International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2141-2149
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0285

Drug Repurposing Using Graph Neural Networks (GNN)

Jinson Jiji[1], Linda Francis[2], Nayan Suresh[3], Vishnu Dev S[4], Muhammed Azharudeen Sahib S[5]
[1,2,3,4,5]Department of Computer Science & Engineering, College of Engineering Perumon Kollam, India.
Email ID: *jinsonjijy50@gmail.com[1], francislinda287@gmail.com[2], nayanyaju@gmail.com[3], vishnusudevan07@gmail.com[4], muhdazscep@gmail.com[5]*

Abstract
The discovery of new drugs is a time-consuming, expensive, and uncertain task. Drug repurposing is an efficient solution to this problem, which aims to find new uses for existing drugs. This paper introduces an AI-powered framework for computational drug repurposing based on Graph Neural Networks (GNNs) for predicting disease-drug associations. The biomedical knowledge graph is represented as a heterogeneous graph of diseases, genes, and drugs, which are linked by validated biological interactions. The framework leverages multimodal node representations, such as molecular embeddings for drugs and protein sequence embeddings from a pre-trained protein language model. A two-layer heterogeneous GraphSAGE model is used to aggregate information from the graph to learn biologically informative representations. Drug-disease similarity computation is performed by cosine similarity, and the model is trained with margin ranking loss and negative sampling. Experimental evalua- tion on AUROC indicates excellent performance in distinguishing actual therapeutic associations, which indicates a scalable and biologically informed approach to accelerate drug repurposing.

Keywords*: Drug Repurposing, Graph Neural Networks (GNNs), Biomedical Knowledge Graph, Disease–Drug Associa- tion Prediction, Molecular Embeddings, Protein Embeddings.*

1. Introduction
The process of identifying, validating, and approving new drugs is not only time-consuming but also very costly, with drug development taking more than a decade and costing billions of dollars, while many candidates fail in clinical trials [1], [2]. The high cost and uncertainty of drug development emphasize the need for more efficient methods to accelerate therapeutic discovery. Drug repurposing, which involves finding new therapeutic indications for existing drugs, provides a viable and cost- effective solution. Since repurposed drugs already possess known safety, toxicity, and pharmacological properties, they can be moved towards clinical application faster and with less risk [1], [3], [4]. Repurposing has been an essential component of recent large-scale research and emergency drug assessments, including the screening of COVID-19 therapeutic candidates [5]– [8]. The progress in biomedical research has led to the creation of large-scale datasets for gene-disease associations, drug-target interactions, molecular pathways and multi-omics data. These datasets are highly interconnected and complex. Con- ventional machine learning approaches tend to model samples independently, which makes them less capable of dealing with the relational and network-structured nature of biological systems. Diseases, genes, and drugs are interconnected in the form of networks, which can be better represented using graph structures [9], [10]. Graph Neural Networks (GNNs) offer a robust framework for dealing with such relational biomedical data. By mapping entities to nodes and relations to edges, GNNs can capitalize on both the attributes of nodes and the graph topology, thereby facilitating information propagation and biologically meaning- ful representation learning [11], [12]. Heterogeneous GNNs, in particular, can handle multi-relational biological interactions like disease-gene, gene-drug, and drug-drug interactions [10], [12]. To create a representation space that is

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2141-2149
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0285

biologically diverse we create a heterogeneous biomedical graph with nodes rep- resenting diseases, genes, and drugs, which are connected via disease-gene associations and gene-drug interactions. To represent more complex biological phenomena, we combine multimodal embeddings: gene nodes are represented by pro- tein sequence embeddings obtained from the pretrained ESM- 2 model, drug nodes are represented by pretrained molecular embeddings, and disease nodes have learnable embeddings. Multimodal biological embeddings have been demonstrated to enhance prediction performance in biomedical AI systems [13]–[15]. For learning task-specific representations, we use a two-layer heterogeneous GraphSAGE model, where each type of relation has its own learnable message transformation [11]. After the propagation and nonlinear transformation, we normalize the node representations to guarantee the stability of similarity computation. The similarity between drug and disease is computed by cosine similarity, and the model is optimized by a margin ranking loss to maximize the similarity of the actual therapeutic pairs compared to randomly sampled negatives. Through the integration of heterogeneous graph representation learning with pre-trained biological embeddings, the pro- posed framework offers a scalable and biologically informed solution to computational drug repurposing. The marriage of multimodal representation learning with network medicine facilitates the continued use of artificial intelligence to accelerate therapeutic discovery [10], [16], [17].

2. Related Work

Drug repurposing has been extensively investigated in the field of biomedicine research using computational, experimen- tal, and hybrid strategies [3], [4]. The initial strategies for drug repurposing were primarily based on serendipity or manual biological research [1]. Nevertheless, with the growing number of high-through put data sources, there has been a rise in the development of algorithmic methods with the ability to systematically discover new associations between drugs and diseases [5], [10]. These data sources, including genomics, proteomics, structural biology, and chemical libraries, have promoted the shift from intuition-based discoveries to data- driven algorithms that are able to process large and complex biomedicine networks [15]. Traditional computational approaches include similarity- based comparisons, guilt-by-association models, molecular docking, and network-based inference [9], [10]. Similarity- based approaches compare the chemical or biological simi- larity of drugs, which assumes that similar drugs are used to treat similar diseases. Although useful to a certain extent, these approaches ignore the larger biological context and tend to oversimplify the complex mechanisms of drug action [9]. Network-based approaches utilize protein-protein interactions, gene co-expression, and drug-target interaction networks to predict potential candidates for drug repurposing [10], [18]. Although these approaches introduce a degree of biological plausibility, their predictions remain largely reliant on the patterns of available relationships and tend to perform poorly on noisy, sparse, or incomplete datasets, which are often the case in biomedical knowledge graphs [19]. With the advent of deep learning techniques, Graph Neural Networks have recently appeared as a promising approach for modeling complex biomedical relationships [11], [12]. GCNs, GATs, and RGCNs facilitate message passing between heterogeneous biological entities, thereby modeling multi-hop relationships such as disease-gene-drug pipelines [6], [12]. Earlier systems such as DMSR have shown that structural drug features improve predictive performance [6]. Nevertheless, most of these models rely on unimodal or shallow features, such as one-hot vectors or simple molecular fingerprints, which lack biological expressiveness and interpretability [16]. These drawbacks become more pronounced when a model is faced with “cold-start” nodes, which are entities with very few or no known connections [11]. DRAGON moves the state of the art forward by proposing a multi-modal feature integration approach [17]. Contrary to previous approaches, DRAGON incorporates both the chemi- cal (drug) and biological (protein) modalities into a heteroge- neous graph, which better represents the complex biochemical interactions [17]. This immediately remedies all the major shortcomings pointed out in previous approaches: the lack of feature expressiveness, the insufficient treatment of sparse biological graphs, and

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2141-2149
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0285

the heavy dependence on the graph structure without resorting to powerful pre-trained embeddings [16]. By integrating large-scale protein language models with molecular embeddings, DRAGON locates itself at the cross- roads of GNNs evolution and multimodal biomedical data fusion, providing a more robust, biologically interpretable, and scalable solution for computational drug repurposing [17].

3. Research Gap

Although Graph Neural Networks (GNNs) have been demonstrated to be promising in the biomedical link prediction and drug repurposing tasks, there are still some shortcomings in the current methods. Most of the existing methods have been focusing on the graph structure with limited or manu- ally designed node features. Although the graph structure is informative, it is not sufficient to represent the biochemical and functional aspects of drugs and proteins in the therapeu- tic relationship. Moreover, some existing models have been employing homogeneous graph structures or oversimplified relation modeling for biomedical entities. These models might not be able to distinguish between different biological inter- actions, such as drug targeting and inhibition or activation. Another important limitation is the lack of proper integration of pretrained biological features. Recent breakthroughs in pro- tein language models and molecular embeddings offer strong biochemical prior knowledge, but most graph-based drug repo- sitioning models fail to properly integrate these multimodal cues into heterogeneous graph learning. Thus, there is a need for a relation-aware heterogeneous GNN model that properly integrates pretrained molecular and protein embeddings while maintaining the structural dependencies between diseases, genes, and drugs. This can help in deriving more biologically valid features and better predictive models in computational drug repositioning.

4. Methodology

The methodology of the proposed framework is built upon constructing a heterogeneous biomedical graph, integrating multimodal node embeddings, designing a heterogeneous Graph Neural Network architecture, and training the model using similarity-based ranking techniques.

4.1.Dataset Sources and Preprocessing

The heterogeneous biomedical graph adopted by this re- search is created by incorporating various publicly available databases. The main datasets for graph construction are listed in Table 1

Table 1 Biomedical Data Sources Used for Graph Construction

Database	Data Type	Purpose
DisGeNET	Disease–Gene Associations	Identify genes related to diseases
DrugBank	Drug–Gene Interactions	Drug target information
STRING	Protein–Protein Interactions	Gene functional relationships
KEGG	Biological Pathways	Pathway validation

Prior to graph construction, several preprocessing steps were applied to ensure data consistency across the integrated datasets. The preprocessing pipeline is summarized in Table 2.

Table 2 Data Preprocessing Steps

Step	Description
Gene Standardization	Gene identifiers converted to offi- cial gene symbols
Duplicate Removal	Duplicate interactions removed across datasets
Data Cleaning	Incomplete or inconsistent records removed
Disease Normalization	Disease names standardized across sources
PPI Filtering	Low-confidence protein interac- tions removed
Node Mapping	Entities mapped to unique graph node identifiers

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2141-2149
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0285

4.2.Graph Construction

Trusted sources data connects diseases to genes, as well as drugs to genes within the model, there are three types of points: for diseases, genes, and drugs. The connections between the points represent the relation between diseases and genes, drugs and genes,and how they can affect gene functions, such as blocking or turning on The paths formed by these connections represent how a disease can be treated through common genetic links Shown in Table 3 & 4.

Table 3 Heterogeneous Graph Statistics

Category	Count
Diseases	84
Genes	1257
Drugs	888
Disease–Gene edges	1685
Gene–Drug edges (targets)	1191
Gene–Drug edges (inhibits)	1031
Gene–Drug edges (activates)	279
Total edges	**4186**

4.3.Feature Engineering

The features of the nodes are enhanced by using embed- dings that are pretrain. The embeddings for the drugs (644- dimensional) are obtained from the molecular representation. The embeddings for the proteins (644-dimensional) are ob- tained from the ESM-2 pretrained protein language model, which captures the structural and functional information of the amino acid sequences. The disease nodes are initialized with learnable disease embeddings (644-dimensional).

4.4.Graph Neural Network

The model uses a two-layer heterogeneous GraphSAGE encoder implemented with PyTorch Geometric's HeteroConv framework. There is an SAGEConv layer for each relation type (disease to gene association, gene to drug targeting, inhibition, and activation), which allows for relation-specific message passing. ReLU activation and dropout are used in between layers to enhance non-linearity and generalization. The final embeddings are L2-normalized to enable computation of co- sine similarity.

4.5.Link Prediction

Drug–disease similarity is computed using cosine similarity between their final embedding's.

$$S(d, r) = \frac{\mathbf{h}_d^{\top}\mathbf{h}_r}{\|\mathbf{h}_d\|_2 \|\mathbf{h}_r\|_2}$$

where $\mathbf{h_d}$ and $\mathbf{h_r}$ represent the final normalized embeddings of disease d and drug r, respectively. A margin ranking loss function is applied to enforce higher similarity for known disease–drug pairs compared to randomly sampled negative pairs. This ranking-based objective enables effective drug repurposing prediction.

4.6.Training Protocol

The model is trained using mini batch sampling of positive disease–drug pairs combined with negative sampling. The training objective is defined as:

$$\mathbf{L = \max 0,\ m - S\ (d,\ r^{+}) + S\ (d,\ r^{-})}$$

where r^{+} denotes a positive drug, r^{-} denotes a randomly sampled negative drug, and m is the margin hyper parameter (set to 0.5 in our implementation). Optimization is performed using the Adam optimizer. During each epoch, the model updates relation-specific trans- formation weights to improve ranking performance between diseases and drugs.

Table 4 Heterogeneous Graphsage Model Configuration

Parameter	Value
Input Feature Dimension	644
Hidden Dimension	256
Number of GNN Layers	2
Convolution Type	Relation-specific SAGEConv
Aggregation Function	Mean
Activation Function	ReLU
Dropout Rate	0.3
Embedding Normalization	L2 (dim=1)

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2141-2149
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0285

Similarity Function	Cosine Similarity
Loss Function	Margin Ranking Loss
Margin (m)	0.5
Optimizer	Adam
Learning Rate	0.001
Epochs	100
Batch Size	512

5. Results and Discussion

5.1.Model Performance

The performance of the proposed framework is evaluated using the Area Under the Receiver Operating Characteristic Curve (AUROC), Precision–Recall Area Under Curve (PR- AUC), and ranking-based metrics such as Hits@K and Precision@K. AUROC measures the ability of the model to rank true disease–drug associations higher than randomly generated negative pairs, while PR-AUC provides a more reliable assess- ment under class imbalance conditions. Table 5 summarizes the quantitative evaluation results of the model.

Table 5 Performance Evaluation of the Proposed Heterogeneous GNN Model

Metric	Value
AUROC	0.884
PR-AUC	**0.917**
Hits@20	20
Precision@20	**1.000**
Recall@20	0.256

Evaluation conducted on a heterogeneous drug–disease dataset containing 888 drugs. The dataset was split into 310 training pairs and 78 testing pairs using an 80/20 train–test split. Ranking metrics (Hits@20, Precision@20, and Re- call@20) measure the ability of the model to identify correct drug candidates within the top predicted results. The heterogeneous biological relationships between dis- eases, genes, and drugs are integrated into the proposed framework. Gene nodes are initialized using 644-dimensional feature vectors derived from pretrained protein embeddings generated using the ESM-2 model, which capture rich structural and evolutionary information about proteins . Drug node features are represented using 644-dimensional molec- ular embeddings derived from the SMILES representation of drugs, capturing the chemical and structural properties of the molecules. Disease nodes are represented using learned disease embeddings that encode associations between diseases and their related genes in the heterogeneous biomedical network [15]. These feature representations allow the model to capture complex biological interactions among diseases, genes, and drugs for effective drug repurposing prediction [6], [16].

5.2.Comparison with Existing Drug Repurposing Methods

To evaluate the effectiveness of the proposed heteroge- neous graph neural network framework, its performance was compared with several existing drug repurposing approaches reported in the literature, including recommender-based meth- ods, graph-based integration models, and deep learning frame- works. Recommender-based approaches infer drug–disease associations using collaborative filtering and similarity-based techniques, while graph-based integration models combine large-scale drug–protein interaction networks to identify po- tential therapeutic candidates [10], [18]. Deep learning ap- proaches have also been applied to drug repurposing tasks by learning complex nonlinear relationships in biomedical data and integrating heterogeneous biological information [5], [19]. Among recent graph neural network approaches, the DRAGON framework integrates drug and protein embeddings within a GNN architecture for predicting drug–disease associations [17]. In contrast, the proposed framework incorporates heterogeneous biological relationships among diseases, genes, and drugs, enabling richer representation learning within the biomedical network. Table 6 presents a comparison of the proposed model with existing methods based on AUROC and PR-AUC metrics. Although the DRAGON framework presents slightly higher PR-AUC results, the proposed heterogeneous GNN presents competitive results with better AUROC results. Unlike the DRAGON framework, which is based only on drug-protein interactions, the proposed model considers various hetero- geneous biological interactions, such as disease-gene, gene-

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2141-2149
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0285

drug, and protein-protein interactions, for better representation learning of the drug-disease prediction problem.

Table 6 Comparison with Existing Drug Repurposing Methods

Model	AUROC	PR-AUC
RCDR	0.78	0.73
Graph-based Drug–Protein Model	0.81	0.76
Deep Learning Drug Repurposing	0.83	0.79
DRAGON	0.87	0.94
Proposed Heterogeneous GNN	**0.88**	**0.91**

5.3.Case Study: Breast Cancer Drug Repurposing

To demonstrate the practical applicability of the proposed framework, a case study was conducted for breast cancer drug repurposing. Table 7 lists the top six predicted drug candidates ranked by cosine similarity between disease and drug embeddings.

Table 7 Top-6 Predicted Drug Candidates for Breast Cancer

Rank	Drug Name	DrugBank ID	Score
1	Estradiol	DB00783	0.4071
2	Paclitaxel	DB01229	0.3593
3	Procarbazine	DB01168	0.3030
4	Amantadine	DB00915	0.3015
5	Dronabinol	DB00470	0.2962
6	Disulfiram	DB00822	0.2808

Paclitaxel is an effective chemotherapeutic drug used in the treatment of breast cancer and has shown significant improvement in clinical outcomes in breast cancer patients [20]. Estradiol plays a crucial role in the progression of breast cancer mediated through the estrogen receptor and is strongly associated with breast cancer carcinogenesis [21]. Repurposed drugs such as Disulfiram have demonstrated anti-tumor activ- ity in breast cancer by targeting aldehyde dehydrogenase and cancer stem cells [22]. Similarly, cannabinoids such as Dronabinol have shown anti-proliferative effects in breast cancer cells through modulation of cannabinoid signaling pathways [23]. Chemothera- peutic agents such as Procarbazine have also been investigated in combination therapies for breast cancer treatment [24]. Additionally, antiviral drugs such as Amantadine have been explored in drug repurposing studies for potential anticancer applications.

6. System Architecture

The architecture of the proposed framework is illustrated in Figure 1, which depicts the complete workflow from raw biomedical data acquisition to drug repurposing prediction and evaluation.

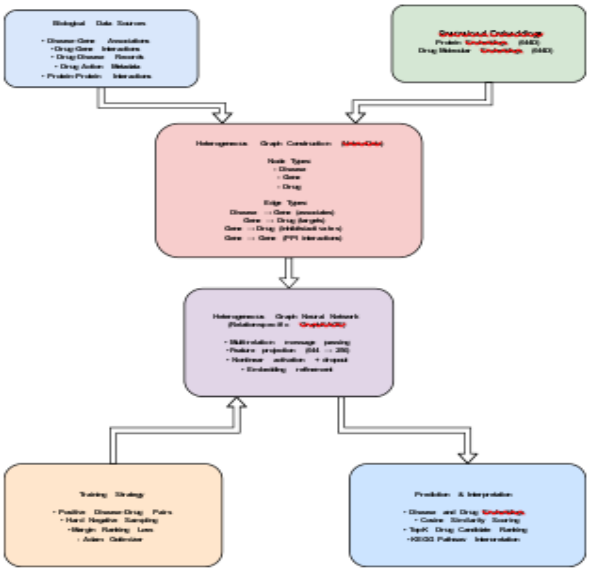

Figure 1 System Architecture of the Proposed Heterogeneous Drug Repurposing Framework

The process begins with the collection of curated biomedical data, which includes information about disease–gene associa- tions, drug–gene interactions, and validated drug data [3], [4], [10]. The data is then preprocessed to ensure that it is consis- tent and biologically meaningful. This includes standardizing diseases, converting gene information into numerical form, and filtering interactions that are specific to humans [10]. To further enhance the capabilities of representation, learning, pretrained embeddings are integrated into the system. Gene nodes are initialized with protein embeddings of di- mensionality 644, which are produced by the ESM-2 model, while drug

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2141-2149
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0285

nodes are initialized with molecular embeddings of dimensionality 644, which are derived from the SMILES representation of each drug [5], [6]. These embeddings capture biological and chemical information that goes beyond what can be represented through graph structures alone. With these processed datasets, we construct a heterogeneous biomedical graph that contains three types of nodes: diseases, genes, and drugs. There are various types of relationships between these nodes, including disease–gene associations and gene–drug relationships such as targeting, inhibition, and activation [10], [19]. Such a multi-relational graph structure is able to represent the intricate therapeutic processes in biological systems without reducing the data to a uniform graph structure [15]. For the analysis of this graph structure, we propose a two-layer heterogeneous graph neural network with relation- specific GraphSAGE convolutions [11]. Each type of relation- ship has its own learnable transformation function to enable information mixing based on biological relevance. Node fea- tures are mapped to a common latent space with 256 features. After that, features are passed through nonlinear activation and dropout to improve generalization. Node embeddings are then normalized using L2 normalization to ensure stability in cosine similarity calculations between disease and drug embeddings [6]. During training, true disease–drug pairs are obtained from known therapeutic relationships, while negative pairs are generated by pairing diseases with drugs that are not known to be associated with them. The model is trained using a margin ranking loss function that encourages the score of a positive disease–drug pair to be higher than that of a negative pair. After training, the model computes similarity scores between diseases and candidate drugs and ranks potential therapeutic associations, enabling the prediction of the Top-K drug candidates for a given disease [6], [10]. After predictions are generated by the heterogeneous graph- based framework, a post-inference pathway interpretation step is performed. For a given disease, the genes associated with that disease are identified from the heterogeneous graph. These genes are then mapped to biological pathways using pathway databases such as KEGG. By analyzing the relationships between disease-associated genes and biological pathways, the framework identifies the most relevant pathways involved in the mechanisms of the disease. This step provides addi- tional biological validation for the predicted drug candidates by linking them to disease-related pathways and molecular mechanisms [15], [19].

Conclusion

This study offers a heterogeneous Graph Neural Network framework that aims to overcome the shortcomings of con- ventional drug repurposing strategies. The proposed model combines molecular drug embeddings and protein sequence embeddings into a single graph representation, which enriches the feature space used for representation learning and allows for biologically meaningful and interpretable embeddings.The multimodal integration of molecular drug embeddings and pro- tein sequence embeddings helps to overcome the shortcomings of topology-only approaches, which tend to lack relational in- formation and biochemical context.The proposed model shows excellent discriminative capability when tested using AUROC, indicating the model's efficiency in ranking positive disease- drug therapeutic pairs above randomly selected negatives. The heterogeneous graph representation and the use of pre- trained biological embeddings improve predictive performance by jointly modeling interaction patterns and molecular infor- mation. The learned disease and drug embeddings encode bi-ologically meaningful relationships that correspond to known therapeutic mechanisms.This framework offers a significant improvement in computational drug repurposing strategies by leveraging heterogeneous biomedical data and powerful graph learning methods. The model architecture is highly scalable and can be extended to include other biomedical modalities in future studies. In general, this framework enables more informed decision-making in the prioritization of repurposable drug candidates, leading to more efficient and biologically sound drug discovery strategies.

Proposed Solutions

Although the proposed heterogeneous Graph Neural Net- work framework demonstrates effective ranking performance, several directions remain for future improvement. First, in- corporating additional

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2141-2149
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0285

biomedical modalities such as gene expression data, pathway activity information, or adverse effect profiles could further enrich node representations. These complementary data sources may provide deeper biological context beyond molecular and protein sequence embeddings. Second, exploring alternative heterogeneous GNN architectures may improve the modeling of complex multi-relational interactions. Attention-based message passing or transformer- inspired graph models could enable adaptive weighting of different biological relations, potentially enhancing representation quality. Third, extending the framework to account for temporal dynamics represents a promising research direction. Since biomedical knowledge evolves continuously, integrating time- aware modeling approaches may allow the system to better re- flect emerging therapeutic evidence and updated disease–drug associations. Finally, improving interpret ability and validating top- ranked predictions remain important next steps. Developing visualization techniques for embedding similarity and rela- tional pathways could increase transparency, while biological or literature-based validation of predicted candidates would strengthen confidence in the model's practical utility.

References

[1]. T. T. Ashburn and K. B. Thor, "Drug repositioning: identifying and de- veloping new uses for existing drugs," Nature Reviews Drug Discovery, vol. 3, pp. 673–683, 2004.

[2]. T. S. Andrews and M. Hemberg, "Identifying cell populations with scrnaseq," IEEE Transactions on Molecular Biology and Medicine, vol. 59, pp. 114–122, 2018.

[3]. S. Pushpakom, F. Iorio, P. E. Eyers, K. Escott, S. Hopper, A. Wells, Doig, P. Guilliams, T. Latimer, C. McNamee, G. C. Norris, R. Sanseau, M. Cavalla, and A. P. Pirmohamed, "Drug repurposing: progress, challenges and recommendations," Nature Reviews Drug Discovery, vol. 18, pp. 41–58, 2019.

[4]. V. Parvathaneni, N. S. Kulkarni, A. Muth, and V. Gupta, "Drug repur- posing: a promising tool to accelerate the drug discovery process," Drug Discovery Today, vol. 24, no. 10, pp. 2076–2085, 2019.

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

[6]. M. Zitnik, M. Agrawal, and J. Leskovec, "Modeling polypharmacy side effects with graph convolutional networks," Bioinformatics, vol. 34, no. 13, pp. i457–i466, 2018.

[7]. S. Kaliamurthi, G. Selvaraj, C. Selvaraj, S. K. Singh, D. Q. Wei, and G. H. Peslherbe, "Structure-based virtual screening reveals ibrutinib and zanubrutinib as potential repurposed drugs against covid-19," Interna- tional Journal of Molecular Sciences, vol. 22, no. 13, p. 7071, 2021.

[8]. A. Khan, S. S. Ali, M. T. Khan, S. Saleem, A. Ali, M. Suleman, Z. Babar, A. Shafiq, M. Khan, and D. Q. Wei, "Combined drug repurposing and virtual screening strategies with molecular dynamics simulation identi- fied potent inhibitors for sars-cov-2 main protease (3clpro)," Journal of Biomolecular Structure and Dynamics, vol. 39, no. 13, pp. 4659–4670, 2021.

[9]. E. Guney, J. Menche, M. Vidal, and A.-L. Barabasi, "Network-based in silico drug efficacy screening," Nature Communications, vol. 7, p. 10331, 2016.

[10]. F. Cheng, R. J. Desai, D. E. Handy, R. Wang, S. Schneeweiss, and J. Baraba´si, "Network-based approach to prediction and population-based validation of in silico drug repurposing," Nature Communications, vol. 9, no. 1, pp. 1–12, 2018.

[11]. W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 1024–1034.

[12]. M. Schlichtkrull, T. N. Kipf, P. Bloem, R. van den Berg, I. Titov, and M. Welling, "Modeling relational data with graph convolutional networks," arXiv preprint arXiv:1703.06103, 2017.

[13]. P. Paplomatas, M. G. Krokidis, P. Vlamos, and

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 04 April 2026
Page No: 2141-2149
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0285

A. G. Vrahatis, "An ensemble feature selection approach for analysis and modeling of transcriptome data in alzheimer's disease," Applied Sciences, vol. 13, no. 4, p. 2353, 2023.

[14]. H. T. N. Tran, K. S. Ang, M. Chevrier, X. Zhang, N. Y. S. Lee, M. Goh, and J. Chen, "A benchmark of batch-effect correction methods for single- cell rna sequencing data," Genome Biology, vol. 21, pp. 1–32, 2020.

[15]. A.-L. Baraba´si, N. Gulbahce, and J. Loscalzo, "Network medicine: a network-based approach to human disease," Nature Reviews Genetics, vol. 12, pp. 56–68, 2011.

[16]. A. Ayuso-Mun˜oz and R. Artin˜ano-Mun˜oz, "Artificial intelligence in drug repurposing," Artificial Intelligence in Medicine, vol. 138, p. 102522, 2023.

[17]. R. Artin˜ano-Mun˜oz, A. Ayuso-Mun˜oz, and M. Smith, "Dragon: Drug re- purposing via graph neural networks," in IEEE International Conference on Biomedical and Health Informatics (CBMS), 2024.

[18]. Y. Zhou, Y. Hou, J. Shen, Y. Huang, W. Martin, and F. Cheng, "Network- based drug repurposing for novel coronavirus 2019-ncov/sars-cov-2," Cell Discovery, vol. 6, no. 1, pp. 1–18, 2020.

[19]. D. M. Gysi, I. Do Valle, M. Zitnik, A. Ameli, X. Gan, O. Varol, S. D. Ghiassian, J. J. Patten, R. A. Dave, J. Loscalzo, and A.-L. Baraba´si, "Network medicine framework for identifying drug-repurposing oppor- tunities for covid-19," Proceedings of the National Academy of Sciences, vol. 118, no. 19, 2021.

[20]. I. C. Henderson, D. A. Berry, and G. D. Demetri, "Improved outcomes from adding paclitaxel to adjuvant therapy for breast cancer," Journal of Clinical Oncology, vol. 21, no. 6, pp. 976–983, 2003.

[21]. J. D. Yager and N. E. Davidson, "Estrogen carcinogenesis in breast cancer," New England Journal of Medicine, vol. 354, no. 3, pp. 270–282, 2006.

[22]. Z. Skrott, M. Mistrik, and K. K. Andersen, "Alcohol-abuse drug disul- firam targets cancer via npl4 inhibition," Nature, vol. 552, pp. 194–199, 2017.

[23]. M. M. Caffarel, C. Andradas, and E. Perez-Gomez, "Cannabinoids reduce erbb2-driven breast cancer progression," Molecular Cancer, vol. 9, p. 196, 2010.

[24]. V. T. DeVita, A. A. Serpick, and P. P. Carbone, "Combination chemother- apy in the treatment of advanced breast cancer," Cancer, vol. 35, no. 1, pp. 98–107, 1975.