

# XplainScreen: Unveiling the Black Box of Graph Neural Network Drug Screening Models with a Unified XAI Framework


GeonHee Ahn*
Ewha Womans University
Seoul, South Korea
justamoment@ewhain.net

Md Mahim Anjum Haque*
Luminary-AI
Blacksburg, VA, USA
mahim@vt.edu

Subhashis Hazarika
Fujitsu Research of America
Santa Clara, CA, USA
shazarika@fujitsu.com

Soo Kyung Kim†
Ewha Womans University
Seoul, South Korea
sookim@ewha.ac.kr



## Abstract

Despite the powerful capabilities of GNN-based drug screening model in predicting target drug properties, the black-box nature of these models poses a challenge for practical application, particularly in a field as critical as drug development where understanding and trust in AI-driven decisions are important. To address the interpretability issues associated with GNN-based virtual drug screening, we introduce XplainScreen: a unified explanation framework designed to evaluate various explanation methods for GNN-based models. XplainScreen offers a user-friendly, web-based interactive platform that allows for the selection of specific GNN-based drug screening models and multiple cutting-edge explainable AI methods. It supports both qualitative assessments (through visualization and generative text descriptions) and quantitative evaluations of these methods, utilizing drug molecules in SMILES format. This demonstration showcases the utility of XplainScreen through a user study with pharmacological researchers focused on virtual screening tasks based on toxicity, highlighting the framework's potential to enhance the integrity and trustworthiness of AI-driven virtual drug screening. A video demo of XplainScreen is available at https://youtu.be/Q4yobrTLKec, and the source code can be accessed at https://github.com/GeonHeeAhn/XplainScreen.



## CCS Concepts

• **Human-centered computing** → **Visualization toolkits**; **Scientific visualization**; **Visualization design and evaluation methods**; *User studies*; **Web-based interaction**; *Displays and imagers*; **Text input**; **User interface toolkits**; **Visual analytics**; **Information visualization**; *Accessibility systems and tools*; • **Computing methodologies** → **Information extraction**; **Natural language generation**; • **Applied computing** → **Chemistry**; • **General and reference** → **Reliability**; **Metrics**; **Validation**; **Performance**.

## Keywords

Explainable AI, Graph Neural Networks, Drug Screening, Scientific Machine Learning


*Both authors contributed equally to this research.
†Corresponding author.




This work is licensed under a Creative Commons Attribution International 4.0 License.



*CIKM '24, October 21–25, 2024, Boise, ID, USA*


© 2024 Copyright held by the owner/author(s).


ACM ISBN 979-8-4007-0436-9/24/10
https://doi.org/10.1145/3627673.3679236

**ACM Reference Format:**
GeonHee Ahn, Md Mahim Anjum Haque, Subhashis Hazarika, and Soo Kyung Kim. 2024. XplainScreen: Unveiling the Black Box of Graph Neural Network Drug Screening Models with a Unified XAI Framework. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM '24), October 21–25, 2024, Boise, ID, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3627673.3679236


## 1 Introduction

Drug discovery and development is an intensive process that typically spans 12 to 18 years and costs between 1 to 3 billion, largely due to the high rate of failure in early stages [11, 21, 23]. Most pharmaceutical companies are now incorporating artificial intelligence (AI) to enhance the efficiency and success rates of their drug development pipelines [19]. Deep learning (DL) models have been particularly successful in quickly identifying promising drug candidates [16]. However, these models are black boxes, making it difficult to understand how they make their predictions. To build trust and ensure transparency in AI-driven drug discovery, it's crucial to clearly explain how these models arrive at their decisions in terms that align with scientific understanding [13]. Unfortunately, current explainable AI (XAI) methods, which help to interpret these decisions, tend to lack robustness and can be sensitive to minor changes in the model or data [10]. To address this issue, we present ***XplainScreen***, *a visual framework that allows for the effective examination of drug screening model predictions using various XAI methods*, thus providing more reliable explanations for AI-based drug screening models.

Graph Neural Networks (GNNs) are extensively employed in AI-driven drug screening due to their capacity to represent molecular graphs with high accuracy and fidelity [5, 29, 31]. GNNs learn a continuous representation directly from the molecular structure of chemical compounds (represented as graphs) to predict their properties. The challenge of interpreting those GNN-based methods is the quality of the explanations can vary depending on the network architecture and complexity of the XAI methods. They are generally not very robust and stable to small changes to model architectures and input perturbations [4, 8, 17]. Moreover, the programming and data preprocessing overhead required to deploy AI models present significant challenges for scientists. Therefore, integrating a user-friendly framework that evaluates and compares multiple XAI methods for GNNs into the drug screening workflow is essential. Despite the clear need, current research lacks such integrated, user-friendly XAI frameworks specifically designed for drug screening.

Addressing this gap, we present **XplainScreen**, a web-based user-friendly explain framework for GNN-based drug screening

that consolidates various XAI methods into a single platform. The main contributions of XplainScreen are as follows: **(1)** ***Unified explanation framework:*** XplainScreen allows the end-users to efficiently apply and compare *multiple XAI methods*, to qualitatively and quantitatively *investigate*, *compare*, and *interpret* the underlying AI model predictions. XplainScreen enables users even without domain knowledge to understand AI models by providing four types of explanation, including visualization, quantitative evaluation, comparision analysis and textual interpretation. **(2)** ***Efficient Visualization Scheme with Unified Data Structure:*** To back this front-end system, we propose a unified data structure for the heterogeneous graph-level explanation artifacts (nodes, edges, graph walks of different lengths), facilitating analysis of explanation results from different methods and models in the same platform. **(3)** ***User-friendly framework:*** XplainScreen offers the user-friendly web-based platform that end-users can easily load different types of pre-trained GNN-based drug screening models and apply XAI methods in real time without programming and data preprocessing overheads. **(4)** ***User-study:*** To demonstrate the utility of XplainScreen in the drug discovery workflow, we conducted a systematic user study interviewing multiple pharmacological scientists.

## 2 Related Work

The rise of AI models in computational drug design and development necessitates the development of explanation strategies for these models [12, 15, 20]. Several XAI methods have been recently proposed for deep learning architectures like GNNs which are widely used in the drug modeling domain [30, 34]. Jimenez–Luna et al [13] elaborates on the state-of-the-art, and future direction of XAI in drug discovery. However, the explanation of these XAI methods are not always robust. They are highly susceptible to the model architecture, minor input perturbations and noise. Therefore while trying to understand the prediction results of a trained model it is important to compare explanation results across different XAI methods. To the best of our knowledge, such a framework to explain and analyze the model predictions from different GNN-based drug screening models using diverse pool of XAI methods is still missing. Rao et al. [25] provides a quantitative benchmark datasets to evaluate XAI methods on specific molecular properties of drug candidates. Yuan et al. [18, 34] offers a generic XAI evaluation framework for GNNs in general domain. Our proposed framework is the first attempt to provide an unified explanation framework for both qualitative (via. visualization and generative text description) and quantitative evaluation of explanation results for pre-trained GNN-based drug screening models.

## 3 Solution Architecture

XplainScreen offers a feature to explain GNN-based drug candidate screening models pretrained with various datasets using different XAI methods. As illustrated in Fig 1, the workflow of XplainScreen is divided into three stages: Input Selection Stage, Explanation Stage, and Presentation Stage.

(1) **Input Selection Stage:** The end-user inputs the target drug molecule in SMILES string format [6] and selects four parameters: *1-the target dataset, 2-the desired XAI method, 3-the pre-trained GNN-based drug screening model*, and *4-the visualization scheme.*

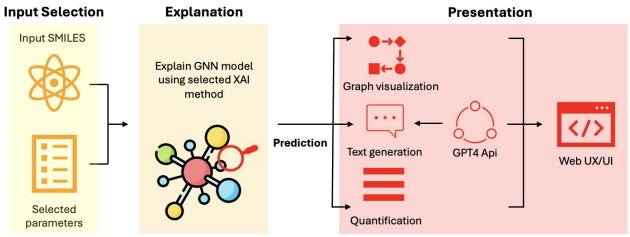



**Figure 1: Overview of XplainScreen workflow**

(2) **Explanation Stage:** The chosen XAI method is applied to the selected GNN model, which has been pre-trained on the specified dataset.

(3) **Presentation Stage:** The results from the Explanation Stage are visualized using the chosen scheme, displaying the molecular graph where each color represents the contribution of corresponding sub-graphs to the GNN's prediction. These colored molecular graphs are then analyzed using GPT-4 Vision, which interprets the scientific significance of the visualization based on its knowledge base.

## 4 Input Components

For XplainScreen prototype, here we utilized publicly available drug property datasets alongside existing pre-trained GNN models. The platform integrates four different state-of-the-art XAI methods, showing visual and textual description of explanation results for the selected XAI method. In this section, we will describe the key input components selected during the input selection stage, including the datasets, pre-trained GNN models, XAI methods, and the visualization scheme used in our prototype.

### 4.1 Dataset

In our prototype framework, we focus on the downstream task of screening drug candidates based on low toxicity, a critical criterion for defining drug viability [1]. We utilized two drug toxicity datasets: ***Tox21*** [2] and ***ClinTox*** [22]. Our focus is on models predicting toxicity for the AhR (aryl hydrocarbon receptor), a protein crucial for mediating toxicity.

### 4.2 GNN-based Drug Screening Method

We used two widely used GNN models pre-trained on Tox21 and ClinTox datasets as GNN-based drug screening models [18]:

- **Graph Convolution Networks (GCN):** GCN integrates structural information by utilizing a localized first-order approximation of spectral graph convolutions. It aggregates and processes neighboring node features through a neural network for effective training [28].
- **Graph Isomorphism Networks (GIN):** GIN implements a robust neighbor aggregation strategy. This design enhances its discriminative capabilities, making it superior to GCN and other graph neural network variants in detecting isomorphisms [32].

### 4.3 XAI Methods for Graph Neural Networks

Although post-hoc XAI methods for GNNs vary in their underlying approaches [34], they all aim to ***highlight important sub-spaces***

***in the input domain that have a positive or negative influence on model predictions***. We have incorporated four such XAI methods suitable for GNN models:

- **GNNExplainer:** GNNExplainer [33] uses an optimization approach to generate explanations, learning soft masks for edges and hard masks for node features to pinpoint significant inputs. Masks are optimized to align the mutual information between the original and new predictions.
- **GNN-LRP (Layerwise Relevance Propagation):** A decomposition based method, GNN-LRP [26] assesses the importance of graph walks using high-order Taylor decomposition to relate predictions back to specific graph walks. GNN-LRP is unique in that it unrolls the message-passing layers of GNNs to extract important graph walks.
- **DeepLIFT:** Adapted for deep graph models from the image classification task [27], DeepLIFT decomposes predictions to node contributions, serving as an efficient proxy for Shapley values.
- **GradCAM:** Initially for image classification [24], GradCAM is adapted to graph models to assess node significance by integrating hidden feature maps with gradients to highlight nodes critical to predictions.

## 4.4 Visualization Schemes

XplainScreen provides two visualization schemes including (1) the standard `RDKit` [6] style and (2) the custom style for efficient visualization employing novel unified explanation data-structure.

- **The Standard `Rdkit` Style:** Important sub-spaces within molecular graphs can be visualized and highlighted by the `highlightAtoms` feature of `RDKit` by specifying the relevant portions of the SMILES string [6].
- **The Custom Style using Unified Explanation Data structure:** A particular challenge while working with the different explanation methods to create unified evaluation framework is to represent the explanation results in a common data structure, both for post-hoc analysis as well as consistent visualization around the molecular graph in single platform frame. As described in section-**4.3** above, out of the four methods, GNN-LRP produces graph walks of different lengths as its explanation results, while the other methods produce finer node-level and/or edge-level results. *To represent and visualize such different XAI methods in single platform, here we implement results from four XAI methods as a dictionary structure which stores the graph sub-structure as the key and corresponding explanation/importance scores from the methods as value.* Graph walks are stored as a sequence of node ids (i.e, atom ids for molecular graphs) corresponding to the walk sequence. For explanation methods producing node and edge level information, we expand the sub-structure keys in our dictionary to match the length of the longest graph walk sequence. Nodes are expanded to a repeated sequence of the same node id and edges are expanded to a sequence with the two node ids associated with the edge. This facilitates efficient query of node and edge combinations for explanation evaluation from the sub-structure keys and still maintain a consistent representation for different explanation results. Fig 2 illustrates the idea of our proposed datastructure to store different levels of explanation results. This is essentially a dictionary which stores the graph substructure as the *key* and corresponding explanation/importance scores from the methods as *value*. Graph walks are stored as a sequence of node ids (i.e, atom ids for molecular graphs) corresponding to the walk sequence. This is illustrated with the green walk in Fig 2. For explanation methods producing node and edge level information, we expand the sub-structure keys in our dictionary to match the length of the longest graph walk sequence. Nodes are expanded to a repeated sequence of the same node id (highlighted in color blue) and edges are expanded to a sequence with the two node ids associated with the edge (highlighted with orange color). This facilitates efficient query of node and edge combinations for explanation evaluation from the sub-structure keys and still maintain a consistent representation for different explanation results.

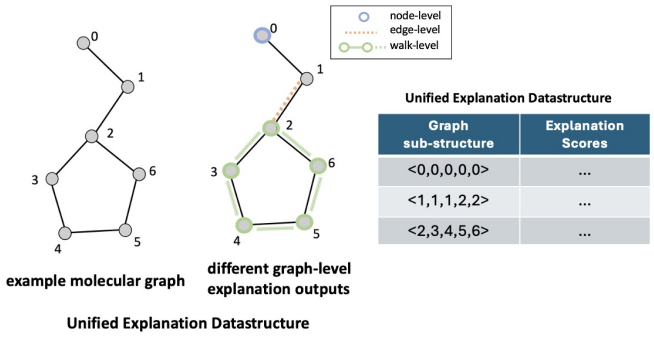



**Figure 2: Our unified explanation data-structure**

# 5 Demonstration Platform

Our platform operates on a web-based client-server architecture developed using streamlit [14], whereby the user interacts with the models through a web browser. All processes, including the execution of XAI algorithms, and the visualization and interpretation of results, are conducted on the server side. This architecture enables users to utilize server computing resources for model execution while the client side manages input selections and visualization tasks. It should be noted that a single execution of XplainScreen can be completed within just a few seconds on a single processor; therefore, users can feasibly perform both the execution of XAI algorithms and visualization on a modern laptop. A snapshot of the user interface is shown in Fig 3. It consists of two sub-components: (1) a control panel, (2) an explanation panel. We describe each component in this section.

## 5.1 Control Panel

The leftmost column of the interface (Fig 3, left) is designed for configuring four input parameters and interacting with the models. Initially, an end-user selects their desired dataset, explanation method, type of pre-trained GNN model, and visualization style. The control panel also provides an option to either choose a SMILES string from the test dataset dropdown menu (select from Tox21/ClinTox Database box) or enter a custom SMILES string manually (Provide Custom SMILE box).

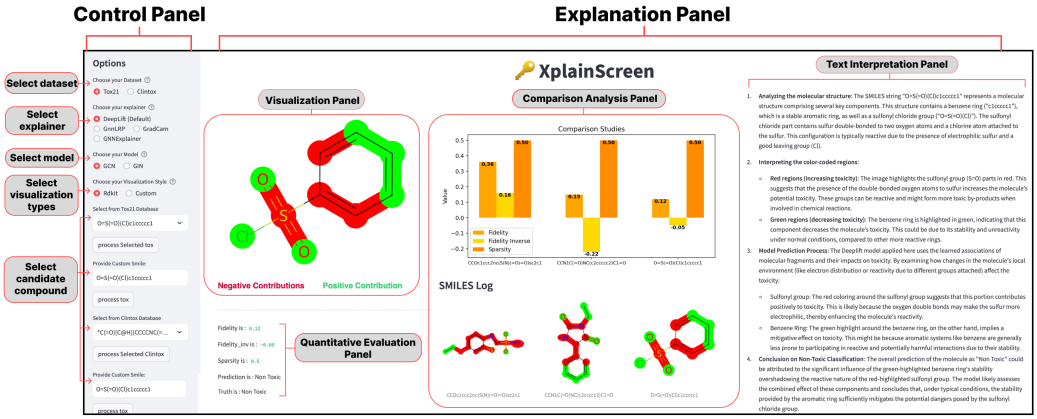



**Figure 3: User interface for the demonstration platform of XplainScreen. Left panel: Control. Right panel: Explain.**

## 5.2 Explanation Panel

Explanation panel is divided by four sub-panels: **(1)** ***Visualization Panel,*** **(2)*Quantitative Evaluation Panel***, **(3)*Comparison Analysis Panel***, and **(4)*Textual Interpretation Panel***, showing four types of explanation, all based on the assessment of the target XAI algorithms. Explanation in each panel is described as follows:

**(1) Visualization Panel.** XplainScreen enables users to comprehend why a particular drug screening model classifies a drug as toxic or non-toxic. This ***model-derived explanation*** includes visualizations that highlight crucial sub-spaces in the molecular graph, using colors to indicate influence on the model prediction result: green for positive (non-toxic in Fig 3) and red for negative (toxic in Fig 3) influences. This allows users to discern which sub-graphs are critical in determining toxicity according to the model.

**(2) Quantitative Evaluation Panel.** Qualitative evaluations alone, such as visualization and text descriptions, can be subjective and may not provide reliable model explanations. Therefore, we include several quantitative metrics to assess the quality of explanations. We employ the following three quantitative metrics [9, 24, 34] - ***Fidelity+ :*** The difference of accuracy between the original predictions and the new predictions after masking out important input features. ***Fidelity- :*** The fraction of prediction change by keeping important input features and removing unimportant features as highlighted by the explanation results. ***Sparsity:*** The fraction of features selected as important by explanation methods. Good explanations should be sparse, capturing the most important input features and ignore the irrelevant ones.

**(3) Comparison Analysis Panel.** To compare explanation results for the target molecule in the current XplainScreen session with those from previous sessions, we have included a histogram that compares quantitative metrics across different sessions. Additionally, we show a colored molecular-graph contribution visualization for each session below histogram. As end-users progress through sessions for different molecules, the comparison analysis panel is updated, with new columns for histogram bars and corresponding graph visualization for each tested molecule.

**(4) Textual Interpretation Panel.** The visualization results are scientifically interpreted by GPT-4 with Vision (GPT-4V) [3, 7]. GPT-4V generates interpretations based on the colored graph visualization. For the prompt, we direct to provide an interpretation that includes four aspects: an analysis of the molecular substructure using the provided SMILES string, an interpretation of the toxicity's negative (red) and positive (green) segments, a description of how the chosen XAI method assesses these substructures to determine toxicity, and an overview of the model's prediction. This approach aims to improve users' comprehension of model predictions, regardless of their domain expertise.

## 6 User Study

To evaluate how XplainScreen supports end-users in virtual drug screening, we conducted a user study with seven pharmacological scientists. Initially, we introduced the detailed features of XplainScreen and demonstrated the toxicity analysis for 10 candidate drugs (5 toxic and 5 non-toxic drugs). Participants were then asked to rank the usefulness of XplainScreen on a scale from 0 (not useful) to 5 (very useful). The average score is **4.14**. We also collected feedback on the pros and cons of XplainScreen. Participants highlighted its capability to validate AI-model alignment with scientific interpretation, enhancing trust in the AI model. They suggested further integration of this framework with scientific drug simulations to compare AI predictions with physics-driven simulations.

## 7 Conclusion

We showcase the XplainScreen prototype to evaluate explanation methods for GNN-based drug screening models, offering a user-friendly, web-based interactive platform that allows for the selection of specific GNN-based drug screening models and multiple cutting-edge explainable AI methods.

## Acknowledgement

This work was supported by the Ewha Womans University Research Grant of 2024, and was partly supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.RS-2022-00155966, Artificial Intelligence Convergence Innovation Human Resources Development (Ewha Womans University)) and Korea Evaluation Institute of Industrial Technology(KEIT) grant funded by the Korea government(MOTIE) (No.RS-2022-00155966)

## References


[1] [n. d.]. Drug-likeliness. https://en.wikipedia.org/wiki/Druglikeness.

[2] [n. d.]. Tox21 Challenge. https://tripod.nih.gov/tox21/challenge/.

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

[4] Chirag Agarwal, Owen Queen, Himabindu Lakkaraju, and Marinka Zitnik. 2023. Evaluating explainability for graph neural networks. *Scientific Data* 10, 1 (2023), 144.

[5] Luiz Anastacio Alves, Natiele Carla da Silva Ferreira, Victor Maricato, Anael Viana Pinto Alberto, Evellyn Araujo Dias, and Nt Jose Aguiar Coelho. 2022. Graph neural networks as a potential tool in improving virtual screening programs. *Frontiers in Chemistry* 9 (2022), 787194.

[6] A Patrícia Bento, Anne Hersey, Eloy Félix, Greg Landrum, Anna Gaulton, Francis Atkinson, Louisa J Bellis, Marleen De Veij, and Andrew R Leach. 2020. An open source chemical structure curation pipeline using RDKit. *Journal of Cheminformatics* 12 (2020), 1–16.

[7] Jiawen Deng, Kiyan Heybati, and Matthew Shammas-Toma. 2024. When vision meets reality: Exploring the clinical applicability of GPT-4 with vision. , 110101 pages.

[8] Andreas Holzinger, Bernd Malle, Anna Saranti, and Bastian Pfeifer. 2021. Towards multi-modal causability with graph neural networks enabling information fusion for explainable AI. *Information Fusion* 71 (2021), 28–37.

[9] Sara Hooker, Dumitru Erhan, Pieter-Jan Kindermans, and Been Kim. 2019. A benchmark for interpretability methods in deep neural networks. *Advances in neural information processing systems* 32 (2019).

[10] Wei Huang, Xingyu Zhao, Gaojie Jin, and Xiaowei Huang. 2022. SAFARI: Versatile and Efficient Evaluations for Robustness of Interpretability. *arXiv preprint arXiv:2208.09418* (2022).

[11] James P Hughes, Stephen Rees, S Barrett Kalindjian, and Karen L Philpott. 2011. Principles of early drug discovery. *British journal of pharmacology* 162, 6 (2011), 1239–1249.

[12] B Iswarya and K Manimekalai. 2022. Drug Discovery With XAI Using Deep Learning. In *Principles and Methods of Explainable Artificial Intelligence in Healthcare*. IGI Global, 131–149.

[13] José Jiménez-Luna, Francesca Grisoni, and Gisbert Schneider. 2020. Drug discovery with explainable artificial intelligence. *Nature Machine Intelligence* 2, 10 (2020), 573–584.

[14] Mohammad Khorasani, Mohamed Abdou, and J Hernández Fernández. 2022. Web Application Development with Streamlit. *Software Development* (2022), 498–507.

[15] Jintae Kim, Sera Park, Dongbo Min, and Wankyu Kim. 2021. Comprehensive Survey of Recent Drug Discovery Using Deep Learning. *International Journal of Molecular Sciences* 22, 18 (2021). https://doi.org/10.3390/ijms22189983

[16] Antonio Lavecchia. 2019. Deep learning in drug discovery: opportunities, challenges and future prospects. *Drug discovery today* 24, 10 (2019), 2017–2032.

[17] Yiqiao Li, Jianlong Zhou, Sunny Verma, and Fang Chen. 2022. A survey of explainable graph neural networks: Taxonomy and evaluation metrics. *arXiv preprint arXiv:2207.12599* (2022).

[18] Meng Liu, Youzhi Luo, Limei Wang, Yaochen Xie, Hao Yuan, Shurui Gui, Haiyang Yu, Zhao Xu, Jingtun Zhang, Yi Liu, Keqiang Yan, Haoran Liu, Cong Fu, Bora M Oztekin, Xuan Zhang, and Shuiwang Ji. 2021. DIG: A Turnkey Library for Diving into Graph Deep Learning Research. *Journal of Machine Learning Research* 22, 240 (2021), 1–9. http://jmlr.org/papers/v22/21-0343.html

[19] Kit-Kay Mak and Mallikarjuna Rao Pichika. 2019. Artificial intelligence in drug development: present status and future prospects. *Drug discovery today* 24, 3 (2019), 773–780.

[20] Grégoire Montavon, Wojciech Samek, and Klaus-Robert Müller. 2018. Methods for interpreting and understanding deep neural networks. *Digital Signal Processing* 73 (2018), 1–15. https://doi.org/10.1016/j.dsp.2017.10.011

[21] Rick Ng. 2015. *Drugs: from discovery to approval*. John Wiley & Sons.

[22] Paul A Novick, Oscar F Ortiz, Jared Poelman, Amir Y Abdulhay, and Vijay S Pande. 2013. SWEETLEAD: an in silico database of approved drugs, regulated chemicals, and herbal isolates for computer-aided drug discovery. *PloS one* 8, 11 (2013), e79568.

[23] Elina Petrova. 2013. Innovation in the pharmaceutical industry: The process of drug discovery and development. In *Innovation and Marketing in the Pharmaceutical Industry: Emerging Practices, Research, and Policies*. Springer, 19–81.

[24] Phillip E Pope, Soheil Kolouri, Mohammad Rostami, Charles E Martin, and Heiko Hoffmann. 2019. Explainability methods for graph convolutional neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10772–10781.

[25] Jiahua Rao, Shuangjia Zheng, and Yuedong Yang. 2021. Quantitative Evaluation of Explainable Graph Neural Networks for Molecular Property Prediction. *arXiv preprint arXiv:2107.04119* (2021).

[26] Thomas Schnake, Oliver Eberle, Jonas Lederer, Shinichi Nakajima, Kristof T Schütt, Klaus-Robert Müller, and Grégoire Montavon. 2020. Higher-order explanations of graph neural networks via relevant walks. *arXiv preprint arXiv:2006.03589* (2020).

[27] Avanti Shrikumar, Peyton Greenside, and Anshul Kundaje. 2017. Learning important features through propagating activation differences. In *International conference on machine learning*. PMLR, 3145–3153.

[28] Max Welling and Thomas N Kipf. 2016. Semi-supervised classification with graph convolutional networks. In *J. International Conference on Learning Representations (ICLR 2017)*.

[29] Oliver Wieder, Stefan Kohlbacher, Mélaine Kuenemann, Arthur Garon, Pierre Ducrot, Thomas Seidel, and Thierry Langer. 2020. A compact review of molecular property prediction with graph neural networks. *Drug Discovery Today: Technologies* 37 (2020), 1–12.

[30] Oliver Wieder, Stefan Kohlbacher, Mélaine Kuenemann, Arthur Garon, Pierre Ducrot, Thomas Seidel, and Thierry Langer. 2020. A compact review of molecular property prediction with graph neural networks. *Drug Discovery Today: Technologies* 37 (2020), 1–12. https://doi.org/10.1016/j.ddtec.2020.11.009

[31] Jiacheng Xiong, Zhaoping Xiong, Kaixian Chen, Hualiang Jiang, and Mingyue Zheng. 2021. Graph neural networks for automated de novo drug design. *Drug discovery today* 26, 6 (2021), 1382–1393.

[32] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. 2018. How powerful are graph neural networks? *arXiv preprint arXiv:1810.00826* (2018).

[33] Zhitao Ying, Dylan Bourgeois, Jiaxuan You, Marinka Zitnik, and Jure Leskovec. 2019. Gnnexplainer: Generating explanations for graph neural networks. *Advances in neural information processing systems* 32 (2019).

[34] Hao Yuan, Haiyang Yu, Shurui Gui, and Shuiwang Ji. 2022. Explainability in Graph Neural Networks: A Taxonomic Survey. *IEEE transactions on pattern analysis and machine intelligence* PP (2022).