All You Need is LOVE: Large Optimized Vector Embeddings Network for Drug Repurposing

Sepehr Asgarian
Klick Inc.
Klick Applied Sciences
Toronto, Ontario, Canada
sasgarian@klick.com

Sina Akbarian
Klick Inc.
Klick Applied Sciences
Toronto, Ontario, Canada
sakbarian@klick.com

Jouhyun Jeon
Klick Inc.
Klick Applied Sciences
Toronto, Ontario, Canada
cjeon@klick.com

Abstract

Traditional drug development is a resource-intensive and time-consuming process with a high rate of failure. To expedite this process, researchers have turned to computational approaches to construct comprehensive graphs of drug-disease associations and explore drug repurposing, finding novel therapeutic applications for existing medications. In parallel, the rapid advancement of the machine-learning field, coupled with the evolution of Natural Language Processing, shows capabilities for reasoning and extracting relationships across various domains. In this paper, we introduce LOVENet (Large Optimized Vector Embeddings Network), a new framework maximizing the synergistic effects of knowledge graphs and large language models (LLMs) to discover novel therapeutic uses for pre-existing drugs. Specifically, our approach fuses information from pairs of embedding from Llama2 and heterogeneous knowledge graphs to derive complex relations of drugs and diseases. To empirically validate our methodology, we conducted benchmarking experiments against state-of-the-art algorithms, utilizing three distinct datasets. Our results demonstrate that LOVENet consistently outperforms all other baselines. The code for this project is available at https://github.com/KlickInc/brave-foundry-drug-repurposing.

1 Introduction

The traditional experiment-guided drug development is a time-consuming and labor-intensive process. On average, the journey from a new drug discovery to market takes 12 years and costs about 3 billion with only a small success rate Kraljevic et al. [2004]. As more biological and chemical knowledge is accumulated over time, computational approaches, which can integrate and analyze large high-throughput data, are considered as an alternative method to accelerate the process of drug discovery and time to market. Indeed, machine-learning or deep-learning based researches have explored several pharmaceutical topics including drug-drug interaction prediction Nyamabo et al. [2022], drug adverse effect prediction Dey et al. [2018], as well as drug repurposing Kang et al. [2023], Zhang et al. [2018].

The NeurIPS 2023 Workshop on New Frontiers of AI for Drug Discovery and Development (AI4D3 2023), New Orleans, LA, USA, 2023.

Drug repurposing, also known as drug reprofiling or drug repositioning, is the process of discovering new indications for existing drugs that were developed for different therapeutic purposes. Drug repurposing provides a variety of advantages during the process of drug discovery, such as lower risk of failure, less investment, and a shorter development time frame Kulkarni et al. [2023]. In the beginning, most repurposing drugs were accidentally discovered in experimental or clinical settings. With the continuous updating of large-scale biological, chemical, and pharmacological datasets, computational technology has enabled researchers to predict drug-disease associations in data-driven ways.

There has been an unprecedented surge of interest in Artificial Intelligence (AI) algorithms, particularly advanced large-language models (LLMs). These LLMs have demonstrated an unparalleled ability to transform and analyze text data, making them capable of revolutionizing research across biological and medical domains, such as ligand/drug design and DNA/protein sequencing Parrot et al. [2023], Bagabir et al. [2022]. In this paper, we introduce the LOVENet architecture: a novel approach that leverages the potential of LLM models and knowledge graph data to find new indications of existing drugs. Our novel methodology harnesses the power of LLM models to extract comprehensive representations of drugs and diseases. These representations are then integrated with a knowledge graph, resulting in a potent fusion of textual and structured data sources. To validate our approach, we benchmarked our results against state-of-the-art methods across three distinct datasets. The findings from our study not only underscore the transformative potential inherent to LOVENet but also provide insights on its capacity to push the boundaries of drug discovery.

2 Related Work

2.1 Drug repurposing prediction

Inferring potential new uses for approved or investigational drugs is an essential step in the drug development process. With the accumulation of extensive biological, chemical, and medical datasets, there has been growing interest in utilizing computational approaches to identify potential drug repurposing candidates. The most popular computational approaches include (1) complex network methods, (2) molecular docking, and (3) machine- or deep-learning methods.

Complex network methods evaluate the features of biological networks and recognize distinctive network patterns capable of representing associations between drugs and diseases Lotfi Shahreza et al. [2018]. Cheng et al. developed the method to quantify the network proximity between molecular determinants of diseases and drug targets within the human protein-protein interaction network and predict potential repurposing drugs that are closely connected to disease modules within the network Cheng et al. [2018]. This network-driven drug discovery algorithm has been used to identify potential repurposing drugs for COVID-19 Fiscon et al. [2021].

Molecular docking techniques involve extensive simulations of the physical interactions between drugs and disease-related proteins, aiming to identify the most favorable binding conformation of drugs for given disease-related proteins Kukol et al. [2008]. To conduct reliable and precise docking simulations, it is crucial to have access to three-dimensional (3D) structures of both chemical ligands and protein targets. The limited availability of 3D structures for many biologically significant proteins has constrained the applicability and effectiveness of docking simulations.

Machine-learning methods appear more favorable than docking simulations, as they can examine multiple large-scale datasets and identify a list of potential promising drug candidates for further experimental screening and validation. Rodrigues et al. proposed a machine-learning framework that uses large-scale pathological datasets to quantify the association between Alzheimer's disease pathology and human genes and predicted target affinities of drugs Rodriguez et al. [2021]. Physicochemical characteristics of drug compounds were fed into the Recurrent Neural Network (RNN) model to identify potential repurposing drugs that share similar molecular characteristics of a given disease Duvenaud et al. [2015]. Moreover, graph neural networks (GNN) have been proposed to find potential drug-disease associations. These neural networks are a form of representation learning that attempts to learn node embedding in heterogeneous graphs and further facilitate downstream applications such as multi-labeled drug repurposing Sadeghi et al. [2022].

2.2 Generative AI models for drug discovery

In the brief period following the introduction of ChatGPT, there has been a surge in the proposal of large language models (LLMs) and generative AI systems for the purpose of conceiving and characterizing potential drug candidates Vert [2023]. One research team has introduced an AI-driven platform for generative drug design by utilizing a neural network to estimate the synthesizability of molecules Parrot et al. [2023]. The synergy between generative AI and active learning facilitates the design and refinement of chemical ligands with high affinity for binding to CDK2 and KRAS Filella-Merce et al. [2023]. Furthermore, AI-generated antibody variants have been proposed to target human epidermal growth factor receptor 2 (HER 2) and treat breast cancer Shanehsazzadeh et al. [2023]. Although generative AI technology has been applied to diverse drug discovery fields, there are a limited number of studies investigating its application to drug repurposing.

3 Methods

Predicting repurposing drugs (drug-disease associations) in a graph setting can be conceptualized as a task of link prediction. This task seeks to estimate the likelihood of an existing association between a specified drug v^r and a disease v^d node.

3.1 Benchmark dataset

We compiled three well-known drug-disease association benchmark datasets that are commonly used for drug repurposing studies (colloquially termed as B-, C-, and F-datasets). These datasets have different network characteristics including network connectivity and positive/negative data ratios Gu et al. [2022]. B-dataset is proposed by Zhang et al. and comprises 269 drugs and 598 diseases with 18,416 drug-disease associations Zhang et al. [2018]. C-dataset has 2,352 known drug-disease associations which are composed of 663 drugs registered in DrugBank and 409 genetic diseases listed in the Online Mendelian Inheritance in Man (OMIM) database Luo et al. [2016]. The F-dataset is composed of 593 drugs from DrugBank and 313 associated diseases in the OMIM database (in total, 1,933 associations) Gottlieb et al. [2011].

3.2 Construction of drug-disease heterogeneous network

Three types of biological associations were used to construct a heterogeneous graph representing drugs and diseases. First, drug-drug associations were determined utilizing the Simplified Molecular Input Line-Entry System (SMILES) to extract binarized molecular fingerprints, with pairwise drug similarities computed using Tanimoto metrics Godden et al. [2000]. Second, disease-disease associations were defined by applying the MinMiner algorithm, which gauged the co-occurrence of medical subject headings (MeSH) terms within disease descriptions. Then, drug-disease associations were characterized using an adjacency matrix, which detailed node connections, and a node feature matrix describing each node's individual attributes.

3.3 Architecture of LOVENet

LOVENet is a fused algorithm that integrates graph convolutional network and LLM model (Figure 1). This architecture enabled the training of both knowledge graph and LLM embedding simultaneously. We adopted and modified the heterogeneous graph network architecture proposed by Gu et al. [2022]. The knowledge graph representation is as follows:

$$Z_{i,j}^{KG} = \text{GNN-Layer}(v_i^r, v_j^d). \tag{1}$$

where i and j are indexes to distinguish between various diseases and drugs. The embedding $Z_{i,j}^{KG}$ denotes the KG model's output for a specific disease-drug pairing in the knowledge graph.

Next, we employed Llama2 with 13 billion parameters Touvron et al. [2023] as our LLM model. To extract embedding for every drug-disease association, the following sentence structure: "Is this drug (v^r) related to this disease (v^d) answer yes or no?" was fed into the Llama2. This constructed sentence was the foundation for generating the next word embedding probability. Our workflow

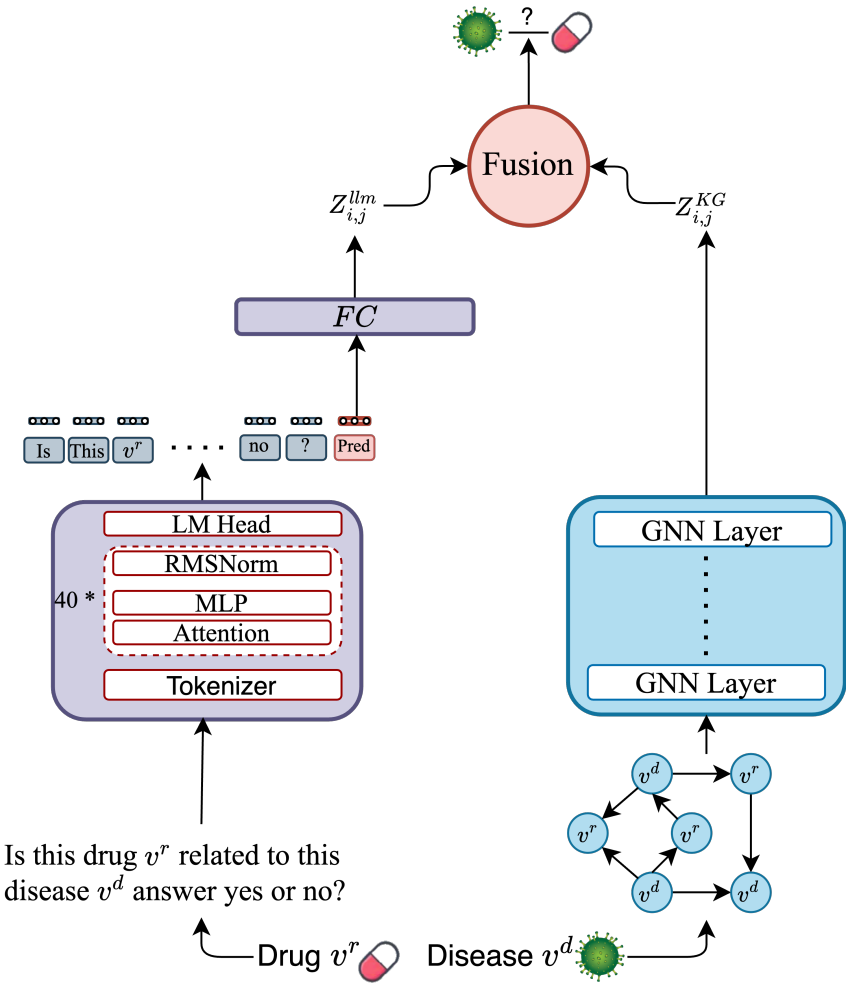

Figure 1: The architecture of LOVENet. The LOVENet employs a heterogeneous graph representing drug-disease ($v^r - v^d$) associations and integrates it with Llama2 embedding. Getting representation from Llama2 includes sentence creation, tokenization, which pass through a series of Llama2 blocks. Within the fusion layer, GNN and Llama2 embedding are concatenated together, and fed into a fully connected network. Next, the output of the fusion layer is used to predict link probabilities between drugs (v^r) and diseases (v^d).

began by applying tokenization to the input sentence, followed by passing it through the first 40 blocks of Llama2. We then extracted the output from the Language Model (LM) head layer, which provided word probability distributions. Using this approach, we could create a dictionary where each drug-disease association was key, and its Llama2 embedding was a corresponding value (purple box in Figure 1). The Llama2 embedding is described as follows:

$$R_{i,j}^1, \ldots, R_{i,j}^{n+1} = \text{Llama2}\left(w^1, w^2, v_i^r, ..., v_j^d, ..., w^n\right). \tag{2}$$

where sequence w represents the input word sequence for the Llama2 models, where n denotes each word, including drugs and diseases. The $(R_{i,j}^1, \ldots, R_{i,j}^{n+1})$ represents the output of Llama2, and $R_{i,j}^{n+1}$ represents a tensor with size of 32,000 showing the probability of the next word.

Our procedure followed by applying L2 normalization to the output of the heterogeneous GNN. Simultaneously, L2 normalization is applied to the Llama2 embedding, which is then processed through a fully connected network. These two results are subsequently fused and passed through a sigmoid function. The following equation shows how knowledge graphs and Llama2 embedding

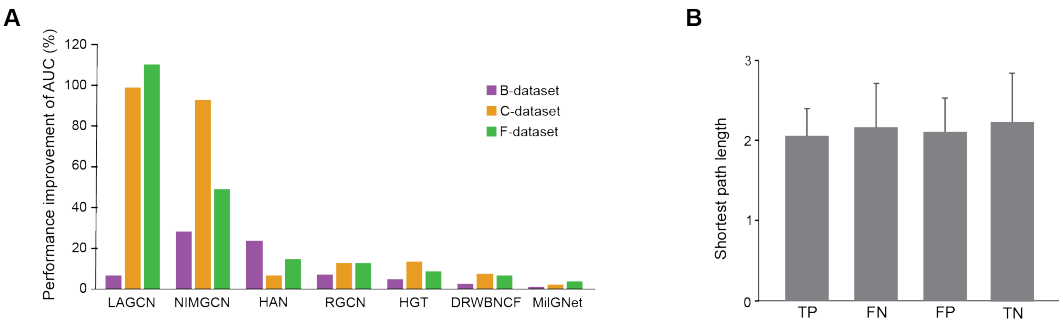

Figure 2: Prediction performance of LOVENet. A. Percentage performance improvement of LOVENet. The performance of seven state-of-the-art methods were compared with the performance of LOVENet. Purple, orange, and green indicated performance improvement from B-, C- and F-benchmark datasets, respectively. B. Shortest path lengths of true positives (TP), false negatives (FN), false positives (FP), and true negatives (TN).

integrate.

$$Z_{i,j}^{llm} = \text{Fully-Connected}(R_{i,j}^{n+1}), \tag{3}$$

$$Z_{i,j}^{Fusion} = \text{Fusion-Layer}(Z_{i,j}^{llm}, Z_{i,j}^{KG}). \tag{4}$$

where $Z_{i,j}^{Fusion}$ is the fusion of the concatenation of the outputs of the fully connected layer after Llama2 embedding $Z_{i,j}^{llm}$ and KG model embedding $Z_{i,j}^{KG}$.

3.4 Training LOVENet

We fully trained the LOVENet network, including the heterogeneous GNN and LLM embedding utilizing cross-entropy loss function. To show the robustness of our findings, we applied a 5-fold cross-validation, and repeated the entire experiment 10 times with distinct random seeds.

3.5 Performance evaluation

To evaluate the performance of LOVENet, seven state-of-the-art methods that were recently proposed to identify repurposing drugs (i.e. drug-disease associations) were extracted from Gu et al. [2022], including (NIMGCN Li et al. [2020], LAGCN Yu et al. [2021], DRWBNCF Meng et al. [2022], RGCN Kipf and Welling [2016], HAN Wang et al. [2019], HGT Hu et al. [2020], and MilGNet Gu et al. [2022]). We then compare their performance with LOVENet. However, due to space constraints in the paper, we focus our comparison primarily on MilGNet, the best algorithm in the benchmark, and provide the rest in Appendix 1.

To ensure a comprehensive evaluation of our approach, we compare our performance with the sole use of Llama2. In this scenario, we used the same structure of Llama2 as discussed in the previous section without containing GNN followed by the same training procedure as LOVENet.

For further evaluation, LOVENet model trained by the B-dataset was applied to an independent validation dataset. This independent validation set was created from manually curated drug-disease associations, and composed of 266 drugs and 571 diseases with 5,427 drug-disease associations. Drugs and disease in the independent validation dataset are part of B-dataset but their associations are not included in the B-dataset. For the performance evaluation, AUROC (Area Under the Receiver Operating Characteristic curve), AUPR (Area Under the Precision-Recall curve), and F1-score to ensure a thorough and meaningful appraisal of our model's capabilities.

To understand drug-disease associations within a biological system, we created a network of drugs, genes, pathways, and diseases. Drug-gene, drug-pathway, disease-gene, and disease-pathway associations were extracted from the Gene Ontology (GO) database Ashburner et al. [2000] and Kyoto Encyclopedia of Genes and Genomes (KEGG) Kanehisa et al. [2012].

4 Results and Discussion

4.1 Performance of LOVE predictive model for drug repurposing

We trained LOVENet models using three established benchmark datasets (B, C, and F-datasets), and compared their performance with seven other state-of-the-art methods. Through a 10-repetitive 5-fold cross-validation, we observed that LOVENet consistently outperformed other state-of-the-art methods (Figure 2A). LOVENet achieved 0.91±0.05 of AUROC which improved 24.52±33.78% compared to other seven state-of-the-art methods across all three benchmark datasets (Figure 2A and Table A.1). On average, LOVENet achieved 0.52±0.06 of AUPR, and 0.52±0.05 of F1-score (Table 1). Notably, LOVENet exhibited an improvement of 2.34% in AUROC, 11.79% in AUPR, and 5.52% in F1 compared to MilGNet, which had shown the best performance among seven other methods (Table 1).

We also compared dataset-specific prediction performance. LOVENet achieved reasonable performance across all B, C, and F datasets. LOVENet showed 0.86 (B-dataset), 0.95 (C-dataset), and 0.93 (F-dataset) of AUROC (Table 1). We also observed that LOVENet showed a significant performance improvement in C- and F-data in terms of AUPR and F1-score when compared to MilGNet. Median improvements were 32.90% for AUPR, and 56.97% for F1-score compared to the other seven methods (Table A.1).

Moreover, a synergistic performance enhancement was observed by integrating the large language model with the heterogeneous knowledge graph. Compare to the performance of Llama2, in the overall evaluation, LOVENet showed improvements of 3.41%, 30%, and 30% in AUROC, AUPR and F1-score, respectively (Table 1). This indicates that LOVENet consistently delivered strong performance, maintaining stability and high accuracy irrespective of the data type, including network density or size.

Table 1: Performance of LOVENet, MilGNet, and Llama2 on 3 benchmark datasets

Dataset	AUROC			AUPR			F1-score		
	LOVENet	MilGNet	Llama2	LOVENet	MilGNet	Llama2	LOVENet	MilGNet	Llama2
B-data	0.86	0.85	0.84	0.52	0.47	0.47	0.51	0.49	0.47
C-data	0.95	0.93	0.91	0.58	0.52	0.39	0.59	0.55	0.36
F-data	0.93	0.89	0.91	0.46	0.41	0.35	0.48	0.46	0.38
Average	0.91	0.89	0.88	0.52	0.47	0.40	0.52	0.50	0.40

Average AUROC, AUPR, and F1-score were tabulated. Standard deviation from 10-repetitive 5-CV is presented at Table A.1

For further evaluation, the LOVENet model trained by B-dataset was applied to an independent validation dataset (details in section 3.5). This independent validation dataset was never used for training and testing the LOVENet models, and provided literature evidence that represents the confidence of a given drug-disease association. From the independant validation, LOVENet demonstrated not only a commendable accuracy of 78.78% in identifying drug-disease associations but also an improvement in predictive accuracy, particularly when dealing with highly reliable drug-disease associations (Table 2). From 525 drug-disease associations supported by at least 5 literature evidence, LOVENet correctly identified 518 associations, and achieved 98.67% of accuracy. Furthermore, when considering drug-disease associations supported by at least 10 literature evidence, LOVENet achieved a perfect 100% accuracy.

Table 2: Performance of LOVENet on an independent validation dataset

Association confidence*	# of drug-disease associations	Accuracy (%)	Correctly predicted drug-disease association	Incorrectly predicted drug-disease associations
0	5,427	78.78	4,273	1,154
1	2,099	90.94	1,909	190
5	525	98.67	518	7
10	259	100	259	0
15	163	100	163	0

Association confidence indicates the number of literature references that describe the association between drugs and diseases.

4.2 Network properties of drug-disease associations within a biological network

Drugs experience chemical modifications through various biological pathways, thereby influencing or changing the function of target genes, which subsequently leads to an impact on the disease treatment. To gain comprehensive insights into drug-disease associations, we measured the shortest path lengths between drugs and diseases within the network of drugs, biological pathways, genes, and diseases. LOVENet model trained on the B-dataset was considered for further analyses since over 80% of drugs and diseases in the B-dataset were included in the network (Table 2).

True positives, which represent the correctly predicted positive drug-disease associations, tend to be closely located in the network. Their shortest path length was 2.06 ± 0.34 (Figure 2B). In contrast, true negatives (correctly predicted negative associations) displayed the longest shortest path length at 2.24 ± 0.61 (P-value between TP and TN = 1.80×10^{-235}, Mann-Whitney U test). It suggested that in cases where drugs are genuinely effective in addressing disease pathology, they tend to be closely connected to the associated diseases with the biological network. Both false negatives and false positives showed intermediate levels of shortest path length suggesting that misclassifications may arise from various factors such as complex interactions within the biological network, uncharacterized biological pathways, or limitation of network models.

4.3 Case studies

To further demonstrate the ability of LOVENet in uncovering novel drug repurposing, we selected top scored drug-disease associations and scrutinized their supporting experimental evidences. Quinidine, a class IA anti-arrhythmic agent, is currently used to treat heart rhythm disturbances. It was highlighted by LOVENet for a new association with seizure. Notably, empirical data reveal that early Quinidine treatment significantly reduced seizure burden, and improved quality of life in patients Dilena et al. [2018]. Also, Quinidine reduced seizure frequency among individuals with KCNT-1 positive epilepsy Mikati et al. [2015].

Buspirone is employed in treating anxiety disorders. LOVENet has revealed a novel association with substance withdrawal syndrome. Studies showed that buspirone played a role in alleviating opioid withdrawal Buydens-Branchey et al. [2005] and was effective in reducing withdrawal symptoms that follow opiate use cessation Rose et al. [2003].

Fluoxetine is a selective serotonin reuptake inhibitor and approved for the treatment of major depressive disorder, anxiety, and panic disorder. LOVENet predicted a novel association with attention deficit disorder with hyperactivity (ADHD). Indeed, in ADHD children, fluoxetine monotherapy has been shown to significantly improve inattentivenss and hyperactivity Barrickman et al. [1991]. Off-label prescriptions of fluxetine for ADHD are common.

4.4 Limitations

While LOVENet demonstrated superior performance compared to existing drug repurposing methods, this study has several limitations. Firstly, the model complexity of LOVENet is higher than that of other approaches due to the incorporation of LLM inferences into existing heterogeneous networks for enhanced model optimization. Secondly, there is room for reevaluating the definition and creation of meta-paths that represent potential drug-disease associations, with a focus on obtaining more dependable meta-path-level instances. For instance, this could involve generating meta-paths based on shared association proteins for drug-disease pairs.

5 Conclusion

In this study, we demonstrated the effectiveness of LOVENet, a fusion approach between AI-powered language models and knowledge graphs for drug repurposing. The knowledge graph offers valuable information regarding the relationships between drugs and diseases, and the large language model utilizes its capacity to infer drug-disease associations from extensive textual data it has been trained on. Future development should explore the integration of LOVENet with other biological associations and datasets and its performance on larger datasets to evaluate its reliability across various experimental scenarios.

6 Acknowledgments

Thanks to all of the project members, supporters, and researchers at Klick Inc. for the successful development, implementation, and evaluation of this research. On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Michael Ashburner, Catherine A Ball, Judith A Blake, David Botstein, Heather Butler, J Michael Cherry, Allan P Davis, Kara Dolinski, Selina S Dwight, Janan T Eppig, et al. Gene ontology: tool for the unification of biology. *Nature genetics*, 25(1):25–29, 2000.

Sali Abubaker Bagabir, Nahla Khamis Ibrahim, Hala Abubaker Bagabir, and Raghdah Hashem Ateeq. Covid-19 and artificial intelligence: Genome sequencing, drug development and vaccine discovery. *Journal of Infection and Public Health*, 15(2):289–296, 2022.

Les Barrickman, Russell Noyes, Samuel Kuperman, Elizabeth Schumacher, and Michele Verda. Treatment of adhd with fluoxetine: a preliminary trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 30(5):762–767, 1991.

Laure Buydens-Branchey, Marc Branchey, and Christine Reel-Brander. Efficacy of buspirone in the treatment of opioid withdrawal. *Journal of clinical psychopharmacology*, 25(3):230–236, 2005.

Feixiong Cheng, Rishi J Desai, Diane E Handy, Ruisheng Wang, Sebastian Schneeweiss, Albert-László Barabási, and Joseph Loscalzo. Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nature communications*, 9(1):2691, 2018.

Sanjoy Dey, Heng Luo, Achille Fokoue, Jianying Hu, and Ping Zhang. Predicting adverse drug reactions through interpretable deep learning framework. *BMC bioinformatics*, 19(21):1–13, 2018.

Robertino Dilena, Jacopo C DiFrancesco, Maria Virginia Soldovieri, Antonella Giacobbe, Paolo Ambrosino, Ilaria Mosca, Maria Albina Galli, Sophie Guez, Monica Fumagalli, Francesco Miceli, et al. Early treatment with quinidine in 2 patients with epilepsy of infancy with migrating focal seizures (eimfs) due to gain-of-function kcnt1 mutations: functional studies, clinical responses, and critical issues for personalized therapy. *Neurotherapeutics*, 15:1112–1126, 2018.

David K Duvenaud, Dougal Maclaurin, Jorge Iparraguirre, Rafael Bombarell, Timothy Hirzel, Alán Aspuru-Guzik, and Ryan P Adams. Convolutional networks on graphs for learning molecular fingerprints. *Advances in neural information processing systems*, 28, 2015.

Isaac Filella-Merce, Alexis Molina, Marek Orzechowski, Lucía Díaz, Yang Ming Zhu, Julia Vilalta Mor, Laura Malo, Ajay S Yekkirala, Soumya Ray, and Victor Guallar. Optimizing drug design by merging generative ai with active learning frameworks. *arXiv preprint arXiv:2305.06334*, 2023.

Giulia Fiscon, Federica Conte, Lorenzo Farina, and Paola Paci. Saverunner: A network-based algorithm for drug repurposing and its application to covid-19. *PLoS computational biology*, 17 (2):e1008686, 2021.

Jeffrey W Godden, Ling Xue, and Jürgen Bajorath. Combinatorial preferences affect molecular similarity/diversity calculations using binary fingerprints and tanimoto coefficients. *Journal of Chemical Information and Computer Sciences*, 40(1):163–166, 2000.

Assaf Gottlieb, Gideon Y Stein, Eytan Ruppin, and Roded Sharan. Predict: a method for inferring novel drug indications with application to personalized medicine. *Molecular systems biology*, 7(1): 496, 2011.

Yaowen Gu, Si Zheng, Bowen Zhang, Hongyu Kang, and Jiao Li. Milgnet: a multi-instance learning-based heterogeneous graph network for drug repositioning. In *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 430–437. IEEE, 2022.

Ziniu Hu, Yuxiao Dong, Kuansan Wang, and Yizhou Sun. Heterogeneous graph transformer. In *Proceedings of the web conference 2020*, pages 2704–2710, 2020.

Minoru Kanehisa, Susumu Goto, Yoko Sato, Miho Furumichi, and Mao Tanabe. Kegg for integration and interpretation of large-scale molecular data sets. *Nucleic acids research*, 40(D1):D109–D114, 2012.

Hongyu Kang, Li Hou, Yaowen Gu, Xiao Lu, Jiao Li, and Qin Li. Drug–disease association prediction with literature based multi-feature fusion. *Frontiers in Pharmacology*, 14:1205144, 2023.

Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.

Sandra Kraljevic, Peter J Stambrook, and Kresimir Pavelic. Accelerating drug discovery: Although the evolution of '-omics' methodologies is still in its infancy, both the pharmaceutical industry and patients could benefit from their implementation in the drug development process. *EMBO reports*, 5(9):837–842, 2004.

Andreas Kukol et al. *Molecular modeling of proteins*, volume 443. Springer, 2008.

VS Kulkarni, V Alagarsamy, VR Solomon, PA Jose, and S Murugesan. Drug repurposing: An effective tool in modern drug discovery. *Russian Journal of Bioorganic Chemistry*, pages 1–10, 2023.

Jin Li, Sai Zhang, Tao Liu, Chenxi Ning, Zhuoxuan Zhang, and Wei Zhou. Neural inductive matrix completion with graph convolutional networks for mirna-disease association prediction. *Bioinformatics*, 36(8):2538–2546, 2020.

Maryam Lotfi Shahreza, Nasser Ghadiri, Sayed Rasoul Mousavi, Jaleh Varshosaz, and James R Green. A review of network-based approaches to drug repositioning. *Briefings in bioinformatics*, 19(5):878–892, 2018.

Huimin Luo, Jianxin Wang, Min Li, Junwei Luo, Xiaoqing Peng, Fang-Xiang Wu, and Yi Pan. Drug repositioning based on comprehensive similarity measures and bi-random walk algorithm. *Bioinformatics*, 32(17):2664–2671, 2016.

Yajie Meng, Changcheng Lu, Min Jin, Junlin Xu, Xiangxiang Zeng, and Jialiang Yang. A weighted bilinear neural collaborative filtering approach for drug repositioning. *Briefings in bioinformatics*, 23(2):bbab581, 2022.

Mohamad A Mikati, Yong-hui Jiang, Michael Carboni, Vandana Shashi, Slave Petrovski, Rebecca Spillmann, Carol J Milligan, Melody Li, Annette Grefe, Allyn McConkie, et al. Quinidine in the treatment of kcnt 1-positive epilepsies. *Annals of neurology*, 78(6):995–999, 2015.

Arnold K Nyamabo, Hui Yu, Zun Liu, and Jian-Yu Shi. Drug–drug interaction prediction with learnable size-adaptive molecular substructures. *Briefings in Bioinformatics*, 23(1):bbab441, 2022.

Maud Parrot, Hamza Tajmouati, Vinicius Barros Ribeiro da Silva, Brian Ross Atwood, Robin Fourcade, Yann Gaston-Mathé, Nicolas Do Huu, and Quentin Perron. Integrating synthetic accessibility with ai-based generative drug design. *Journal of Cheminformatics*, 15(1):1–17, 2023.

Steve Rodriguez, Clemens Hug, Petar Todorov, Nienke Moret, Sarah A Boswell, Kyle Evans, George Zhou, Nathan T Johnson, Bradley T Hyman, Peter K Sorger, et al. Machine learning identifies candidates for drug repurposing in alzheimer's disease. *Nature communications*, 12(1):1033, 2021.

Judith S Rose, Marc Branchey, Leah Wallach, and Laure Buydens-Branchey. Effects of buspirone in withdrawal from opiates. *American Journal on Addictions*, 12(3):253–259, 2003.

Shaghayegh Sadeghi, Jianguo Lu, and Alioune Ngom. An integrative heterogeneous graph neural network–based method for multi-labeled drug repurposing. *Frontiers in Pharmacology*, 13:908549, 2022.

Amir Shanehsazzadeh, Sharrol Bachas, Matt McPartlon, George Kasun, John M Sutton, Andrea K Steiger, Richard Shuai, Christa Kohnert, Goran Rakocevic, Jahir M Gutierrez, et al. Unlocking de novo antibody design with generative artificial intelligence. *bioRxiv*, pages 2023–01, 2023.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*, 2023.

Jean-Philippe Vert. How will generative ai disrupt data science in drug discovery? *Nature Biotechnology*, pages 1–2, 2023.

Xiao Wang, Houye Ji, Chuan Shi, Bai Wang, Yanfang Ye, Peng Cui, and Philip S Yu. Heterogeneous graph attention network. In *The world wide web conference*, pages 2022–2032, 2019.

Zhouxin Yu, Feng Huang, Xiaohan Zhao, Wenjie Xiao, and Wen Zhang. Predicting drug–disease associations through layer attention graph convolutional network. *Briefings in bioinformatics*, 22 (4):bbaa243, 2021.

Wen Zhang, Xiang Yue, Weiran Lin, Wenjian Wu, Ruoqi Liu, Feng Huang, and Feng Liu. Predicting drug-disease associations by using similarity constrained matrix factorization. *BMC bioinformatics*, 19:1–12, 2018.

Appendix A

Table A.1: Prediction performance comparison

	RGCN	HAN	HGT	NIMGCN	LAGCN	DRWBNCF	MilGNet	LOVENet
				AUROC				
B-dataset	0.801±0.003	0.695±0.016	0.820±0.001	0.669±0.008	0.806±0.119	0.838±0.001	0.850±0.001	**0.859 ± 0.002**
C-dataset	0.842±0.003	0.889±0.008	0.837±0.002	0.492±0.005	0.477±0.112	0.884±0.002	0.927±0.004	**0.949 ± 0.001**
F-dataset	0.819±0.004	0.806±0.005	0.851±0.010	0.620±0.038	0.440±0.190	0.865±0.003	0.893±0.003	**0.925± 0.003**
Avg.	0.821	0.797	0.836	0.594	0.574	0.862	0.890	**0.911**
				AUPR				
B-dataset	0.389±0.006	0.256±0.014	0.416±0.003	0.234±0.006	0.516±0.042+	0.455±0.002	0.472±0.004	**0.517±0.007**
C-dataset	0.117±0.008	0.310±0.018	0.083±0.002	0.008±0.000	0.467±0.157	0.541±0.004	0.520±0.018	**0.584± 0.008**
F-dataset	0.093±0.003	0.061±0.004	0.198±0.019	0.037±0.007	0.445±0.189	0.398±0.008	0.407±0.019	**0.462± 0.001**
Avg.	0.200	0.209	0.232	0.093	0.476	0.465	0.466	**0.521**
				F1-Score				
B-dataset	0.419±0.005	0.323±0.019	0.443±0.002	0.290±0.008	0.471±0.097	0.474±0.003	0.485±0.002	**0.507±0.005**
C-dataset	0.183±0.009	0.376±0.016	0.137±0.003	0.019±0.001	0.050±0.099	0.561±0.003	0.549±0.014	**0.585± 0.004**
F-dataset	0.156±0.005	0.108±0.004	0.273±0.020	0.079±0.021	0.076±0.174	0.433±0.009	0.457±0.017	**0.482± 0.001**
Avg.	0.253	0.269	0.284	0.129	0.199	0.489	0.497	**0.524**