

ExplainMIX: Explaining Drug Response Prediction in Directed Graph Neural Networks With Multi-Omics Fusion

Ying Xiang , Xiaodi Li, Qian Gao, Junfeng Xia , and Zhenyu Yue 

Abstract—The intricacies of cancer present formidable challenges in achieving effective treatments. Despite extensive research in computational methods for drug response prediction, achieving personalized treatment insights remains challenging. Emerging solutions combine multiple omics data, leveraging graph neural networks to integrate molecular interactions into the reasoning process. However, effectively modeling and harnessing this information, as well as gaining the trust of clinical professionals remain complex. This paper introduces ExplainMIX, a pioneering approach that utilizes directed graph neural networks to predict drug responses with interpretability. ExplainMIX adeptly captures intricate structures and features within directed heterogeneous graphs, leveraging diverse data modalities such as genomics, proteomics, and metabolomics. ExplainMIX goes beyond prediction by generating transparent and interpretable explanations. Incorporating edge-level, meta-path, and graph structure information, it provides meaningful insights into factors influencing drug response, supporting clinicians and researchers in the development of targeted therapies. Empirical results validate the efficacy of ExplainMIX in prediction and interpretation tasks by constructing a quantitative evaluation ground truth. This approach aims to contribute to precision medicine research by addressing challenges in interpretable personalized drug response prediction within the landscape of cancer.

Index Terms—Drug response prediction, directed graph neural networks, explanation, multi-omics.

I. INTRODUCTION

Due to the intricate characteristics of cancer, effectively treating this disease continues to pose a significant challenge. It requires the meticulous selection of treatment strategies tailored for different cancer subtypes, commonly referred to as personalized cancer therapy [1]. Yet the majority of individuals with cancer do not currently benefit from such precision approaches [2]. In recent years, numerous scientific communities have made significant contribution by providing vast amounts of pharmacogenomics data, such as Genomics of Drug Sensitivity in Cancer (GDSC) [3], Cancer Cell Line Encyclopedia (CCLE) [4] and Cancer Therapeutics Response Portal (CTRP) [5]. These projects offer extensive molecular profiling data, including somatic mutation, copy number aberration, proteomic data, among others, for a diverse array of cancer cell lines. Additionally, they encompass data on the response of these cell lines to various targeted therapies and chemotherapy. These invaluable resources have accelerated further investigations, with many researchers currently focusing on personalized cancer therapy using computational models [6], [7].

One particular area of interest is the study of drug response prediction [8], [9], which holds immense significance and poses numerous challenges in both bioinformatics and translational medicine. Some researchers have proposed the use of Graph Neural Networks (GNN) as a solution to this problem [10], [11], [12]. GNNs can capture nonlinear relationships and recursively propagate messages through the edges of the input graph, enabling the implementation of sophisticated prediction functions by leveraging prior knowledge [13], [14]. However, despite their notable strengths, there are still some stumbling blocks, such as the high dimensionality of omics data and feature redundancy. Furthermore, entrusting patients' medical decisions entirely to black box models has also sparked considerable controversy both ethically and practically. In addition, constructing the ground truth for explainable algorithms remains in the exploratory phase, making interpretation challenging.

The wealth of omics data on cancer cell lines adds complexity, necessitating a thorough analysis of its impact on prediction models [12]. Researchers have made notable progress, paving the way for leveraging graph structural information in cancer drug sensitivity prediction [15], [16], [17]. Each node, representing a cell line or a drug, holds unique characteristics within distinct feature spaces [18]. Most existing methods have

Received 11 February 2024; revised 30 January 2025; accepted 7 March 2025. Date of publication 11 March 2025; date of current version 4 July 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 62472005, and Grant 62102004, in part by the Open Fund of National Engineering Laboratory for Big Data System Computing Technology under Grant SZUBDSC-OF2024-18, and in part by the Graduate Education Quality Project of Anhui under Grant 2022xsx052 and Grant 2023jyjxggyjY072. (Corresponding authors: Zhenyu Yue; Junfeng Xia.)

Ying Xiang, Xiaodi Li, Qian Gao, and Zhenyu Yue are with the School of Information and Artificial Intelligence, Anhui Agricultural University, Hefei 230036, China (e-mail: yingxiang@stu.ahu.edu.cn; lixiaodi@stu.ahu.edu.cn; gaoqian@stu.ahu.edu.cn; zhenyuyue@ahu.edu.cn).

Junfeng Xia is with the Institutes of Physical Science and Information Technology, Anhui University, Hefei 230601, China (e-mail: jfxia@ahu.edu.cn).

The dataset and source code of ExplainMIX are publicly available at <https://github.com/AhauBioinformatics/ExplainMIX>.

Digital Object Identifier 10.1109/JBHI.2025.3550353

treated the cell line-drug bipartite network as an undirected graph. Nevertheless, the edges in the network represent either sensitivity or resistance, which indicates that it can be better conceptualized as a directed graph. It is crucial to incorporate heterogeneous and directed factors into research and analysis. The diverse interactions among nodes demonstrate asymmetry, with these asymmetries encapsulating vital structural information within the graph [19]. This is essential for advancing our understanding of personalized medicine and improving treatment outcomes [20].

Furthermore, a critical bottleneck lies in practitioners' trust and understanding of the GNN model, identifying and rectifying systematic patterns of mistakes made by models before deploying them in the real world [21]. GNNs face challenges in transparency, as they do not readily facilitate human-intelligible explanations for their predictions [22]. To advance precision medicine, it is essential to focus not only on improving clinical utility by enhancing accuracy in predicting drug responses for specific patients but also on increasing the reliability and translatability of computational methods [23]. While much work has focused on model interpretability [24], [25], [26], [27], [28], there is limited research on directed heterogeneous graphs. In these networks, nodes have distinct roles and connections, complicating the explanation of relationships [29]. Existing algorithms often treat homogeneous and heterogeneous edges equally, making it difficult to capture the causal influences between nodes. Furthermore, few studies have applied explanation methods to drug response prediction.

In this paper, we present a novel method for explaining drug response prediction, called ExplainMIX, which considers heterogeneity present in both the prediction and explanation processes. In particular, ExplainMIX incorporates various types of genomic data, including gene expression, somatic mutations, copy number variation, DNA methylation, proteomics and metabolomic data, as input to capture the complex biological information involved in drug response. Moreover, it aims to effectively capture the complex structure information within a heterogeneous graph. ExplainMIX employs the Relational Graph Convolutional Network (RGCN) [30] to model and leverage relationships among different nodes in the heterogeneous graph, enabling predictions of drug responses across various types. Finally, and most important, ExplainMIX goes beyond prediction and provides explanations for its results. It achieves this by leveraging edge-level information and meta-path logical semantics within the heterogeneous graph. These explanations help elucidate the underlying reasons for the predicted drug responses, providing valuable insights for clinical researchers.

The contributions of our work are summarized as follows.

- 1) Application of domain knowledge in prediction and explanation: Our method integrates edge-level, meta-path, and graph structure information to provide transparent and interpretable results, facilitating better understanding in the context of precision medicine.
 - 2) Development of a quantitative evaluation ground truth: We establish an objective framework for evaluating the quality of generated explanations, providing a reliable metric for interpretability.
 - 3) Improving prediction accuracy with directed graph neural networks: By modeling edges as distinct relationships between cells and drugs, we treat sensitivity and resistance separately, constructing a directed graph to address data imbalance and enhance prediction accuracy.
- 4) Leveraging diverse omics data for deeper insights: Our method integrates multiple omics modalities (e.g., genomics, proteomics, metabolomics) and uses perturbation experiments to explore their impact on prediction outcomes, offering a more comprehensive view of the mechanisms at play.

II. RELATED WORKS

A. Prediction Methods of Drug Response

The identification of drug response in individuals plays a pivotal role in advancing precision cancer medicine. Various categories of methods have been developed to address this problem [8], [31]. Each of these approaches offers unique advantages and insight in predicting drug response and hold promise for advancing our understanding in this area. Existing methods can be broadly classified into three categories: matrix factorization methods, network differential analysis methods and deep learning methods.

- i) Matrix factorization methods: These methods typically involve the decomposition of a drug response matrix into two matrices representing the cell lines and drugs, respectively. By extending the concept of matrix factorization, the similarity between cell lines and drugs [32] or the associations between pathways and drug responses [33] are learned to enhance the overall performance. While they have not thoroughly explored the unique features and characteristics of drugs and cell lines, their performance is notably deficient when predicting responses for new drugs or cell lines.
- ii) Network differential analysis methods: These methods take into account the interactions between entities within networks and primarily focus on predicting the response of cell lines to the given drugs [34], [35]. However, limitation of these methods is their limited scalability and slow performance when it comes to predicting response for a large number of drugs [8].
- iii) Deep learning methods: In order to address the aforementioned challenges, researchers have turned to deep learning techniques. Tuan Nguyen et al. developed a neural network architecture that directly represents drugs as molecular graphs to capture the bonds among atoms to predict the response value of each drug-cell line pair [14]. Wei Peng et al. proposed the MOFGCN algorithm to predict drug response in cell lines by integrating multi-omics data and Graph Convolution Network with a linear correlation coefficient decoder [15]. Yiheng Zhu et al. proposed a novel approach (TGSA) for drug response prediction using Twin Graph Neural Networks and a Similarity Augmentation module [36]. These methods aim to leverage the power of deep neural networks to tackle the complexities of drug response prediction. While these models have shown promising performance in predicting responses for unseen datasets, researchers are actively working on improving the interpretability and transferability of these prediction methods.

It is crucial to ensure that the models can be reproduced across various settings and effectively transferred to clinical applications. This entails not only improving the robustness and generalizability of the models but also addressing the interpretability aspect to provide meaningful explanations for the predictions.

B. Explanation for Link Prediction in Graph Neural Network

GNNs are widely used in fields like disease diagnosis, image object detection and precision medicine in cancer, but their “Black-Box” nature limits transparency and traceability, hindering human-AI understanding. Explainable AI (XAI) addresses this by offering insights into GNN decisions, aiming to reduce bias, enhance accountability, and build trust in AI systems [37], [38]. Current XAI methods for GNNs in link prediction involve various approaches, including subgraph formulation [25], [38] and counterfactual explanations [27], [25].

GnnExplainer [25], [39] aim to select the important subgraph by maxing the mutual information (MI) between the model predictions and possible subgraphs. Given the well-trained GNN model and its predictions as inputs, the method returns the most important subgraph or features that contribute to the model’s prediction. This approach provides insights into the specific graph components that play a pivotal role in the GNN’s decision-making process. ExplainNE [27], [25], as a counterfactual explanations-based approach, quantifies the change in predicting probability by manipulating or perturbing the embeddings. This method evaluates the influence of modifying graph links on model predictions, offering shaping the model’s output. It provides insights into the significance and functionality of specific links or nodes in the graph structure, enhancing our understanding of how diverse connections impact overall predictions and improving the interpretability of GNNs. However, these explanation methods for graph neural networks face limitations in effectively handling diverse types of nodes and edges, especially when applied to heterogeneous graphs.

III. MATERIALS AND PREDICTION

A. Data Collection and Processing

In this study, we present a novel approach for drug response prediction (see Fig. 1) by leveraging multi-omics fusion and drug response data. We frame the problem as a link prediction task by integrating molecular data from the CCLE dataset (<https://sites.broadinstitute.org/cdle/datasets>) and drug response data from the GDSC dataset (<https://sites.broadinstitute.org/cdle/datasets>). Initially, we compiled data on 1406 cancer cell lines and 449 drugs, classifying drug responses into three categories: sensitivity, resistance, and no correlation. After rigorous data cleaning, 192 cell lines with complete multi-omics data—including transcriptomics, proteomics, copy number variations, mutations, DNA methylation, and metabolomics—were retained. Using multi-omics fusion, we integrated these diverse data types to calculate cell line similarities. This comprehensive approach provides valuable insights into biological systems, informing medical research and practice. Additionally, drug

features were obtained from PubChem (<https://pubchem.ncbi.nlm.nih.gov>), yielding a dataset of 352 drugs, 3001 sensitive, and 26696 insensitive data points, which serve as the foundation for our experiments. The test set consists of 10% of all drug-cell line connections. Leveraging the distinct characteristics of the drugs, we calculate drug similarities to enhance the prediction process.

B. Prediction Method

In this section, we incorporate drug response relation data and similarity relation data to construct a comprehensive heterogeneous graph network. To capture the intricate structural information embedded within this network, we utilize the Relational Graph Convolutional Network (RGCN), a powerful framework specifically designed for modeling graph data that contains multiple types of relations. ExplainMIX, our proposed method, leverage the power of RGCN to effectively diffuse the similarity relations and response information, unveiling the intricate relationships between cell lines and drugs.

1) Drugs Similarity Relation: To ensure the availability of chemical structure information of the 352 drugs, we computed a comprehensive set of 1875 molecular descriptors for each of them using the PaDEL software [40]. These descriptors encompass various aspects of the molecular structure, including atom-type electrotopological state descriptors, McGowan volume, molecular linear free energy relation descriptors, and ring counts, among others. To quantify the similarity between drugs, we employed the cosine similarity metric, which calculates the cosine of the angle between two molecular descriptor vectors.

2) Cell Lines Similarity Relation: To obtain a comprehensive view of cell lines, we employed the Similarity Network Fusion (SNF) method [41], which leverages networks of samples to integrate multiple omics data. SNF operates by constructing different omics networks and fusing them to extract useful information. One of the advantages of SNF is the ability to derive meaningful insights even when the number of samples is small compared to the number of features. Additionally, during the procedure, weak similarity edges are eliminated, reducing noise in the data. In our study, we combined six data types for 192 cancer cell lines, including gene expression, somatic mutation, copy number variation, DNA methylation, proteomics and metabolomic data. We normalized the features and calculated pairwise distances using cosine similarity. As depicted in Fig. 1(c), each data type generated distinct patterns of cell line similarity. Next, we computed affinity matrix that represents the neighborhood graph of data points based on the generated distance matrices. We set the number of nearest neighbors (K) to 20, variance for local model (sigma) to 0.8. Finally, we fused the multiple views of the network together to construct an overall status matrix. The fusion process was performed for 50 iterations ($t = 50$). Through this process, we were able to provide a comprehensive view of a cohort of cell lines.

As shown in Fig. 1(d), we acquired four types of relations that can be used as input for the Relational Graph Convolutional Network (RGCN). The RGCN model is specifically designed to handle the complexity of multi-relational data features, making it well-suited for prediction tasks in heterogeneous graphs

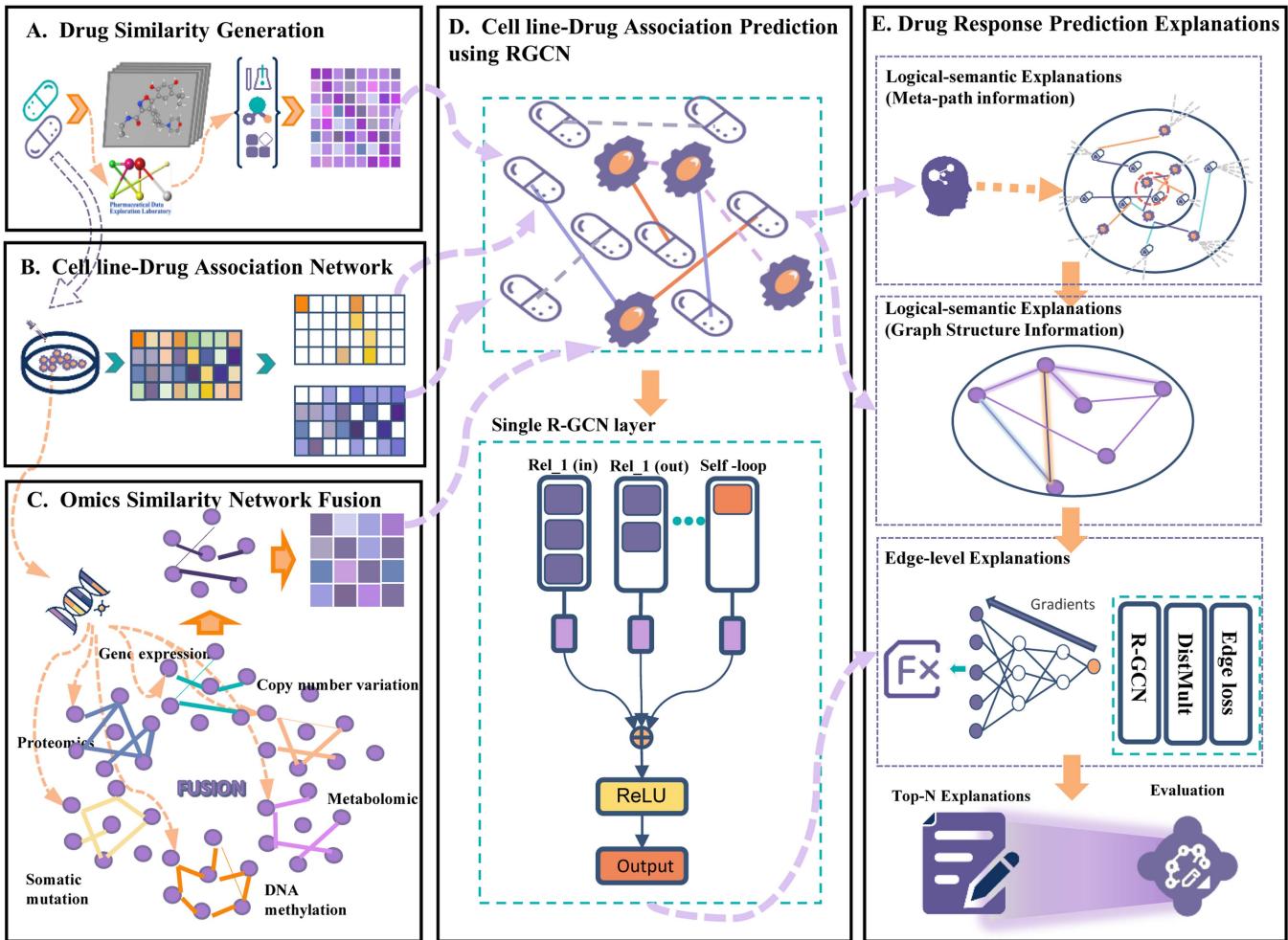


Fig. 1. Workflow of the proposed computational method for explaining drug response prediction. **(a)** Illustrates the process of constructing a drug similarity network based on the 3D structures of drugs. **(b)** Depicts the generation of a correlation network that categorizes cellular responses to drugs into sensitive and resistant categories. **(c)** Demonstrates the creation of a cell line similarity network by integrating cell line omics data. **(d)** Outlines the prediction of drug-response relationships in cell lines. **(e)** Showcases the generation of explanatory insights by leveraging edge prediction results within a directed heterogeneous network structure.

that involve diverse relationships between drugs and cell lines [30]. This propagation model works by iteratively calculating a forward-pass update for each entity v_i considering the relations and corresponding links:

$$h_i^{l+1} = \sigma \left(\sum_{r \in R} \sum_{j \in N_i^r} \frac{1}{c_{i,r}} W_r^{(l)} h_j^{(l)} + W_0^{(l)} h_i^{(l)} \right) \quad (1)$$

where R represents a specific set of relationships, which captures the different types of connections between nodes. N_i^r denotes the set of neighbors of a given node v_i with respect to a particular relationship. $c_{i,r}$ is a constant associated with a certain problem. Similar to the Graph Convolutional Network (GCN) framework, $h_j^{(l)}$ is the hidden state of the node v_i in the neural network's l th layer. The weight matrix W is employed to transform the input features or hidden states. $\sigma(\cdot)$ is the activation function that is applied element-wise to introduce non-linearity to the model.

To mitigate overfitting on rare relationships and handling large model sizes, RGCN implements parameter sharing between weight matrices. This helps limiting the total number of

parameters and preventing overfitting on infrequent relations. In our experimental setup, aimed at learning node features within the drug-cell line heterogeneous network, we utilized a three-layer RGCN architecture. To optimize the model's performance in terms of accuracy, we set the learning rate of 0.001, which yielded the best results among the tested values of {0.1, 0.01, 0.001, 0.0001}. Subsequently, we trained the RGCN for 400 epochs, optimizing the model based on the loss between predicted results and true labels in the training dataset. Lastly, we employed the DistMult factorization as the scoring function to determine the likelihood of edges belonging to specific drug response relations.

IV. EXPLANATION

A. Explanation Method

This work focuses on exploring explanation methods for drug response prediction in heterogeneous information networks, aiming to generate accurate top-N personalized explanations. Heterogeneous graphs offer richer information and semantics

compared to homogeneous graphs, as different node types reside in distinct feature spaces [18]. For drug response, cell line features relate to disease type, age, and gender, while drug features involve molecular structure, pharmacokinetics, and target. The varying importance of relations in cell-drug responses underscores the need to learn the subtle differences in node connections. Treating a heterogeneous network as homogeneous risks misinterpreting weak edges as important while neglecting key explanatory ones.

In this section, we introduce a novel approach for generating logic semantic explanations in heterogeneous graph link predictions. Our model adopts a hierarchical explanation structure, spanning from edge to semantic levels, enabling comprehensive and interpretable explanations for predicted graph links. By integrating logical reasoning and semantic analysis, our approach enhances understanding of underlying relationships, boosting interpretation effectiveness in link predictions.

B. Edge-Level Explanations

Firstly, we introduce an edge-level explanation approach, called ExplaiNE [27], that learns the importance of a link. It quantifies edge influence on the predicted link, identifying the most impactful edges as explanations for the result. The intuitive strategy is to observe the impact of the specific link $e\{a, b\}$, $e \in E$, on the probability of the existence of another link $e\{i, j\}$. If certain edges exhibit a higher impact compared to others in the graph, it indicates a strong relationship between these edges and the predicted result. Thus, these edges can be considered as having explanatory roles, providing valuable insights into the underlying patterns and contributing factors influencing the prediction.

For a given heterogenous drug-cell lines network $G = (V, E)$, where $V = D \cup C$, D and C are the sets of drugs, cell lines respectively, while $E = E_{d,c}^+ \cup E_{d,c}^- \cup E_{dd} \cup E_{cc}$ are the sets of links in G . $E_{d,c}^+$ denotes the sensitive link between cell lines and drugs, while $E_{d,c}^-$ represents the resistant link. E_{dd} and E_{cc} indicate drug and cell line similarity, respectively. Its adjacency matrix is A , and its optimal embedding is X^* . g_{ij} is a differentiable score function that predicts the existence probability of an edge existing between the node pair i and j . α_{kl} is the link between two nodes k, l , serving as an explanation candidate. The edge-level explanation method investigates the infinitesimal change to α_{kl} impact on the probability function g_{ij} , by calculating the gradient of it with respect to the adjacency matrix A . The effect score assigned to α_{kl} is determined using the chain rule:

$$\frac{\partial g_{ij}}{\partial \alpha_{kl}}(A) = \nabla_X g_{ij}(X^*)^T \frac{\partial X^*}{\partial \alpha_{kl}}(A) \quad (2)$$

There are studies and experiments [27] have demonstrated that the most informative explanations for a predicted link $e\{i, j\}$ often involve links $e\{a, b\}$, that are directly connected to node i and j . This suggests that the links adjacent to the predicted nodes have a more direct impact on the prediction outcome, whereas other links may have a secondary or indirect influence. To ensure scalability in our explanation generation process, we

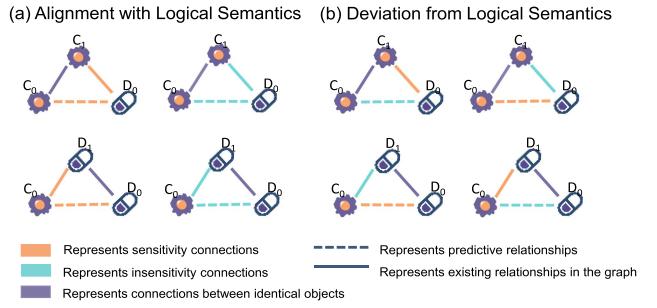


Fig. 2. Visualization of logic semantic rule explanations. This visualization depicts the process of explaining logic semantic rules. **(a)** It provides a representation of the paths that adhere to the logical pattern. **(b)** It provides a representation of the paths that deviate from the logical pattern.

specifically focus on the links directly connected to the nodes within the predicted edge.

C. Logical-Semantic Explanations (Meta-Path Information)

Drug response prediction models that rely solely on gradient-based methods focus on local structures while neglecting the overall model, heterogeneity, and domain knowledge, resulting in incomplete explanations. Attention mechanisms [42] are nowadays ubiquitous in the deep learning, and their potential to offer explanations for model predictions has generated significant interest [43]. In this section, we propose a novel attention mechanism consisting of two modules that incorporate domain knowledge to enhance and complement edge-level explanation results. The first module focuses on semantic relations based on the meta-paths, providing a contextual understanding of node connections leveraging biological prior knowledge. The second module aggregates structural information, offering a broader perspective on the overall graph structure. We incorporate domain expertise into our explanation method, delivering more comprehensive and specialized explanations for drug response predictions. This approach enhances the model's ability to explain predictions by combining both local and global perspectives, driven by relevant biological knowledge.

The meta-path based module provides a structured approach to navigate the network and explore specific paths or patterns of relationships between different types of objects [44]. By leveraging meta-paths, we can extract meaningful insights and address the interdependence of objects in the heterogeneous network. In our case, where the focus is on generating top-N explanations for the relationship between the drug D_0 and the cell line C_0 . To enhance clarity and facilitate understanding, we present the module for semantic-level explanation using a 2-length meta-path as an illustrative example. The predicted edge that requires explanation, and interpretation edges comprises three nodes: C_1, T_2, D_1 . T_2 represents a node distinct from both C_1 and D_1 , and it can either be a drug or a cell line. If the meta-path $C_1 \xrightarrow{R_1} T_2 \xrightarrow{R_2} D_1$ does not exist, it implies that there is no interaction between the nodes along the path. In such cases, we do not consider it for further analysis, and assign a

meta-level attention score $MLap = 0$. And, if the 2-length meta-path exists, we proceed to apply semantic logic judgment for evaluation and interpretation. Fig. 2 visualizes the meta-path and indicates (a) which paths conform to the logical semantic and (b) which path do not.

When the signal between cell lines and drugs in the meta-path aligns with the signal of the predicted edge, we consider interpretation edge to be logical and assign a logical weight of 1, $MLap = 1$. Conversely, if the signal between cell lines and drugs in meta-path contradicts with the signal of the predicted edge, we consider the interpretation edge to deviate from the logical pattern and assign a logical weight of -1, $MLap = -1$.

D. Logical-Semantic Explanations (Graph Structure Information)

Existing explanation methods often emphasize capturing local information. To achieve a more comprehensive explanation, we introduced a structural-level weight that automatically learns the importance of different explanations and combines them for specific task in the second module. The calculation of the structural-level weights involves measuring the interconnectedness between the explained objects and interpretive objects across the entire network [45]. By considering the relatedness between objects throughout the network, we can leverage the complete network topology and incorporate a broader range of information into the model's explanations. In the real-world scenario, two objects are considered highly related if their relationship strength is high and they have fewer relationships with other objects. The relevance measurement between the explained V_0 and interpretive object T_i is determined using a formula:

$$PRel(V_0, T_i) = \frac{N(V_0, T_i) \left(\frac{1}{\deg(V_0)} + \frac{1}{\deg(T_i)} \right)}{\frac{1}{\deg(V_0)} \sum_p N(V_0, V_p) + \frac{1}{\deg(T_i)} \sum_p N(T_i, V_p)} \quad (3)$$

where $V_0 \in C_0$ or D_0 , T_i is a vertex of the explaining edges but not the explained edges. $N(V_0, T_i)$ is the number of k-length paths connecting objects V_0 and T_i , k is a hyperparameter. V_p represents the neighboring nodes connect to V_0 or T_i . $\deg(V_0)$ and $\deg(T_i)$ are node degrees of objects V_0, T_i . The value of the $PRel$ is non-negative and less or equal to one.

In summary, the proposed method in this study utilizes a hierarchical structure to integrate edge-level and logical semantic-level explanations in heterogeneous graphs. By considering both local and global perspectives, we provide a comprehensive understanding of the predicted links. For each predicted link $e\{i, j\}$, we calculate the final score of the explanation by combining the edge-level and logical semantic-level explanations:

$$\text{Score} = \frac{\partial g_{ij}}{\partial \alpha_{kl}} (A) * MLap(V_k, T_l) * PRel(V_k, T_l) \quad (4)$$

where g_{ij} is a function representing the probability between nodes i and j . α_{kl} denotes a candidate explaining edge between nodes k and l , where k is either i or j . V_k and T_l are the

TABLE I
SUMMARY OF THE INFORMATION IN THE TEST DATA OF FOUR EXPERIMENTAL SETTINGS

Experimental setting	Cell line	Drug	Association
Tp	Known	Known	New
Tb	Known	Known	New
Td	Known	New	New
Tc	New	Known	New
Tn	New	New	New

respective nodes. $\frac{\partial g_{ij}}{\partial \alpha_{kl}}$ (A) represents the edge-level influence of the edge α_{kl} on the interpreted link. $MLap(V_k, T_l)$ captures meta-path-level influence, reflecting semantic relations grounded in biological insights, while $PRel(V_k, T_l)$ represents structural-level influence, focusing on graph-based connections. Together, these components enhance the logical-semantic understanding, offering a more nuanced and comprehensive explanation of drug response predictions by integrating both biological knowledge and graph structure. This approach allows for a more nuanced and interpretable assessment of the relationships in the heterogeneous graph.

V. EXPERIMENT SETTINGS AND TASKS

A. Experimental Design for Prediction

In our predictive experiments, we focused on investigating the following research questions:

RQ1: Performance and robustness comparison: Does ExplainMIX outperform state-of-the-art prediction methods across various tasks, and how robust is its performance across different data and experimental conditions?

RQ2: Impact of new omics data: What is the influence of integrating new omics data, such as proteome and metabolome, on the prediction outcomes, and how does this affect the model's generalizability?

To address these questions, we conducted a comprehensive analysis using a carefully designed experimental setup. Our dataset was divided into a training set and a non-overlapping test set. We formulated five different experiment settings (as show in Table I) to evaluate the performance of ExplainMIX:

- 1) Task pairs (Tp): Predicting links in the test set without any special preprocessing.
- 2) Task balance (Tb): Predicting links with the balanced datasets.
- 3) Task drug (Td): Predicting links in the test set while deliberately masking the drug links information contained in the test set from the training set.
- 4) Task cell lines (Tc): Predicting links in the test set while deliberately masking the cell line links information contained in the test set from the training set.
- 5) Task new pairs (Tn): Predicting links in the test set while deliberately masking both the drug and cell line links information contained in the test set from the training set.

By comparing the performance of ExplainMIX with other state-of-the-art methods across all tasks, we assessed its stability and effectiveness. Additionally, to assess the significance of new omics data for prediction, our experiment employed a feature selector that selectively masked the relevant omics data.

B. Experimental Design for Explanation

In our explanation experiments, we aimed to address these key research questions:

RQ3: Performance comparison: Does ExplainMIX surpass state-of-the-art explanation methods based on global ground truth evaluation metrics?

RQ4: Interpretability of main components: What is the impact of the major components of ExplainMIX?

We conducted experiments using an arbitrary selected set of ten prediction results, comprising both insensitive and sensitive cases from the Tp task. For each result, we generated the top five explanations using three different explanation methods, including ExplainNE, GnnExplainer and our ExplainMIX. Through these experiments and analyses, we aimed to evaluate the performance of ExplainMIX in comparison to other explanation methods. Additionally, we performed in-depth analyses on two specific cases to further mine practical information and insights. By delving into these cases, we aimed to uncover valuable information and gain a deeper understanding of the implications of the explanations.

VI. RESULTS

A. Performance in Different Predicting Tasks

In order to evaluate the accuracy of ExplainMIX, we compared it with four state-of-art methods of drug response prediction, namely SRMF [32], MOFGCN [16], GraphDRP [46], and TGSA [36]. The performance of these methods was evaluated in various tasks using the test dataset. In order to assess their prediction accuracy, we computed three evaluation metrics (AUC, AUPR and F1). The obtained results have been summarized and presented in Table II.

Based on the results in Table II, it is evident that ExplainMIX demonstrates superior performance across the four tasks when compared to the other methods. In the Tp task, all methods achieve impressive results. However, it is important to note that the performance may vary when predicting new data. Some methods experienced a significant decline in performance metrics, particularly on the Td and Tn experiment sets.

Nevertheless, overall, our results outperformed other methods for most of the evaluation metrics. As for the more challenging Tn tasks, where the test set represents new drugs and new cell lines, ExplainMIX exhibits significantly superior performance compared to other methods in terms of AUC, AUPR, and F1 metrics, as evident from Table II (4). This indicates that the heterogeneous model utilized in ExplainMIX is more realistic than the isomorphic model in drug response prediction. And it can enhance the results through its internal mechanisms, even for drug or cell line that did not appear in the training set. (**RQ1**)

TABLE II

(1) COMPARISON RESULTS FOR THE Tp TASK

(1) COMPARISON RESULTS FOR THE Tp TASK			
Methods	AUC	AUPR	F1
SRMF	0.8676	0.8776	0.8452
MOFGCN	0.9758	0.9741	0.9712
GraphDRP	0.9517	0.9437	0.9380
TGSA	0.9108	0.9203	0.9021
HGT	0.9544	0.8734	0.9198
REDDA	0.9000	0.4080	0.4480
ExplainMIX-sixo	0.9964	0.9830	0.9571
ExplainMIX-noMO	0.9966	0.9845	0.9665
ExplainMIX-noMu	0.9995	0.9951	0.7319
ExplainMIX-noP&MO	0.9943	0.9879	0.9639
ExplainMIX(noP)	0.9971	0.9864	0.9642

(2) COMPARISON RESULTS FOR THE Tb TASK

Methods	AUC	AUPR	F1
SRMF	0.8534	0.9234	0.8282
MOFGCN	0.6456	0.8052	0.7513
GraphDRP	0.9217	0.9550	0.9171
TGSA	0.9321	0.9635	0.9271
HGT	0.9465	0.8721	0.9198
REDDA	0.8700	0.4000	0.4190
ExplainMIX-sixo	0.9999	0.9999	0.9629
ExplainMIX-noMO	0.9675	0.9845	0.9665
ExplainMIX-noMu	0.9842	0.9998	0.6897
ExplainMIX-noP&MO	0.9950	0.9697	0.9629
ExplainMIX(noP)	0.9999	0.9984	0.9656

(3) COMPARISON RESULTS FOR THE Td TASK

Methods	AUC	AUPR	F1
SRMF	0.7876	0.8044	0.7303
MOFGCN	0.6546	0.4059	0.3188
GraphDRP	0.8444	0.8514	0.8090
TGSA	0.6476	0.6855	0.4565
HGT	0.9432	0.8015	0.8554
REDDA	0.9190	0.4300	0.4190
ExplainMIX-sixo	0.9345	0.7894	0.6805
ExplainMIX-noMO	0.9484	0.8951	0.7390
ExplainMIX-noMu	0.9479	0.9221	0.7321
ExplainMIX-noP&MO	0.9483	0.9222	0.6604
ExplainMIX(noP)	0.9707	0.9132	0.8504

(4) COMPARISON RESULTS FOR THE Tc TASK

Methods	AUC	AUPR	F1
SRMF	0.7740	0.7948	0.7080
MOFGCN	0.8316	0.7017	0.6849
GraphDRP	0.7806	0.7161	0.6746
TGSA	0.8310	0.8490	0.7966
HGT	0.9308	0.7924	0.8572
REDDA	0.7880	0.3030	0.3620
ExplainMIX-sixo	0.9012	0.9100	0.6809
ExplainMIX-noMO	0.9012	0.9099	0.6809
ExplainMIX-noMu	0.9012	0.9099	0.6254
ExplainMIX-noP&MO	0.9012	0.9099	0.6254
ExplainMIX(noP)	0.9012	0.9099	0.6224

(5) COMPARISON RESULTS FOR THE Tn TASK

Methods	AUC	AUPR	F1
SRMF	0.7174	0.7444	0.6061
MOFGCN	0.7066	0.4818	0.3486
GraphDRP	0.8007	0.7586	0.7172
TGSA	0.7439	0.7697	0.6557
HGT	0.9465	0.8721	0.9198
REDDA	0.8340	0.3660	0.3540
ExplainMIX-sixo	0.8482	0.7183	0.4484
ExplainMIX-noMO	0.9047	0.9122	0.6969
ExplainMIX-noMu	0.9001	0.9016	0.6295
ExplainMIX-noP&MO	0.9044	0.9116	0.6295
ExplainMIX(noP)	0.9047	0.9122	0.6713

Note: sixo includes all six omic data types; noMO excludes metabolomic data; noMu excludes mutation data; noP&MO excludes both proteomic and metabolomic data; noP excludes proteomic data.

TABLE III
(1) TOP-10 EXPLANATIONS COMPARATIVE ANALYSIS

(1) TOP-10 EXPLANATIONS COMPARATIVE ANALYSIS					
Methods	GEAavg	GEAmax	TPzero	Recall	RGT
ExplainMIX	0.1055	0.6667	37	0.9233	1.4081
ExplainMIX_nom	0.0738	0.4375	73	0.6859	0.9961
ExplainMIX_nos	0.1046	0.6250	36	0.9191	1.3971
ExplaiNE	0.0692	0.4375	107	0.6408	0.9376
GnnExplainer	0.0223	0.1190	320	0.2857	0.3158

(2) TOP-15 EXPLANATIONS COMPARATIVE ANALYSIS					
Methods	GEAavg	GEAmax	TPzero	Recall	RGT
ExplainMIX	0.1491	0.7778	26	0.9004	1.3716
ExplainMIX_nom	0.1051	0.4800	45	0.6640	0.9691
ExplainMIX_nos	0.1479	0.7222	27	0.8974	1.3604
ExplaiNE	0.1007	0.4688	60	0.6373	0.9326
GnnExplainer	0.0320	0.1304	152	0.2847	0.3152

(3) TOP-20 EXPLANATIONS COMPARATIVE ANALYSIS					
Methods	GEAavg	GEAmax	TPzero	Recall	RGT
ExplainMIX	0.1818	0.7826	22	0.8706	1.3250
ExplainMIX_nom	0.1348	0.5938	35	0.6548	0.9546
ExplainMIX_nos	0.1807	0.7692	22	0.8687	1.3158
ExplaiNE	0.1268	0.5625	42	0.6200	0.9038
GnnExplainer	0.0406	0.1400	91	0.2825	0.3122

Note: ExplainMIX_nom indicates the ablation of the meta-path module in the ExplainMIX method. ExplainMIX_nos indicates the ablation of the graph structure module in the ExplainMIX method.

To investigate the potential benefits of incorporating less-explored omics data (proteomic or metabolomic data) in drug response prediction, we conducted four feature mask experiments in each setting. The results of these experiments presented in Table II . The results indicate that more is not always better; this means the inclusion of more omics data does not always lead to improved results [47]. In the Tp and Tc tasks, the incorporation of additional omics data resulted in enhanced performance, while in the Td and Tn tasks, using only five omics data yielded superior results. Upon analyzing the data, we observed that the inclusion of each additional omics dataset led to increased variability between cell lines and decreased association links in graph. This suggests that there is a possibility that neglecting important edges could have a negative impact on the prediction outcomes. However, it is noteworthy that simply adding more information links does not necessarily lead to better results. Even the fusion of four omics data did not yield satisfactory performance. These findings provide valuable insights for future experiments, highlighting the need to carefully select omics data based on the specific task in drug response prediction studies. (RQ2)

B. Evaluation of Quality for ExplainMIX Explanations

In this section, we focus on applying three explainable methods to interpret the predicted links from test set in the Tp task network. The main objective is to assess the interpretability of ExplainMIX and determine whether it provides sensible explanations. Evaluating the quality of explanation is a challenge task, as there is often a scarcity of datasets with ground-truth explanations. Following a majority of cell drug prediction studies [48], the following hypothesis exists:

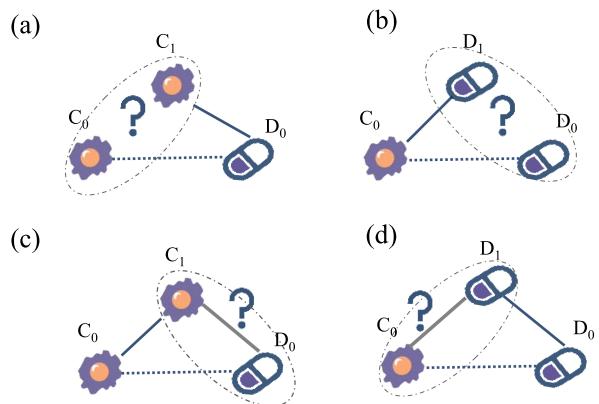


Fig. 3. Evaluation of the reasonableness of explanations. In the diagram above, dashed lines (C_0 to D_0) always represent predicted edges, while solid lines represent explanatory edges generated by the algorithms. (a) If cell line C_1 is sensitive to drug D_0 , the explanatory link (C_1 to D_0) is deemed reasonable upon confirming high similarity between C_0 and C_1 . (b) If cell line C_0 is sensitive to drug D_1 , the explanatory link (C_0 to D_1) is deemed reasonable upon confirming high similarity between D_0 and D_1 . (c) If cell lines C_0 and C_1 are highly similar, the explanatory link (C_1 to D_0) is deemed reasonable upon confirming the sensitivity of C_1 to drug D_0 . (d) If drugs D_0 and D_1 are highly similar, the explanatory link (C_0 to D_1) is deemed reasonable upon confirming the sensitivity of C_0 to drug D_1 .

- 1) If two drugs, D_1 and D_2 , exhibit similar profiles and modes of action, and C_1 is sensitive to drug D_1 , then drug D_2 can be considered as a candidate for cell line C_1 .
- 2) If two cell lines, C_1 and C_2 , share similar profiles, and C_1 is sensitive to drug D_1 , then drug D_1 can be considered as a candidate for cell line C_2 .

This work aligns with these hypotheses and is built upon this premise. We designed an interpretation evaluation framework that allows us to assess the interpretability of the predictions and determine the sensibility and meaningfulness of the explanations provided. To assess the reasonableness of the interpretations, we relied on prior knowledge and domain expertise to ensure their validity. Specifically, the criteria consider four cases based on the type of explanatory links. As shown in the Fig. 3 , the link (C_0 to D_0) always represents the predicted link being explained. (a) If cell line C_1 is sensitive to drug D_0 , the explanatory link (C_1 to D_0) is deemed reasonable upon confirming high similarity between C_0 and C_1 . (b) If cell line C_0 is sensitive to drug D_1 , the explanatory link (C_0 to D_1) is deemed reasonable upon confirming high similarity between D_0 and D_1 . (c) If cell lines C_0 and C_1 are highly similar, the explanatory link (C_1 to D_0) is deemed reasonable upon confirming the sensitivity of C_1 to drug D_0 . (d) If drugs D_0 and D_1 are highly similar, the explanatory link (C_0 to D_1) is deemed reasonable upon confirming the sensitivity of C_0 to drug D_1 . Using these evaluation criteria, we derived the ground truth (the set of edges most suitable for predicting edge interpretations) based on prior knowledge and the overall graph structure. Additionally, we defined several evaluation metrics, including GEAavg, GEAmx, TPzero, Recall, and RGT. GEA is represented in Formula 5, where TP_i denotes the correctly identified edges, defined as the intersection of the interpretive edge set and the ground truth set. FP_i refers to the edges that the interpretation algorithm considers explanatory but are not part of

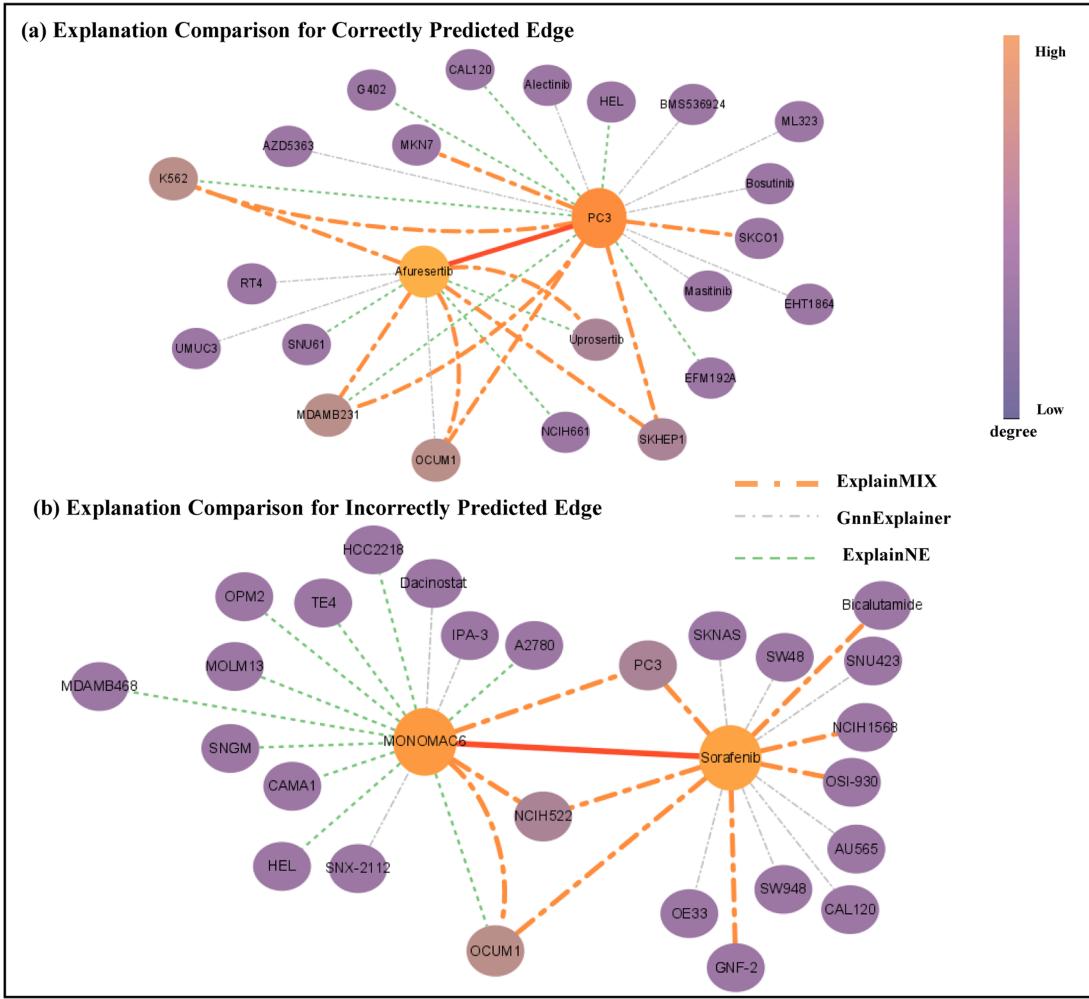


Fig. 4. Case study of explanations for both correct predictions and incorrect predictions. **(a)** Top-ten explanations for the link Afuresertib - PC3 by three methods. **(b)** Top-ten explanations for the link Sorafenib - MONOMAC6 by three methods.

the ground truth, while FN_i indicates the portion of the ground truth not identified by the interpretation algorithm. Formula 6 defines Recall as the probability of correct interpretation. Formula 7 adjusts Recall to account for variations in the number of ground truth edges for each interpretation by standardizing Recall values, where $\bar{\varepsilon}_{gt}$ represents the average ground truth edges, and ε_{gt}^i refers to the ground truth for each edge. This approach helps to assess algorithmic stability.

$$GEA = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FP_i + FN_i} \quad (5)$$

$$Recall = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FN_i} \quad (6)$$

$$RGT = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{(TP_i + FN_i) \bar{\varepsilon}_{gt}^i / \bar{\varepsilon}_{gt}} \quad (7)$$

We conducted a comprehensive analysis of all predicted links to evaluate the validity of the explanations, with comparative

results shown in Table III. Our analysis includes comparisons with advanced methods, ExplaiNE and GnnExplainer, along with ablation experiments on the interpretation algorithm's meta-path and graph structure components. We also assess the top ten, fifteen, and twenty explanations provided by the interpretation algorithm. The table reveals that the GEAvg and GEAmix values achieved by our method are higher than those of other methods, and Recall remains consistently around 0.9, compared to values below 0.7 for other methods. The metric TPzero, which indicates cases where the explanations do not overlap with the ground truth, shows that a lower TPzero corresponds to greater stability; our method demonstrates considerable stability. **(RQ3)**

Additionally, incorporating meta-path (bioinformation) and graph structure information enhances both the accuracy and stability of the explanations generated by our method. The results of the ablation studies on meta-paths and graph structure show a reduction in all evaluation metrics, with the ablation of the graph structure having a more significant impact. This suggests that considering global structural information contributes substantially to enhancing the accuracy of the interpretations. **(RQ4)**

C. Case Studies

The explanation algorithm interprets predicted edges, regardless of whether the predictions are correct or incorrect. In this section, we conduct a detailed analysis of both scenarios. For correct predictions, we investigate the association between the drug Afuresertib and the cell line PC3. As shown in Fig. 4(a), the graph demonstrates significant overlap among the nodes of the explanatory edges, indicating that our algorithm's explanations offer strong associations. Comparing the top 10 explanations produced by the three methods, our method generated 9 explanations matching the ground truth, while ExplainNE provided two, and GNNExplainer provided four. For incorrect predictions, we analyze the association between the drug Sorafenib and the cell line MONOMAC6, where none of the top ten explanations matched the ground truth. In this case, ExplainNE provided one matching explanation and GNNExplainer provided two, underscoring that none of the methods offered satisfactory ground truth explanations for incorrect predictions, which is expected. Furthermore, Fig. 4(b) shows that the edge types in the top ten explanations generated by our method are fairly evenly distributed. In contrast, ExplainNE's explanations mainly focused on cell line similarity edges, while GNNExplainer's explanations centered around drug-cell line edges—a trend consistent across other examples. These findings suggest that the ExplainMIX method offers distinct advantages for explaining heterogeneous graphs.

VII. CONCLUSION

In this paper, we introduce ExplainMIX, a novel interpretable drug response prediction algorithm based on directed graph neural networks and multi-omics fusion, incorporating domain knowledge into the interpretive process. Our proposed method effectively captures complex structures and rich semantics by incorporating prior biological knowledge behind the heterogeneous graph. By integrating edge-level, meta-path and graph structure information, ExplainMIX accurately learns the importance of interpreting edges. Our experimental results, including prediction and explanation tasks, underscore the efficacy of ExplainMIX. Notably, the investigation revealed that augmenting the omics data in cell line drug response prediction does not always lead to improved results, suggesting a need for careful selection and curation of data inputs. By enabling detailed reasoning behind predictions, our method holds great promise for clinical researchers seeking deeper insights into the underlying mechanisms.

Moving forward, our goal is to develop a comprehensive evaluation system for drug response interpretation methods, enabling systematic comparison across different approaches. A key step is establishing a reliable ground truth, which allows accurate assessment of explanation algorithms in terms of both predictive accuracy and interpretative validity. This involves integrating expert annotations, domain knowledge, and empirical data, while leveraging interpretative insights to enhance predictive performance. However, certain limitations must be acknowledged. First, due to data integrity issues, a significant amount of data was excluded during pre-processing. Additionally, our

evaluation relied on heterogeneous and generalized concepts, where expert evaluation could provide more precision. Despite these limitations, we believe our approach provides a strong foundation for advancing drug response prediction and interpretation, with potential for refinement as more data and expert insights become available. In conclusion, our work contributes to directed graph prediction and interpretable machine learning, providing valuable tools for cancer drug response researchers and potential advancements in clinical applications. Continuing our refinement and extension of this approach, we envision a future where interpretable and high-performing methods synergistically contribute to advancements in precision medicine and beyond.

REFERENCES

- [1] V. Prasad, "Perspective: The precision-oncology illusion," *Nature*, vol. 537, no. 7619, 2016, Art. no. S63.
- [2] M. Ali et al., "Global proteomics profiling improves drug sensitivity prediction: Results from a multi-omics, pan-cancer modeling approach," *Bioinformatics*, vol. 34, no. 8, pp. 1353–1362, 2018.
- [3] F. Iorio et al., "A landscape of pharmacogenomic interactions in cancer," *Cell*, vol. 166, no. 3, pp. 740–754, 2016.
- [4] M. Ghandi et al., "Next-generation characterization of the cancer cell line encyclopedia," *Nature*, vol. 569, no. 7757, pp. 503–508, 2019.
- [5] A. Basu et al., "An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules," *Cell*, vol. 154, no. 5, pp. 1151–1161, 2013.
- [6] M. Fazio et al., "Zebrafish patient avatars in cancer biology and precision cancer therapy," *Nature Rev. Cancer*, vol. 20, no. 5, pp. 263–273, 2020.
- [7] J. Tang and T. Attokallio, "Network pharmacology strategies toward multi-target anticancer therapies: From computational models to experimental design principles," *Curr. Pharmaceut. Des.*, vol. 20, no. 1, pp. 23–36, 2014.
- [8] B. Shen et al., "A systematic assessment of deep learning methods for drug response prediction: From in vitro to clinical applications," *Brief. Bioinf.*, vol. 24, no. 1, 2023, Art. no. bbac605.
- [9] J. C. Costello et al., "A community effort to assess and improve drug sensitivity prediction algorithms," *Nature Biotechnol.*, vol. 32, no. 12, pp. 1202–1212, 2014.
- [10] D. Baptista et al., "Deep learning for drug response prediction in cancer," *Brief. Bioinf.*, vol. 22, no. 1, pp. 360–379, 2021.
- [11] S. Chawla et al., "Gene expression based inference of cancer drug sensitivity," *Nature Commun.*, vol. 13, no. 1, 2022, Art. no. 5680.
- [12] H. Sharifi-Noghabi et al., "MOLI: Multi-omics late integration with deep neural networks for drug response prediction," *Bioinformatics*, vol. 35, no. 14, pp. i501–i509, 2019.
- [13] F. Scarselli, M. Gori, A. C. Tsai, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Trans. Neural Netw.*, vol. 20, no. 1, pp. 61–80, Jan. 2009.
- [14] R. Ying et al., "Graph convolutional neural networks for web-scale recommender systems," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2018, pp. 974–983.
- [15] W. Peng, T. Chen, and W. Dai, "Predicting drug response based on multi-omics fusion and graph convolution," *IEEE J. Biomed. Health Inform.*, vol. 26, no. 3, pp. 1384–1393, Mar. 2022.
- [16] W. Peng et al., "Predicting cancer drug response using parallel heterogeneous graph convolutional networks with neighborhood interactions," *Bioinformatics*, vol. 38, no. 19, pp. 4546–4553, 2022.
- [17] H. Zhao et al., "MSDRP: A deep learning model based on multisource data for predicting drug response," *Bioinformatics*, vol. 39, no. 9, 2023, Art. no. btad514.
- [18] X. Wang et al., "Heterogeneous graph attention network," in *Proc. World Wide Web Conf.*, 2019, pp. 2022–2032.
- [19] K. Rohe et al., "Co-clustering directed graphs to discover asymmetries and directional communities," *Proc. Nat. Acad. Sci.*, vol. 113, no. 45, pp. 12679–12684, 2016.
- [20] L. Cui et al., "Deterrent: Knowledge guided graph attention network for detecting healthcare misinformation," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2020, pp. 492–502.

- [21] Y. Zhang, P. Tiňo, A. Leonardis, and K. Tang, "A survey on neural network interpretability," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 5, no. 5, pp. 726–742, Oct. 2021.
- [22] N. Liu et al., "Interpretability in graph neural networks," in *Graph Neural Networks: Foundations, Frontiers, and Applications*. Berlin, Germany: Springer, 2022, pp. 121–147.
- [23] G. Novakovsky et al., "Obtaining genetics insights from deep learning via explainable artificial intelligence," *Nature Rev. Genet.*, vol. 24, no. 2, pp. 125–137, 2023.
- [24] J. Yu et al., "Towards the explanation of graph neural networks in digital pathology with information flows," 2021, *arXiv:2112.09895*.
- [25] Z. Ying et al., "Gnnexplainer: Generating explanations for graph neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, vol. 32, pp. 9240–9251.
- [26] S. Hooker et al., "A benchmark for interpretability methods in deep neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, vol. 32, pp. 1–12.
- [27] B. Kang et al., "Explaine: An approach for explaining network embedding-based link predictions," 2019, *arXiv:1904.12694*.
- [28] X. Wang and H.-W. Shen, "Gnninterpreter: A probabilistic generative model-level explanation for graph neural networks," 2022, *arXiv:2209.07924*.
- [29] B. Aslani and S. Mohebbi, "Ensemble framework for causality learning with heterogeneous directed acyclic graphs through the lens of optimization," *Comput. Operations Res.*, vol. 152, 2023, Art. no. 106148.
- [30] M. Schlichtkrull et al., "Modeling relational data with graph convolutional networks," in *Proc. Semantic Web: 15th Int. Conf.*, Heraklion, Crete, Greece, Jun. 2018, pp. 593–607.
- [31] S.-I. Lee et al., "A machine learning approach to integrate Big Data for precision medicine in acute myeloid leukemia," *Nature Commun.*, vol. 9, no. 1, 2018, Art. no. 42.
- [32] L. Wang et al., "Improved anticancer drug response prediction in cell lines using matrix factorization with similarity regularization," *BMC Cancer*, vol. 17, no. 1, pp. 1–12, 2017.
- [33] M. Ammad-Ud-Din et al., "Drug response prediction by inferring pathway-response associations with kernelized Bayesian matrix factorization," *Bioinformatics*, vol. 32, no. 17, pp. i455–i463, 2016.
- [34] C. Suphavilai et al., "Predicting cancer drug response using a recommender system," *Bioinformatics*, vol. 34, no. 22, pp. 3907–3914, 2018.
- [35] P. Hiort et al., "DrDimont: Explainable drug response prediction from differential analysis of multi-omics networks," *Bioinformatics*, vol. 38, no. Supplement_2, pp. ii113–ii119, 2022.
- [36] Y. Zhu et al., "TGSA: Protein–protein association-based twin graph neural networks for drug response prediction with similarity augmentation," *Bioinformatics*, vol. 38, no. 2, pp. 461–468, 2022.
- [37] A. Holzinger et al., "Towards multi-modal causability with graph neural networks enabling information fusion for explainable AI," *Inf. Fusion*, vol. 71, pp. 28–37, 2021.
- [38] R. Rynazal et al., "Leveraging explainable AI for gut microbiome-based colorectal cancer classification," *Genome Biol.*, vol. 24, no. 1, pp. 1–13, 2023.
- [39] D. Luo et al., "Parameterized explainer for graph neural network," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, vol. 33, pp. 19620–19631.
- [40] C. W. Yap, "PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints," *J. Comput. Chem.*, vol. 32, no. 7, pp. 1466–1474, 2011.
- [41] B. Wang et al., "Similarity network fusion for aggregating data types on a genomic scale," *Nature Methods*, vol. 11, no. 3, pp. 333–337, 2014.
- [42] D. Bahdanau et al., "Neural machine translation by jointly learning to align and translate," 2014, *arXiv:1409.0473*.
- [43] S. Wiegreffe and Y. Pinter, "Attention is not not explanation," 2019, *arXiv:1908.04626*.
- [44] Y. Sun and J. Han, "Mining heterogeneous information networks: A structural analysis approach," *AcM Sigkdd Explorations Newslett.*, vol. 14, no. 2, pp. 20–28, 2013.
- [45] M. Gupta et al., "DPRel: A meta-path based relevance measure for mining heterogeneous networks," *Inf. Syst. Front.*, vol. 21, pp. 979–995, 2019.
- [46] T. Nguyen, G. T. T. Nguyen, T. Nguyen, and D.-H. Le, "Graph convolutional networks for drug response prediction," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 19, no. 1, pp. 146–154, Jan./Feb. 2022.
- [47] D. Chicco et al., "Ten quick tips for avoiding pitfalls in multi-omics data integration analyses," *PLoS Comput. Biol.*, vol. 19, no. 7, 2023, Art. no. e1011224.
- [48] T. N. Jarada et al., "A review of computational drug repositioning: Strategies, approaches, opportunities, challenges, and directions," *J. Cheminformatics*, vol. 12, no. 1, pp. 1–23, 2020.