Interactive Visual Explanations for Deep Drug Repurposing

Qianwen Wang [*1] **Kexin Huang** [*1] **Payal Chandak** [*2] **Nils Gehlenborg** [1] **Marinka Zitnik** [1]

Abstract

Faced with skyrocketing costs for developing new drugs from scratch, repurposing existing drugs for new uses is an enticing alternative that considerably reduces safety risks and development costs. However, successful drug repurposing has been mainly based on serendipitous discoveries. Here, we present a tool that combines a graph transformer network with interactive visual explanations to assist scientists in generating, exploring, and understanding drug repurposing predictions. Leveraging semantic attention in our graph transformer network, our tool introduces a novel way to visualize meta path explanations that provide biomedical context for interpretation. Our results show that the tool generates accurate drug predictions and provides interpretable predictions.

1. Introduction

Developing a new drug to treat a disease, moving the drug forward through clinical trials, and obtaining approval for it is a long, expensive process with a high risk of failure. It currently takes 13-15 years and US $2B to $3B on average (Pushpakom et al., 2019) to get a novel drug to the market, and the costs are going up (Nosengo, 2016). Instead of developing a new drug from scratch, drug repurposing aims to identify therapeutic opportunities of existing drugs that already passed clinical testing, therefore significantly reducing the safety risks and the development cost (Gysi et al., 2021). Because most repurposed drugs have already passed the early phases of development and clinical testing (Ashburn & Thor, 2004), they can potentially get to market in less than half the time and at one-quarter of the cost needed to develop a new drug from scratch (Nosengo, 2016). However, despite considerable advances (Frantzi et al., 2020), drug repurposing remains an open problem driven by serendipitous discoveries.

[*]Equal contribution [1]Harvard University, Cambridge, USA [2]Harvard-MIT Health Sciences and Technology, Cambridge, USA. Correspondence to: Nils Gehlenborg <nils@hms.harvard@edu>, Marinka Zitnik <marinka@hms.harvard.edu>.

Workshop on Interpretable ML in Healthcare at International Conference on Machine Learning (ICML).

Graph neural networks (GNNs) have emerged as a promising approach for drug development and demonstrated the ability to identify promising therapeutic opportunities with unprecedented speed, scale, and accuracy (Li et al., 2021). However, critical challenges remain. (1) The underlying GNN models remain elusive to interpretation by domain experts. Typical GNN models are built by specialists, and endpoint users (e.g., wet lab biologists and physicians) cannot directly engage with the model. How can we enable users to provide feedback on model development and build trust in ML predictions by asking what-if questions and receiving accurate predictions that can be interpreted meaningfully? (2) Biomedical data involve rich multimodal and heterogeneous interactions of many different types, including experimental readouts, curated annotations, and metadata—no single data modality can capture all the factors necessary to identify a successful drug treatment (Zitnik et al., 2019).

Explainability has an important role in addressing the above challenges. However, existing GNN explainers (e.g., Ying et al.; Schlichtkrull et al.; Huang et al.; Vu & Thai), while remarkably powerful, produce explanations that are either too vague or that do not lend themselves to testable biomedical hypotheses. For example, GNNExplainer (Ying et al., 2019) identifies a locally informative subgraph and a subset of node feature dimensions to explain a given GNN prediction. But when applying GNNExplainer to drug repurposing, it remains unclear how to connect these explanations to disease treatment mechanisms. To successfully apply GNNs for drug repurposing, it is crucial to provide explanations that can be easily interpreted by scientists in the context of drug development.

Here, we provide interactive visual explanations that reflect biological mechanisms to assist GNN-based drug repurposing. To this end, we construct a drug repurposing knowledge graph consisting of a variety of molecular interaction data, gene expression data, clinical trials, and drug treatments. We then develop a heterogeneous graph attention message passing GNN architecture and use it with the knowledge graph to make drug repurposing predictions (i.e., indications, contra-indications, or off-label use).

Our key contribution is a novel interactive visualization that provides explanations that can reflect biological mechanisms and be interpreted meaningfully in the context of drug

repurposing. Preliminary results show the approach can accurately predict treatments for molecularly uncharacterized diseases and provide visual explanations that faithfully capture known biological mechanisms. A web demo is available at **http://drugexplorer.gehlenborglab.org**.

2. Desiderata: Explainable Deep Repurposing

To guarantee safe and effective applications, predictions should enable meaningful interpretation by end-users in the intended use cases (i.e., a user-centric view) and be actionable from a biomedical perspective (i.e., a domain-specific view).

A **user-centric view** requires careful consideration about how the explanations are presented to and utilized by the end users. The machine learning expertise of the end users will significantly influence the manner in which the explanations are interpreted. The interpretation should lead to clear actionable insights rather than a hunch.

A **domain-specific view** requires explanations that reflect biological mechanisms of drug action (e.g., Van Maanen et al.; Gonçalves et al.; Lin et al.). These domain-specific explanations can work as an effective way to evaluate the trustworthiness of a prediction, especially when the ground truth of the prediction is unavailable or when the assessment of the prediction requires downstream studies (e.g., clinical studies). Meanwhile, considering our incomplete understanding of molecular pathology and drug actions, explainable deep learning bears the potential to augment our understanding of drugs and maximize the yield of follow-up studies.

3. Heterogeneous Graph Attention Approach

3.1. Molecular Data and Knowledge Graph

We processed a variety of molecular interaction data, including the human interactome assembled from 21 public databases of protein-protein interactions, gene expression data, clinical trials, and information on drug indications, contra-indications, and off-label use across the entire range of 22K+ human diseases and 7K+ drugs. We integrated these data into a knowledge graph of heterogeneous entities and their relations, as shown in Table 1. We have 10 types of nodes and 32 types of relations. Note that a pair of node types can have multiple relation types. For example, the relation between a *drug* node and a *protein* node can be either *carrier*, *enzyme*, *target*, or *transporter*.

Node type	Count	Node type	Count
gene/protein	27,576	biological process	28,110
disease	20,761	phenotype	13,631
anatomy	8,601	molecular function	10,207
drug	7,420	cellular component	3,887
pathway	2,427	exposure	1,336
total	125,956		

Table 1. Statistics about the knowledge graph.

3.2. Drug-Disease Link Prediction

We treat drug repurposing as a link prediction task defined on a knowledge graph. Given drug i and disease j, we aim to predict the type of a relation $r(e_{ij}) \in R_{\text{drug,disease}}$, where $R_{\text{drug,disease}} = \{\text{indication, contra-indication, off-label use}\}$.

In order to infer the missing links (i.e., unknown drug-disease relations) from the heterogeneous knowledge graph, we developed a heterogeneous version of the graph attention neural network (Veličković et al., 2018). This model overcomes the limitation of most existing GNN models that can only learn from fixed and homogeneous graphs.

Meanwhile, this model provides a semantic attention mechanism that enables a node to learn the importance of different neighbors based on the type of their relations. More specifically, the model learns relation-specified weights for the 32 relation types in the knowledge graph. For a node i at layer l, its embedding $\mathbf{h}_i^{(l)}$ is calculated based on embedding from the previous layer using two relation-specified weight matrices $\mathbf{W}_{r,M}^{(l)}$ and $\mathbf{W}_{r,A}^{(l)}$:

$$e_{i,j}^{(l)} = \text{LeakyRelu}(\mathbf{W}_{r,A}^{(l)}((\mathbf{W}_{r,M}^{(l)}\mathbf{h}_i^{(l-1)}) \| (\mathbf{W}_{r,M}^{(l)}\mathbf{h}_j^{(l-1)})))$$

$$\mathbf{h}_i^{(l)} = \sum_{j \in \mathcal{N}_i} \frac{\exp(e_{i,j})}{\sum_{k \in \mathcal{N}_i} \exp(e_{i,k})} \mathbf{W}_{r,M}^{(l)} \mathbf{h}_j^{(l-1)}$$

Probability of a relation $r \in R_{\text{drug,disease}}$ between a drug i and a disease j is then calculated as: $p_{i,j,r} = 1/(1 + \exp(-\text{sum}(\mathbf{h}_i * \mathbf{w}_r * \mathbf{h}_j)))$. Once the model is trained, $e_{i,j}$ indicates the relevance of drug i for disease j.

4. Interactive Visual Explanations

4.1. Visual Explanations from Two Perspectives

We developed an interaction visualization tool to help end users interact with the GNN model and understand the predicted drug indications (Figure 1). Users can search and select a disease of interest in the control panel (Figure 1a). The tool will then predict possible drug indications for the selected disease and visualize the explanations of these predictions (Figure 1b-d). For each predicted drug indication, our tool constructs and visualizes two kinds of explanations: a model-level explanation and a human-level explanation.

Model-level explanations reveal how the model makes a certain prediction. In GNN-based drug repurposing, the model-level explanation includes node embedding, relation-based edge attentions, and subgraphs that are important for

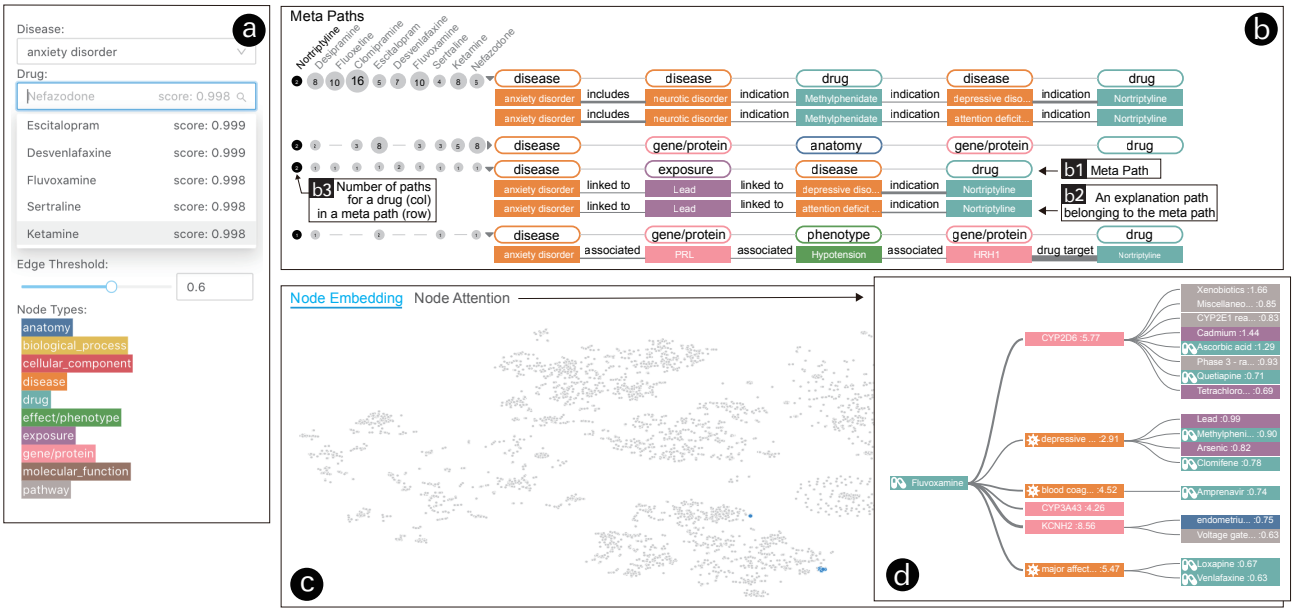

Figure 1. Interface of our system for interactive visual explanation of drug repurposing predictions.

the prediction.

In contrast, human-level explanations mimic how an expert in the application domain would reason about a prediction. In drug repurposing, the human-level explanation can be presented as semantic paths in the knowledge graph that reflect biomedical mechanisms. For example, Ritonavir is a drug originally developed for treatment of AIDS. The indication of Ritonavir for the disease ALS can be explained by a Ritonavir-NR1|2-ALS path: Ritonavir targets the gene NR1|2 which is implied in ALS. This drug-gene-disease meta-path provides a domain-meaningful and human-readable explanation, assisting researchers in assessing the predicted drug candidates and conducting necessary follow-up clinical studies. However, there can be more than ten thousand paths between a drug and a disease node in the knowledge graph. The relation-specified attention can help users effectively identify the important and semantic paths.

4.2. Model-Level Explanations

A model-level explanation is represented using two views: a *Node Embedding* view (Figure 1c) and a *Node Attention* view (Figure 1d). In the *Node Embedding* view, we use t-SNE (Van der Maaten & Hinton, 2008) to present the learned embedding of all drug nodes in the knowledge graph and highlight the predicted drugs for the selected disease. This visualization can help users obtain an overview of the predicted drugs, e.g., the semantic similarity between the predicted drugs, the diversity of the drug prediction. In the *Node Attention* view, we use a tree visualization to present the edge attention and the subgraph structure in message passing. The node color indicates the type of nodes, the edge thickness indicates the attention weights, and the tree structure reflects the message passing. Considering that a node can have thousands of neighbors and most neighbors make little contribution to the prediction, we only show the top-k neighbors with the highest attention weights, where $k = 20/(2^n)$ for n-hop neighbors.

4.3. Human-Level Explanations

For human-level explanations (meta paths in this case), we design a novel yet intuitive visualization for end users.

We present meta-path explanations for top-k drug predictions for a selected disease, as shown in Figure 1b. The rows with round nodes represent meta paths, while the rows with rectangle nodes represent individual explanation paths that belong to a certain meta path. The matrix on the left side summarizes the number of explanation paths for each drug (column) that belong to different meta paths (rows). Users can select a drug and expand all the explanation paths for the selected drug. This visualization effectively summarizes the predicted drugs in terms of a meta path and facilitates the comparison between predicted drugs, especially when their predictions scores are similar.

5. Results

5.1. Performance of Drug Repurposing Predictor

We report here recall scores obtained by the heterogeneous graph attention model on a randomly split dataset of drug-disease relationships. Note that our results focus on demonstrating visual explanations; benchmarking results of the

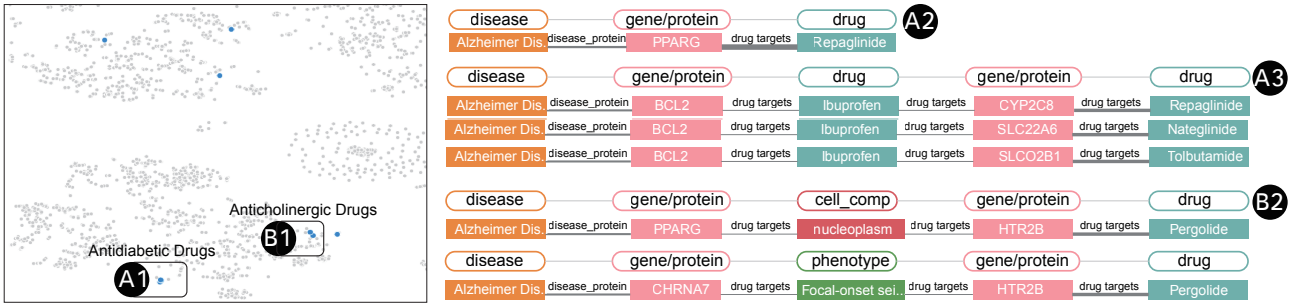

Figure 2. Exploring drug repurposing predictions for the treatment of Alzheimer's Disease.

predictor are not shown due to page limit. We consider disease-centric evaluation: for each disease in the test set, we pair it with the full list of drugs and remove those from the training set. This list is fed to the model to output a score for each drug. We use the scores to produce a ranked list of drugs for each disease. Since only a small number of drugs are available for each disease, we calculate model performance using $recall@K\%$, i.e., how many drugs from the test set (i.e., hits) are found in the top $K\%$ of the ranked list. We aggregate the $recall$ across diseases. Table 2 show the model can accurately predict drug-disease relationships ranking 53.9% of hits in top 1%, 79.5% of hits in top 5%, and 88.9% of hits in top 10%.

5.2. Case Studies for Visual Explanations

We conduct a case study into treatments for Alzheimer's Disease (AD). The GNN model was trained on the full knowledge graph and used to make predictions for drugs, which were not included in the knowledge graph. We selected AD in the visualization tool and explored predicted drugs and their explanations.

The tool automatically produced predictions and updated visualizations for AD (Figure 2). Predicted drugs were scattered in the *Node Embedding* view, indicating that the GNN model produced predictions for a diverse set of drugs.

We first examined the largest cluster of drugs (Figure 2.A1). This cluster included drugs such as *Glyburide*, *Repaglinide*, *Tolbutamide*, and *Metformin*, used to treat Type 2 diabetes (T2D). Drugs found in the cluster were consistent with current scientific understanding of the connections between cognitive impairment and T2D (Sastre et al., 2017). Previous studies have found that the use of antidiabetic treatments among individuals with T2D could mitigate risk for dementia (Akimoto et al., 2020).

$recall@$	indication	contra-indication	off-label use
1%	0.539±0.463	0.339±0.379	0.393±0.379
5%	0.795±0.378	0.615±0.407	0.618±0.476
10%	0.888±0.297	0.759±0.359	0.770±0.412

Table 2. Disease-centric evaluation. Higher values are better.

We then examined explanations for predicted antidiabetic drugs in the *Meta Paths* view. To this end, we first selected *Repaglinide* in the *Meta Paths* view to show a detailed explanation. The shortest meta path is *Disease-Gene/Protein-Drug*. The explanation path below that meta path (Figure 2.A2) showed that *Repaglinide* targets protein *PPARG*, which, in turn, is associated with AD. Based *Disease-Gene /Protein-Drug-Gene/Protein-Disease* meta path (A3), we see that drug *Repaglinide* was predicted partly because it has the same target protein as *Ibuprofen*. *Ibuprofen* targets proteins that are associated with AD and can delay some forms of AD pathology (Lim et al., 2000). Similar instances of meta paths existed in explanations of other antidiabetic drugs, including *Nateglinide* and *Tolbutamide*.

Another cluster (Figure 2.B1) in the *Node Embedding* view comprised of anticholinergic drugs, including *Levodopa*, *Pergolide*, and *Orphenadrine*, which are used to manage Parkinson's disease. The *Meta Paths* view showed explanations for those predictions. By examining explanation paths, we found that a target protein of *Pergolide* interacts with multiple AD-associated proteins through shared cellular phenotypes (B2), an observation consistent with the reported associations between AD and anti-Parkinson's agents (Ono et al., 2006). While some studies (Joung et al., 2019) reported the contraindication of these drugs, the contraindication still reflected the GNN's ability to identify associations unknown in the training graph. This example also highlighted the utility of visual explanations to perform error analysis and identify possible inaccurate predictions.

6. Conclusion and Future Directions

We developed a tool that provides interactive visual explanations for GNN-based drug repurposing predictions. Results show that this tool can effectively identify promising repurposing opportunities and explain predicted drug uses, highlighting the benefits of user-centric and domain-specific visual explanations. Moving forward, we plan to carry out user studies with scientists to systematically evaluate the explainability and usability of the visualization tool.

References

Akimoto, H., Negishi, A., Oshima, S., Wakiyama, H., Okita, M., Horii, N., Inoue, N., Ohshima, S., and Kobayashi, D. Antidiabetic drugs for the risk of alzheimer disease in patients with type 2 dm using faers. *American Journal of Alzheimer's Disease & Other Dementias®*, 35: 1533317519899546, 2020.

Ashburn, T. T. and Thor, K. B. Drug repositioning: identifying and developing new uses for existing drugs. *Nature Reviews Drug Discovery*, 3(8):673–683, 2004.

Frantzi, M., Latosinska, A., Mokou, M., Mischak, H., and Vlahou, A. Drug repurposing in oncology. *The Lancet Oncology*, 21(12):e543, 2020.

Gonçalves, E., Segura-Cabrera, A., Pacini, C., Picco, G., Behan, F. M., Jaaks, P., Coker, E. A., van der Meer, D., Barthorpe, A., Lightfoot, H., et al. Drug mechanism-of-action discovery through the integration of pharmacological and CRISPR screens. *Molecular Systems Biology*, 16 (7):e9405, 2020.

Gysi, D. M., Do Valle, Í., Zitnik, M., Ameli, A., Gan, X., Varol, O., Ghiassian, S. D., Patten, J., Davey, R. A., Loscalzo, J., et al. Network medicine framework for identifying drug-repurposing opportunities for COVID-19. *Proceedings of the National Academy of Sciences*, 118(19), 2021.

Huang, Q., Yamada, M., Tian, Y., Singh, D., Yin, D., and Chang, Y. Graphlime: Local interpretable model explanations for graph neural networks. *arXiv:2001.06216*, 2020.

Joung, K.-i., Kim, S., Cho, Y. H., and Cho, S.-i. Association of anticholinergic use with incidence of alzheimer's disease: population-based cohort study. *Scientific reports*, 9(1):1–10, 2019.

Li, M. M., Huang, K., and Zitnik, M. Representation learning for networks in biology and medicine: Advancements, challenges, and opportunities. *arXiv:2104.04883*, 2021.

Lim, G. P., Yang, F., Chu, T., Chen, P., Beech, W., Teter, B., Tran, T., Ubeda, O., Ashe, K. H., Frautschy, S., et al. Ibuprofen suppresses plaque pathology and inflammation in a mouse model for alzheimer's disease. *Journal of Neuroscience*, 20(15):5709–5714, 2000.

Lin, A., Giuliano, C. J., Palladino, A., John, K. M., Abramowicz, C., Yuan, M. L., Sausville, E. L., Lukow, D. A., Liu, L., Chait, A. R., et al. Off-target toxicity is a common mechanism of action of cancer drugs undergoing clinical trials. *Science Translational Medicine*, 11 (509), 2019.

Nosengo, N. New tricks for old drugs. *Nature*, 534(7607): 314–317, 2016.

Ono, K., Hasegawa, K., Naiki, H., and Yamada, M. Anti-parkinsonian agents have anti-amyloidogenic activity for alzheimer's β-amyloid fibrils in vitro. *Neurochemistry International*, 48(4):275–285, 2006.

Pushpakom, S., Iorio, F., Eyers, P. A., Escott, K. J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C., et al. Drug repurposing: progress, challenges and recommendations. *Nature Reviews Drug Discovery*, 18(1):41–58, 2019.

Sastre, A. A., Vernooij, R. W., Harmand, M. G.-C., and Martínez, G. Effect of the treatment of type 2 diabetes mellitus on the development of cognitive impairment and dementia. *Cochrane Database of Systematic Reviews*, (6), 2017.

Schlichtkrull, M. S., De Cao, N., and Titov, I. Interpreting graph neural networks for nlp with differentiable edge masking. *International Conference on Learning Representations*, 2020.

Van der Maaten, L. and Hinton, G. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(11), 2008.

Van Maanen, J., Retel, J., De Vries, J., and Pinedo, H. Mechanism of action of antitumor drug etoposide: a review. *JNCI: Journal of the National Cancer Institute*, 80(19): 1526–1533, 1988.

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., and Bengio, Y. Graph attention networks. *ICLR*, 2018.

Vu, M. N. and Thai, M. T. PGM-explainer: probabilistic graphical model explanations for graph neural networks. *Advances in Neural Information Processing Systems*, 2020.

Ying, R., Bourgeois, D., You, J., Zitnik, M., and Leskovec, J. GNNExplainer: Generating explanations for graph neural networks. *Advances in Neural Information Processing Systems*, 32:9240, 2019.

Zitnik, M., Nguyen, F., Wang, B., Leskovec, J., Goldenberg, A., and Hoffman, M. M. Machine learning for integrating data in biology and medicine: Principles, practice, and opportunities. *Information Fusion*, 50:71–91, 2019.