superior clustering performance, as indicated by its compact intra-class distances and expansive inter-class distances. Conversely, the other three baselines tend to cluster users into smaller and more disparate groups. Furthermore, NR-HGNN demonstrates significantly higher Normalized Mutual Information (NMI) and Adjusted Rand

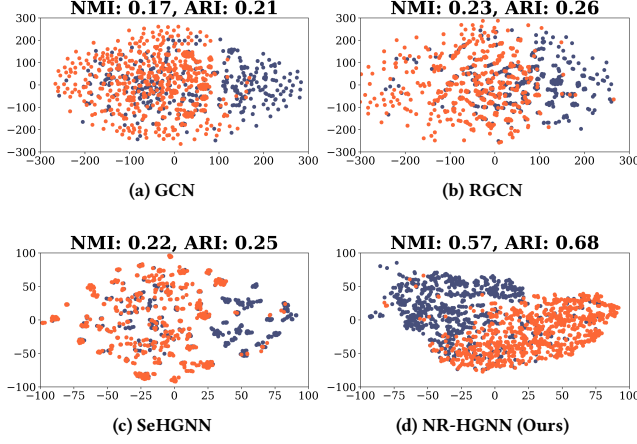


Figure 5: T-SNE Visualization of opioid (blue) and non-opioid (orange) users embeddings.

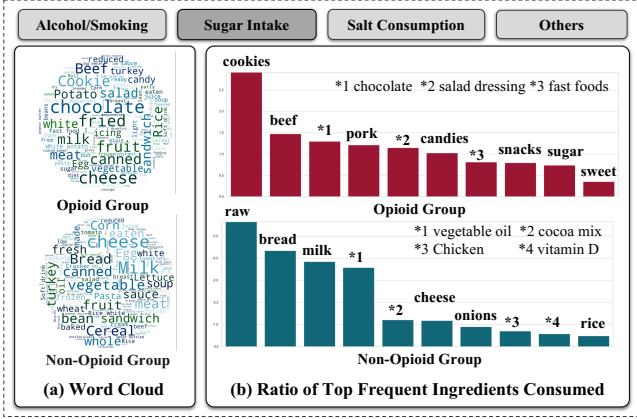


Figure 6: System Demo - Reasoning analysis on food patterns. (a) Word clouds generated from the ingredient names for both opioid and non-opioid groups. (b) The top ten frequent ingredients consumed by opioid and non-opioid groups. The analysis on dietary habits is in Appendix D.

Index (ARI) values, underlining its superiority. This enhanced performance may be attributed to NR-HGNN’s ability to effectively discern reliable patterns within the *NHANES Dietary Graph*.

## 6 DEEP DIVE INTO THE DIETARY PATTERNS

In addition to detecting potential opioid misuses, our Diet-ODIN also provides the provision of interpretive analyses of predictions. To this end, we select a random subset of 250 users, prompted with the dietary patterns with attention and users with shared dietary patterns using NR-HGNN. In alignment with the methodology delineated in Section 4.2, we conduct textual analysis and subsequent evaluations on the output. To vividly demonstrate our findings and the process, we deploy a **demonstration for our system**. One of the primary sections of our system is the statistical analysis panel

Table 3: Nutrient correlation analysis (t-test) reveals a strong correlation between potential opioid users and nutrient intake, corroborating Diet-ODIN’s findings on sugar intake.

| Nutrient | Daily Avg.<br>(Opioid Users) | Daily Avg.<br>(Regular Users) | p-value |
|----------|------------------------------|-------------------------------|---------|
| Sugar    | 111 (g)                      | 99 (g)                        | < 0.001 |
| Calories | 1867 (kcal)                  | 1726 (kcal)                   | < 0.001 |
| Sodium   | 2929 (mg)                    | 2749 (mg)                     | < 0.001 |
| Caffeine | 172 (mg)                     | 78 (mg)                       | < 0.001 |

to manifest the result of our qualitative analysis. We will dive deep into opioid users’ dietary patterns by demonstrating the statistical analysis panel of Diet-ODIN.

### 6.1 Reasoning Analysis

The statistical analysis panel categorizes and visually presents the results from the reasoning module. Take Sugar intake for example, Figure 6.(a) demonstrates a word cloud contrasting keyword frequencies between opioid and non-opioid user groups. As we can see, the non-opioid group demonstrates a wider variety of dietary preferences, whereas the opioid group frequently references items like "chocolate", "cookie". This trend is echoed in the adjacent bar chart (Figure 6.(b)), where the ingredient mentions of the opioid group, normalized based on group population size, reveal a similar pattern on sugar intensive ingredients. These results suggest a potential link between opioid use and sugar consumption. In addition to sugar intake, Diet-ODIN has identified several other noteworthy patterns. For example, dietary habits related to salt consumption, alcohol use, and smoking are prominent in our analysis, suggesting potential associations (as demonstrated in Appendix D). To further identify the reliability of the conclusions drawn from the reasoning process, we provide analyses to show that the observations align with existing literature and statistical analysis in the section below.

### 6.2 Qualitative Analysis

Diet-ODIN identifies key dietary factors—such as *sugar intake*, *salt consumption*, and *alcohol and smoking behaviors*—as significant in detecting opioid misuse. To validate these findings, we analyze the NHANES dataset from 2003 onward, revealing patterns that support Diet-ODIN’s conclusions. Specifically, a t-test comparing nutritional components between opioid users and non-users (Table 3) shows a notable difference in sugar and calorie intake, with a p-value less than 0.001, indicating statistical significance. This is further corroborated by related research [1, 33, 36]. Moreover, a chi-square test is performed to examine dietary habits among opioid users (Table 4), confirming a strong connection between salt consumption and opioid misuse. This association is reinforced by dietary habits related to salt intake and the literature [42, 43], as well as by analyzing daily sodium intake (Table 3). Additionally, the connection between opioid misuse and alcohol and smoking behaviors is affirmed by statistical analysis (Table 4) and literature review [11, 65]. Collectively, these analyses validate Diet-ODIN’s effectiveness and reliability.

**Table 4: Habit correlation analysis (chi-square test) reveals a tight linkage between certain dietary habits and potential opioid misuse, thereby validating Diet-ODIN’s findings on alcohol and smoking behaviors, salt consumption, and new insights not previously explored.**

| Habit Category |                                       | # Opioid Users | # Total Users | p-value |
|----------------|---------------------------------------|----------------|---------------|---------|
| NHANES Data    |                                       | 2728           | 95872         | -       |
| Alco. & Smok.  | Uses tobacco rarely                   | 82             | 1852          | < 0.001 |
|                | Uses tobacco often                    | 383            | 3155          | < 0.001 |
|                | Drinks alcohol less than avg.         | 778            | 26419         | 0.2627  |
|                | Drinks alcohol more than avg.         | 835            | 8634          | < 0.001 |
|                | Light cigarette smoker                | 465            | 11317         | < 0.001 |
|                | Heavy cigarette smoker                | 164            | 3398          | < 0.001 |
| Salt Cons.     | Uses little to no salt in preparation | 764            | 20596         | < 0.001 |
|                | Uses lots of salt in preparation      | 912            | 34341         | < 0.010 |
|                | Adds little to no salt at table       | 674            | 26604         | < 0.001 |
|                | Adds lots of salt at table            | 499            | 9654          | < 0.001 |
| New            | Drinks little or no milk              | 503            | 11717         | < 0.001 |
|                | Drinks lots of milk                   | 1029           | 45901         | < 0.001 |
|                | Takes few or no supplements           | 1161           | 53782         | < 0.001 |
|                | Takes lots of supplements             | 770            | 14027         | < 0.001 |

### 6.3 Dietary Pattern Discovery

Beyond identifying established dietary patterns related to opioid misuse, Diet-ODIN also unveils significant yet previously unexplored factors critical for detecting potential opioid misuse, providing new insights for future research aimed at combating opioid misuse. Notably, Diet-ODIN highlights the unique influence of *milk consumption* and *nutritional supplement intake* on opioid misuse, revealing that users with opioid misuse are (1) less likely to consume milk and (2) more inclined to rely on nutritional supplements instead of traditional food sources for their nutrient intake. This finding is further supported by statistical analysis (Table 4) of the NHANES database. Such observations imply that users with opioid misuse may neglect their diet, however to our best knowledge, little literature has explored these specific directions. These emerging patterns warrant deeper investigation for future research avenues.

## 7 CONCLUSION

In this study, we introduce Diet-ODIN, an innovative framework designed to detect users with opioid misuse through interpretable analysis of dietary patterns. Diet-ODIN establishes a large-scale multifaceted dietary benchmark dataset and proposes the NR-HGNN model, which adeptly utilizes both shared and individual dietary patterns with a denoising refinement process. To uncover critical factors in opioid misuse identification, Diet-ODIN employs LLMs for reasoning and implements two innovative strategies to boost reasoning capabilities. The performance of NR-HGNN in detecting users with opioid misuse significantly exceeds that of various baseline models. Furthermore, our qualitative analysis verifies that the reasoning outcomes align with existing literature and statistical analysis. Moreover, we build a system to demonstrate these

analyses, offering fresh perspectives, illuminating potential, more effective intervention strategies for addressing the opioid crisis.

## ACKNOWLEDGMENTS

This work was partially supported by the NSF under grants IIS-2321504, IIS-2334193, IIS-2340346, IIS-2203262, IIS-2217239, CNS-2203261, and CMMI-2146076. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

## REFERENCES

- [1] Nicole M Avena, Pedro Rada, and Bartley G Hoebel. 2008. Evidence for sugar addiction: behavioral and neurochemical effects of intermittent, excessive sugar intake. *Neuroscience & Biobehavioral Reviews* (2008).
- [2] Andrew Blair-Stanek, Nils Holzenberger, and Benjamin Van Durme. 2023. Can GPT-3 perform statutory reasoning? *arXiv* (2023).
- [3] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. In *NeurIPS*.
- [4] CDC. 2020. *Americans Share Hopeful Stories of Recovery From Opioid Use Disorder*. [https://www.cdc.gov/rxawareness/pdf/articles/TA-T3D2-English\\_MatteArticle\\_Release\\_508.pdf](https://www.cdc.gov/rxawareness/pdf/articles/TA-T3D2-English_MatteArticle_Release_508.pdf).
- [5] CDC. 2024. *National Health and Nutrition Examination Survey*. <https://www.nchs.gov/nhanes/default.aspx>.
- [6] Melody N Chavez and Khary K Rigg. 2020. Nutritional implications of opioid use disorder: A guide for drug treatment providers. *Psychology of Addictive Behaviors* (2020).
- [7] PM Cunningham. 2016. The use of sobriety nutritional therapy in the treatment of opioid addiction. *J Addict Res Ther* (2016).
- [8] Carrie Dennett. 2021. Diet’s Role in Opioid Recovery. *Today’s Dietitian* (2021).
- [9] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Bert: Pre-training of deep bidirectional transformers for language understanding. In *NAACL*.
- [10] Yuxiao Dong, Nitesh V Chawla, and Ananthram Swami. 2017. metapath2vec: Scalable representation learning for heterogeneous networks. In *KDD*.
- [11] Ola Ekholm, Morten Grønbaek, Vera Peuckmann, and Per Sjøgren. 2009. Alcohol and smoking behavior in chronic pain patients: the role of opioids. *European Journal of Pain* (2009).
- [12] Gokhan Ertug, Julia Brennecke, Balázs Kovács, and Tengjian Zou. 2022. What does homophily do? A review of the consequences of homophily. *Academy of Management Annals* (2022).
- [13] Yujie Fan, Yiming Zhang, Yanfang Ye, and Xin Li. 2018. Automatic Opioid User Detection from Twitter: Transductive Ensemble Built on Different Meta-graph Based Similarities over Heterogeneous Information Network.. In *IJCAI*.
- [14] Center for Substance-Abuse-Treatment. 2005. Medication-assisted treatment for opioid addiction in opioid treatment programs. (2005).
- [15] Sajjad Fouladvand, Jeffery Talbert, Linda P Dwoskin, Heather Bush, Amy Lynn Meadows, Lars E Peterson, Ramakanth Kavuluru, and Jin Chen. 2021. Predicting Opioid Use Disorder from Longitudinal Healthcare Data using Multi-stream Transformer. *arXiv* (2021).
- [16] Yinyu Fu, Jiani Zhang, Ziqiao Meng, and Irwin King. 2020. Magnn: Metapath aggregated graph neural network for heterogeneous graph embedding. In *WWW*.
- [17] Aditya Grover and Jure Leskovec. 2016. node2vec: Scalable feature learning for networks. In *KDD*.
- [18] Ja K Gu, Penelope Allison, Alexis Grimes Trotter, Luenda E Charles, Claudia C Ma, Matthew Groenewold, Michael E Andrew, and Sara E Luckhaupt. 2022. Prevalence of Self-Reported Prescription Opioid Use and Illicit Drug Use Among US Adults: NHANES 2005–2016. *Journal of occupational and environmental medicine* (2022).
- [19] Will Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. In *NeurIPS*.
- [20] Dae-Hee Han, Shieun Lee, and Dong-Chul Seo. 2020. Using machine learning to predict opioid misuse among US adolescents. *Preventive medicine* (2020).
- [21] Chuanbo Hu, Bin Liu, Yanfang Ye, and Xin Li. 2023. Fine-grained classification of drug trafficking based on Instagram hashtags. *Decision Support Systems* (2023).
- [22] Ziniu Hu, Yuxiao Dong, Kuansan Wang, and Yizhou Sun. 2020. Heterogeneous graph transformer. In *WWW*.
- [23] Kendall D Jaynes and E Leigh Gibson. 2017. The importance of nutrition in aiding recovery from substance use disorders: A review. *Drug and alcohol dependence* (2017).
- [24] Gyanprakash A Ketwaroo, Vivian Cheng, and Anthony Lembo. 2013. Opioid-induced bowel dysfunction. *Current gastroenterology reports* (2013).
- [25] Ali Kheradmand and Azadeh Kheradmand. 2020. Nutritional status in patients under methadone maintenance treatment. *Journal of Substance Use* (2020).

- [26] Thomas N Kipf and Max Welling. 2017. Semi-supervised classification with graph convolutional networks. In *ICLR*.
- [27] Qingsong Lv, Ming Ding, Qiang Liu, Yuxiang Chen, Wenzheng Feng, Siming He, Chang Zhou, Jianguo Jiang, Yuxiao Dong, and Jie Tang. 2021. Are we really making much progress? revisiting, benchmarking and refining heterogeneous graph neural networks. In *KDD*.
- [28] Nadine Mahboub, Rana Rizk, Mirey Karavetian, and Nanne de Vries. 2021. Nutritional status and eating habits of people who use drugs and/or are undergoing treatment for recovery: a narrative review. *Nutrition reviews* (2021).
- [29] Adriana Matos, David L Bankes, Kevin T Bain, Tyler Ballinghoff, and Jacques Turgeon. 2020. Opioids, polypharmacy, and drug interactions: A technological paradigm shift is needed to ameliorate the ongoing opioid epidemic. *Pharmacy* (2020).
- [30] Elizabeth McDonald and Jennifer Laurent. 2019. Hedonic eating behaviors and food preferences associated with medication-assisted treatment for opioid use disorder. *Journal of Opioid Management* (2019).
- [31] A. Morabia, J. Fabre, E. Ghee, S. Zeger, E. Orsat, and A. Robert. 1989. Diet and Opiate Addiction: a quantitative assessment of the diet of non-institutionalized opiate addicts. *British Journal of Addiction* (1989).
- [32] David J Mysels and Maria A Sullivan. 2010. The relationship between opioid and sugar intake: Review of evidence and clinical applications. *Journal of Opioid Management* (2010).
- [33] David J Mysels and Maria A Sullivan. 2010. The relationship between opioid and sugar intake: review of evidence and clinical applications. *Journal of opioid management* (2010).
- [34] Sepideh Nabipour, AYU Mas, and Mohd Hussain Habil. 2014. Burden and nutritional deficiencies in opiate addiction-systematic review article. *Iranian journal of public health* (2014).
- [35] NIDA. 2024. *Opioids*. <https://www.drugabuse.gov/drug-topics/opioids>.
- [36] Taylor A Ochalek, Jennifer Laurent, Gary J Badger, and Stacey C Sigmon. 2021. Sucrose subjective response and eating behaviors among individuals with opioid use disorder. *Drug and Alcohol Dependence* (2021).
- [37] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. 2014. Deepwalk: Online learning of social representations. In *KDD*.
- [38] Yiyue Qian, Yiming Zhang, Yanfang Ye, Chuxu Zhang, et al. 2021. Distilling meta knowledge on heterogeneous graph for illicit drug trafficker detection on social media. In *NeurIPS*.
- [39] Khary K Rigg and Gladys E Ibañez. 2010. Motivations for non-medical prescription drug use: A mixed methods analysis. *Journal of Substance Abuse Treatment* (2010).
- [40] Andrew Rosenblum, Lisa A Marsch, Herman Joseph, and Russell K Portenoy. 2008. Opioids and the treatment of chronic pain: controversies, current status, and future directions. *Experimental and Clinical Psychopharmacology* (2008).
- [41] 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 *ESWC*.
- [42] Craig M Smith, Joshua BB Garfield, Aparna Attawar, Dan I Lubman, and Andrew J Lawrence. 2019. The influence of opioid dependence on salt consumption and related psychological parameters in mice and humans. *Drug and alcohol dependence* (2019).
- [43] Craig M Smith and Andrew J Lawrence. 2018. Salt appetite, and the influence of opioids. *Neurochemical Research* (2018).
- [44] Guijin Son, Haneul Jung, Moonjeong Hahm, Keonju Na, and Sol Jin. 2023. Beyond Classification: Financial Reasoning in State-of-the-Art Language Models. *arXiv* (2023).
- [45] Jian Tang, Meng Qu, Mingzhe Wang, Ming Zhang, Jun Yan, and Qiaozhu Mei. 2015. Line: Large-scale information network embedding. In *WWW*.
- [46] Ruixiang Tang, Xiaotian Han, Xiaoqian Jiang, and Xia Hu. 2023. Does synthetic data generation of llms help clinical text mining? *arXiv* (2023).
- [47] Lauren J. Tanz, Amanda T. Dinwiddie, Christine L. Mattson, Julie O'Donnell, and Nicole L. Davis. 2022. Drug Overdose Deaths Among Persons Aged 10–19 Years - United States, July 2019–December 2021. *Morbidity and Mortality Weekly Report* (2022).
- [48] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv* (2023).
- [49] Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shriti Bhosale, et al. 2023. Llama 2: Open foundation and fine-tuned chat models. *arXiv* (2023).
- [50] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *NeurIPS*.
- [51] Petar Velicković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2018. Graph attention networks. In *ICLR*.
- [52] Freya Waddington, Mark Naunton, Greg Kyle, and Gabrielle Cooper. 2015. Nutritional intake of opioid replacement therapy patients in community pharmacies: A pilot study. *Nutrition & Dietetics* (2015).
- [53] Xiao Wang, Deyu Bo, Chuan Shi, Shaohua Fan, Yanfang Ye, and S Yu Philip. 2022. A survey on heterogeneous graph embedding: methods, techniques, applications and sources. *TBD* (2022).
- [54] Xiao Wang, Houye Ji, Chuan Shi, Bai Wang, Yanfang Ye, Peng Cui, and Philip S Yu. 2019. Heterogeneous graph attention network. In *WWW*.
- [55] Xiao Wang, Nian Liu, Hui Han, and Chuan Shi. 2021. Self-supervised heterogeneous graph neural network with co-contrastive learning. In *KDD*.
- [56] Zehong Wang, Qi Li, Donghua Yu, Xiaolong Han, Xiao-Zhi Gao, and Shigen Shen. 2023. Heterogeneous graph contrastive multi-view learning. In *SDM*.
- [57] Qianlong Wen, Zhongyu Ouyang, Jianfei Zhang, Yiyue Qian, Yanfang Ye, and Chuxu Zhang. [n. d.]. Disentangled dynamic heterogeneous graph learning for opioid overdose prediction. In *KDD*.
- [58] Megan C Whatnall, Janelle Skinner, Kirrilly Pursey, Katherine Brain, Rebecca Collins, Melinda J Hutchesson, and Tracy L Burrows. 2021. Efficacy of dietary interventions in individuals with substance use disorders for illicit substances or illicit use of pharmaceutical substances: A systematic review. *Journal of Human Nutrition and Dietetics* (2021).
- [59] Siheng Xiong, Ali Payani, Ramana Kompella, and Faramarz Fekri. 2024. Large language models can learn temporal reasoning. *arXiv* (2024).
- [60] Kailai Yang, Shaoxiong Ji, Tianlin Zhang, Qianqian Xie, Ziyun Kuang, and Sophia Ananiadou. 2023. Towards interpretable mental health analysis with large language models. In *EMNLP*.
- [61] Xiaocheng Yang, Mingyu Yan, Shirui Pan, Xiaochun Ye, and Dongrui Fan. 2023. Simple and efficient heterogeneous graph neural network. In *AAAI*.
- [62] Yuan Yang, Siheng Xiong, Ali Payani, Ehsan Shareghi, and Faramarz Fekri. 2023. Harnessing the power of large language models for natural language to first-order logic translation. *arXiv* (2023).
- [63] Chuxu Zhang, Dongjin Song, Chao Huang, Ananthram Swami, and Nitesh V Chawla. 2019. Heterogeneous graph neural network. In *KDD*.
- [64] Jianfei Zhang, Ai-Te Kuo, Jianan Zhao, Qianlong Wen, Erin Winstanley, Chuxu Zhang, and Yanfang Ye. [n. d.]. Rxnet: Rx-refill graph neural network for over-prescribing detection. In *CIKM*.
- [65] Michael J Zvolensky, Andrew H Rogers, Lorra Garey, Justin M Shepherd, and Joseph W Ditte. 2021. Opioid Misuse among Smokers with Chronic Pain: Relations with Substance Use and Mental Health. *Behavioral Medicine* (2021).

## A GRAPH CONSTRUCTION DETAILS

In general, the user information comes from the demographic sections of NHANES data, and the nutrition information is gathered from the dietary sections of NHANES data. The ingredients and category information come from the FNDDS dataset, while the dietary habits information comes from the questionnaire section of NHANES data. We processed the questionnaire data to construct dietary habit features. The details are described as follows:

**Nutritional Information.** The food records in NHANES data are recorded using Food and Nutrient Database for Dietary Studies (FNDDS) food code. This database, integral to USDA, catalogs food and beverage consumption in What We Eat In America (WWEIA) database and NHANES data to help researchers conduct enhanced analysis of nutrient values in dietary intakes. By leveraging this dataset, we link the food items in NHANES dataset with food ingredients and WWEIA food category information, and create *food*, *ingredient* and *category* nodes.

**Dietary Habit Information.** We compile dietary habit data from various NHANES tables, such as Diet Behavior table or Consumer Behavior table. Given the diverse nature of these features, traditional processing methods were inadequate. A team of four reviewers meticulously examines the NHANES data, identifying features indicative of dietary habits, such as the awareness towards healthy diet or the frequency of having frozen food. For each feature, we take the top 10% and bottom 10% of the users and assign these users the habit tags. For example, we take all those respond to milk drinking questionnaire, and assign the 10% users who drink the most milk the habit “drink lots of milk”, and the 10% users who

drink the least the habit “drink little or no milk”. In this way, we summarize 54 distinct habits that serves as the *habit* node in the graph. This habit information is of vital importance and faithfully reflect many insights into users’ dietary patterns as we demonstrate later in the paper.

As for the node features, we use the demographic information as the features for the user nodes, and the nutrient values as the features for the food nodes. To associate other node types with their content, we use the pre-trained BERT embeddings to encode the description of the nodes, serving as the corresponding node representations. The edges record the connections between the nodes and are free of weights or directions.

For the label for our benchmark dataset, we mark a user as an active user with opioid misuse if: 1) a user has records of using heroin within the year, or 2) a user has records of using a prescription opioid medicine continuously over 90 days, which is a criteria commonly employed in the domain [18]. In this way, we generate the labels and form our problem into a binary classification task. The detailed graph statistics can be seen in Table 5.

**Table 5: The statistics of NHANES Dietary Graph**

| Item      | Count                                                                                                        |
|-----------|--------------------------------------------------------------------------------------------------------------|
| Nodes     | # User = 4,826, # Food = 5,896, # Ingredient = 2,792,<br># Category = 174, # Habit = 54                      |
| Relations | # Food - User = 136,967, # Food - Ingredient = 19,410,<br># Food - Category = 5,896, # User - Habit = 46,947 |

## B EXPERIMENTAL SETTINGS

### B.1 Environmental Settings

The experiments are conducted in the Windows 10 operating system with 64GB of RAM. The training process involves the use of one NVIDIA GeForce RTX 3090 GPU and one NVIDIA A40 GPU, with the framework of Python 3.8.18, Pytorch 2.1.0 and Pytorch-geometric 2.4.0.

### B.2 Split Settings

We randomly split the data into a 2-4-4 group (20% training set, 40% valid set and 40% test set), and 4-3-3 group and conduct experiments on these two split settings. The random split is only performed once and all experiments are done on the same split seed.

For the noise reduction phase, since we use the user label to identify whether two users share the same label, we filter the user pairs to make sure the two users always come from the training set to avoid data leak. Also, at this stage, we use the model from the last training epoch instead of the one performs best on the validation set to avoid data leakage on valid data during inference.

### B.3 Hyperparameters

In our experiments, for the macro-aggregation stage, we set the hidden dimension to 64, and in micro-aggregation stage to 256. Both stages involve 4 attention heads. The pooling layer afterwards use a simple MLP with the concatenated hidden dimension of the two model as the input, and has hidden dimension of 256 as output. We perform Adaptive Moment Estimation (Adam) to optimize the

models with a learning rate of 0.001 and L2 regularization as 0.001 for 100 epochs. For the subsequent opioid user prediction module, we train a simple GCN layer for 500 epochs, which has a hidden dimension of 128, a learning rate of 0.0001, a L2 regularization as 0.001 and a dropout rate of 0.6. All experiments utilize the cross entropy as the loss function. In the training, all models are evaluated on the valid sets and the model state with the best valid set performance is used to evaluate the test set. To improve reproducibility, for each split setting, we evaluate the models using the same set of 10 consecutive seeds, and report their mean and standard deviation.

## C FINE-GRAINED CORRELATION BETWEEN DIETARY PATTERNS AND OPIOID MISUSE

To further analyze the fine-grained factors in detecting users with opioid misuse, we would analyze the impact of the ingredients to the foods, enabling us to discover more intricate patterns. In particular, we also utilize the attention score extracted from Contextual Aggregation (Section 4.1.2) to rank the importance of ingredients in detecting opioid users. We employ a similar context prompt, which is formulated as “Food A is <food description>. The model thinks the most important ingredients are: ingredients 1: <>, Ingredient 2: <>, ..., and the remaining ingredients 1: <>, Ingredient 2: <>, ..., and ”. Through the more fine-grained ingredient level analysis, this process helps us navigate to the most important ingredients that contributes to the identification of users with opioid misuse. For example, when cakes with chocolate icing is identified, we can further see it’s the ingredient chocolate, instead of eggs or flour that makes the food indicative.

## D MORE QUALITATIVE ANALYSIS

In this section, we further elaborate on the reasoning analysis and the qualitative analysis we do that demonstrate Diet-ODIN’s efficacy and reliability.

Figure 7 illustrates the analysis of alcohol and smoking habits, as selected via the top buttons in the system. The bar chart below the buttons presents the most frequent habits counts among opioid and non-opioid user groups from the reasoning result. The chart includes four specific habits related to alcohol and cigarette use (e.g., *op1*-drinking alcohol more, *nonop6*-drinking alcohol less than average, *op2*-heavy cigarette smoking, *nonop9*-light cigarette smoking), indicating a significant correlation between opioid usage and alcohol or cigarette consumption. Additionally, a word cloud, positioned next to the bar chart, highlights the most common terms associated with these habits. Notably, terms like “drinks alcohol” and “cigarette smoker” predominate in the opioid user group. Similar patterns can be found in salt consumption. Opioid users tend to have habits related to using more salt, such as *op4*-Adds lots of salt at table, and *op6*-Uses lots of salt in preparation. Similarly, in next paragraphs, we utilize statistical analysis and existing literature to show Diet-ODIN’s reasoning is aligned with real-world scenarios.

**Opioid Use and Salt Consumption.** As evidenced by Table 4, all patterns related to salt consumption are statistically correlated. This finding is further substantiated by the analysis of daily average sodium intake in Table 3, where the p-value indicates a statistically significant difference between the opioid user group and the non-opioid user group. The literature corroborates this pattern as well.

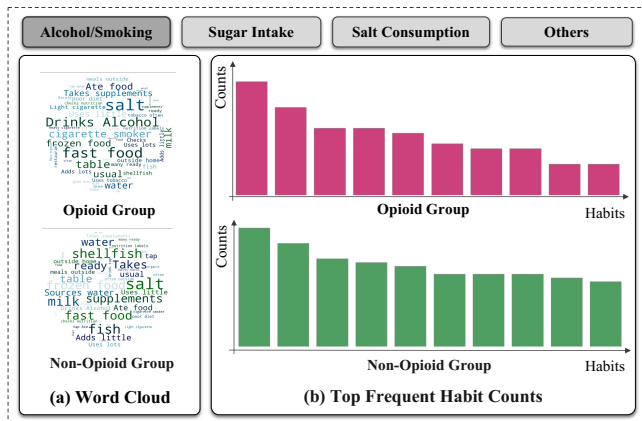


Figure 7: Reasoning analysis for habits: (a) Word cloud generated from the habit names for both opioid and non-opioid groups. (b) The top ten frequent habit counts by opioid and non-opioid groups. Because of the length of the habit names, we hide the name captions and refer by its column sequence. For example, *op1* and *nonop - 1* refers to the leftmost bar of upper and lower barplot respectively.

For example, as Smith et al [42, 43] discussed in their research that opioid dependence can affect dietary salt intake, often resulting in a preference for foods with higher salt concentrations.

**Opioid Use and the Use of Alcohol and Smoking.** In a manner akin to the examination of salt consumption, numerous studies have investigated the relationship between opioid usage and the consumption of other substances. A well-established body of research indicates a positive correlation between opioid use and excessive alcohol consumption and smoking behavior, as noted in studies by [11, 65]. These associations are also detected in our analysis using the NHANES data. The majority of the alcohol and cigarette use patterns identified by our method have been corroborated as correlational. (We amalgamated ‘tobacco use’ and ‘cigarette smoking’ during the reasoning phase.)

## E ETHICS AND PRIVACY STATEMENT

Addressing privacy and ethical considerations is paramount, particularly when dealing with health-related sensitive data. The National Health and Nutrition Examination Survey (NHANES) exemplifies commendable practices in this domain, rigorously adhering to confidentiality safeguards as mandated by public legislation. This foundational commitment to privacy enables the attainment of our core research objectives within the framework of established survey policies. To be specific, the original NHANES dataset has performed an anonymization process to ensure the concealment of personally identifiable information (PII) such as social security number (SSN) or physical address. Despite the absence of PII, the dataset remains a

rich resource for comprehensive analysis, allowing the exploration of dietary patterns associated with opioid misuse introduced in the paper. When it comes to deployment, the usage of prediction results is also privacy-protected as a part of personal medical records. By operating within these parameters, we ensure our research upholds the highest standards of ethical integrity and privacy protection.

## F LIMITATIONS AND DISCUSSION

Although our developed *Diet-ODIN* has been demonstrated its outstanding performance on exploring the complex interplay between opioid misuse and dietary patterns, it could also be subject to certain limitations: data incompleteness and lacking quantitative measure of the interpretations. For data incompleteness, in this work, the quantities of foods are not accessible where analysis bias may exist. Due to the lack of food quantities, the developed model is not able to incorporate edge weights to capture more detailed dietary patterns, which is inclined to answer how frequent a user eats a certain food. Nevertheless, edge weights might be helpful in certain scenarios, for example, our method doesn't consider "coffee" an important factor because it is such a common drinking in daily life that everybody would drink. However, the drinking amounts weigh differently in one's dietary structure. In table-3, we demonstrate that opioid users on average consume more than double amount of caffeine than that of the regular users, which should be an indicator captured by both detection and reasoning tasks. This limitation can be mitigated by implementing edge-weighting scheme on user-food edges. Nevertheless, our experiments and analysis demonstrate the majority of dietary patterns can be captured through edge-weight free methods. Thus, we leave the edge weighting implementation for our future work.

For the lack of quantitative measure of the interpretations, our reasoning experiments mainly rely on statistical qualitative analysis, which posts influence by reflecting real-life patterns and pointing out future research directions. The consistent matching findings between Diet-ODIN and real-life statistics and literature indicate the efficacy and reliability of the patterns. However, the limitation remains that the reasoning process lacks control and quantitative benchmarks to evaluate its efficacy. We argue that our work set a strong baseline by providing insights and setup benchmarks that connect dietary and healthcare information in NHANES data. We appreciate the suggestions from the peer reviewers and agree that our findings would benefit more with quantitative experiments, and we leave this to future work to establish a standardized and quantitative benchmark for properly evaluating the interpretability of our methods against other baselines. Despite these limitations, our work in this paper is the first attempt to exploit AI-driven techniques for comprehensive exploration of the correlations between opioid misuse and users' diets with explainable outputs, which could provide new insights and pave promising research directions for other researchers and practitioners to combat the devastating and lethal opioid epidemic.