SEEJPH Volume XXVI, S2,2025, ISSN: 2197-5248; Posted:03-02-25

Gated Graph Sequence Neural Network-Based Prediction of Drug-Gene Association for Nucleoside Proteins in Oral Cancer

Dr. Soundharya Manogaran [1], Dr. Subasree S [2], Dr. Ramya Ramadoss [3*], Dr. Pradeep Kumar Yadalam [4]

[1] Department of Oral Biology, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India. Email: 152432001.sdc@saveetha.com
[2] Assistant Professor, Department of Periodontics, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India. Email: subasrees.sdc@saveetha.com
[3] Professor, Department of Oral Pathology & Oral Biology, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India. Email: ramyar.sdc@saveetha.com
[4] Professor, Department of Periodontology, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India. Email: pradeepkumar.sdc@saveetha.com
**Corresponding Author: Dr. Ramya Ramadoss, B.D.S., M.D.S, PhD*

KEYWORDS
Oral cancer, drug-gene associations, nucleoside proteins, Gated Graph Sequence Neural Networks (GGSNN), computational drug discovery, precision medicine, machine learning.

ABSTRACT

Background: Oral cancer poses a significant public health challenge due to its increasing incidence, late-stage diagnosis, and resistance to treatment. Nucleoside metabolism plays a vital role in oral cancer biology, influencing DNA synthesis, repair, and cellular proliferation. Drug-gene associations can help identify targeted therapies by analyzing the interactions between nucleoside proteins and drugs. Graph Neural Networks (GNNs), particularly Gated Graph Sequence Neural Networks (GGSNNs), offer a promising approach to model these complex interactions.

Aim: This study aims to predict drug-gene associations for nucleoside proteins in oral cancer using a GGSNN model, with the goal of identifying potential therapeutic targets and improving drug discovery strategies.

Methodology: A dataset comprising drug-gene interaction data was preprocessed to remove missing values and normalize features. Graph data frames were constructed to represent nodes (drugs/genes) and edges (interactions). A GGSNN architecture with two layers was implemented using the GatedGraphConv layer for message passing. The model was trained for 100 epochs using an 80:20 train-test split, with performance evaluated using metrics such as Mean Absolute Error (MAE), R-squared (R^2), precision-recall curves, and F1 scores. Visualization tools such as Cytoscape and dimensionality reduction techniques were used for analysis.

Results: The GGSNN model achieved an average precision score of 76.94%, an F1 score of 72.42%, and a recall rate of 85.95%. The MAE was 0.0685, indicating low prediction error, while the R^2 value of 26.66% highlighted moderate explanatory power. Visualization techniques revealed insights into drug-gene interactions and model learning patterns. The precision-recall curve indicated robust performance across different recall values, with a balanced threshold of -0.0147 optimizing precision and recall.

Conclusion: The GGSNN model demonstrates strong predictive capabilities in identifying drug-gene associations in oral cancer, providing a valuable tool for computational drug discovery. While the model achieves notable precision and recall, its moderate R^2 value suggests areas for improvement. Future work should incorporate additional biological features and validation datasets to enhance model robustness and applicability in precision medicine.

1. Introduction

Oral cancer represents a significant public health challenge with increasing incidence and mortality rates worldwide, particularly in regions with high tobacco and alcohol consumption (1). This malignancy is often diagnosed at advanced stages, leading to poor prognosis and limited treatment options (2). The pathogenesis of oral cancer is multifaceted, involving genetic mutations, environmental exposures, and viral infections, all contributing to its complexity (3). Despite advancements in understanding oral cancer biology, effective therapeutic strategies remain elusive, necessitating novel approaches to address the disease's late-stage presentation, treatment resistance, and molecular heterogeneity (4).

Nucleoside metabolism, including oral cancer, is crucial in cancer biology, as it regulates essential cellular processes like DNA synthesis, repair, and cell proliferation (5). Nucleoside transporters and associated proteins facilitate the uptake and utilization of nucleosides, which are vital for maintaining nucleotide pools necessary for cell growth (6). Aberrations in these processes have been linked to tumorigenesis and cancer cell adaptation to their microenvironment (7), highlighting their significance as potential therapeutic targets (8). Nucleoside analog drugs, which mimic natural nucleosides, have been effective in treating various cancers by disrupting DNA replication and repair mechanisms (9), with the potential to uncover targeted therapies for oral cancer (10). Key components in cancer treatment of nucleoside inhibitors include MOC-2, nucleosides, 5-FU, 6-mercaptopurine, 6-thioguanine, Azacytidine, Decitabine, Cladribine, Cytarabine, Fludarabine, and Gemcitabine. Genetic markers influence treatment outcomes and side effects, with genetic markers affecting treatment outcomes in various cancers.

Drug-gene association is the relationship between specific genes or proteins and drugs' therapeutic effects or mechanisms. Understanding these associations is crucial for identifying molecular targets and elucidating drug interactions. In oral cancer, drug-gene associations highlight the roles of nucleoside proteins in cancer progression and suggest potential interventions with nucleoside analog drugs. Accurately predicting these associations enables researchers to identify promising drug candidates, repurpose existing drugs, and develop tailored treatment strategies based on the unique molecular profiles of diseases. Genomic data and pharmacology studies are crucial for personalized medicine and improving treatment outcomes.

One study highlights the importance of understanding drug-human gene interactions in drug efficacy and genomics. It introduces a representation learning approach using metapath2vec (11) and metapath2vec++ models on adverse drug reaction data. One recent study showed that a new method called SCMFDD (12) uses biological context to predict drug-disease associations. It outperforms existing techniques and has a user-friendly web server. Predicting drug-gene associations (13) is a complex endeavor requiring extensive biological data analysis (16). Given that traditional experimental methods can be time-consuming and resource-intensive, computational approaches are necessary to efficiently process and interpret these datasets (17). Machine learning models (18) have emerged as invaluable tools in this context, capable of uncovering hidden patterns and relationships in the data. Within this framework, graph-based models stand out for their ability to represent molecular interactions as networks. These models can comprehensively understand the molecular landscape by integrating gene expression profiles, protein-protein interactions, and drug-target relationships (19). The accurate prediction of drug-gene associations not only facilitates drug discovery but also aids in identifying biomarkers for early diagnosis and prognosis of oral cancer (20).

Graph Neural Networks (GNNs), particularly Gated Graph Sequence Neural Networks (GGSNNs), have emerged as powerful frameworks for modeling complex biological systems (21). GGSNNs leverage the strengths of graph-based data representation to capture the intricate relationships between drugs, genes, and proteins (23). These models integrate graph structures with sequential data, enabling the dynamic representation of molecular interactions over time (22). The choice of GGSNNs in this study is driven by their ability to handle heterogeneous and interconnected data, which is essential for understanding the multifactorial nature of oral cancer (24). By applying GGSNNs to predict drug-gene associations for nucleoside proteins, this study seeks to uncover critical molecular insights and identify potential therapeutic targets. Integrating GGSNNs in drug-gene prediction represents a significant step toward advancing precision medicine and improving outcomes for patients with oral cancer.

2. Methodology

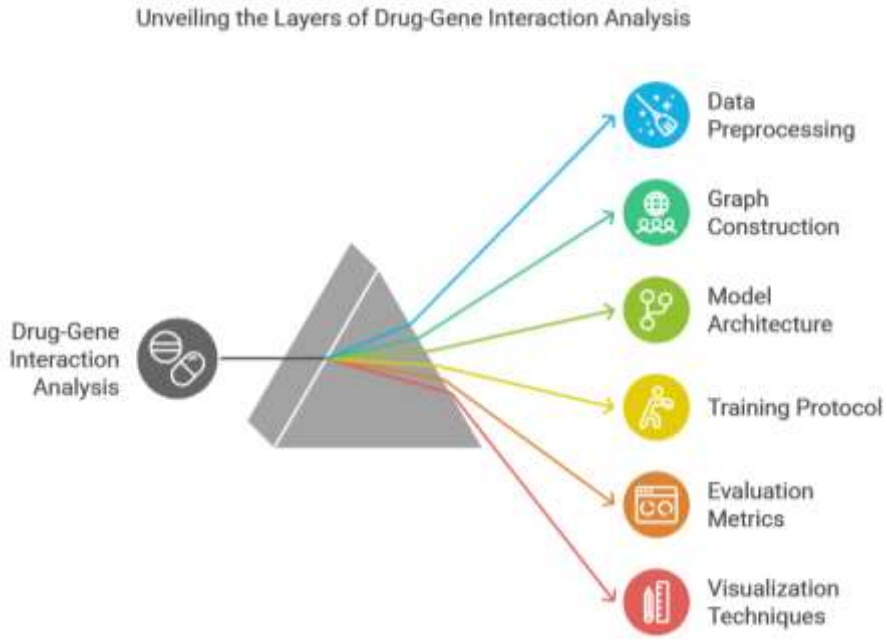

Figure 1 Workflow of Drug Gene Interaction Analysis

Using a probe and drug site,(25) nucleoside drugs and genes were retrieved and normalized, and the missing value was removed. Graph data frame consists of "name" as a node, "gene_name" as a node, "target_type" as edge, "activity_biochemical" as edge weight, and other columns as node features. Data were split into eighty percent and 20 percent test data and subjected to graph neural network architecture(fig-1).

Cytoscape:

Cytoscape is a bioinformatics software platform that visualizes biological networks, analyzes complex biological data, and creates an interactome of drug genes. Using Cytoscape, data were imported and sourced, target nodes were assigned for drugs and genes, and interaction edges were assigned to the protein interactions. The interactome of drug-gene associations was visualized and analyzed for network properties.

Graph Construction:

The data frame was normalized, and missing values and duplicates were removed. The data were split into train and test data. The graph structure enabled the representation of drug-gene interactions in a format suitable for GNN processing.

Model Architecture:

A Gated Graph Neural Network (GGNN) was implemented with two layers. The model consisted of a GatedGraphConv layer for message passing and a fully connected layer for output. The input and output channels were set to 64, matching the node feature dimensions. The model was trained using the Adam optimizer with a learning rate 0.001, and the loss was calculated using Mean Squared Error (MSE). Gradient clipping with a maximum norm of 1.0 was applied to ensure stable training(fig-2).

Training Protocol:

The model was trained for 100 epochs with a batch processing approach, using the entire graph for each forward pass. The training process involved forward propagation through the GGNN, loss calculation, backward propagation, and parameter updates. Loss values were tracked across epochs to monitor training progress, and a seed value of 42 was used to ensure reproducibility.

Evaluation Metrics:

The model's performance was assessed using metrics like MAE, R^2, precision-recall curves, F1 scores, and average precision score, with threshold analysis determining the optimal trade-off between precision and recall.

Visualization Techniques:

The study used PyTorch for deep learning, PyTorch Geometric for graph operations, Scikit-learn for metrics, and Matplotlib/Seaborn for visualizations. Data quality was ensured through value handling, edge weight normalization, and feature scaling. Statistical analysis revealed predictive capabilities and areas for improvement. This robust approach demonstrates reproducibility, scalability, and interpretability in computational drug discovery.

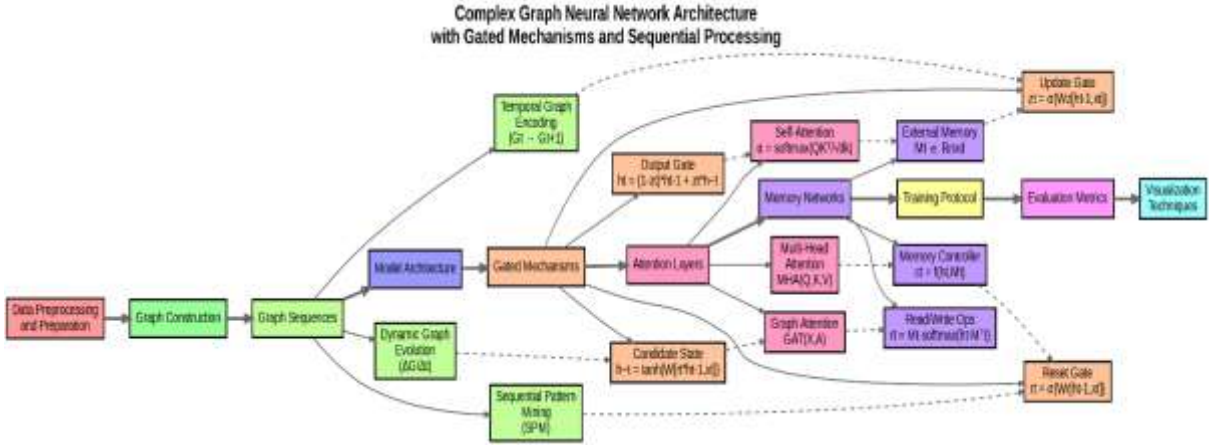

Figure 2 shows the architecture of gated graph neural networks.

3. Results

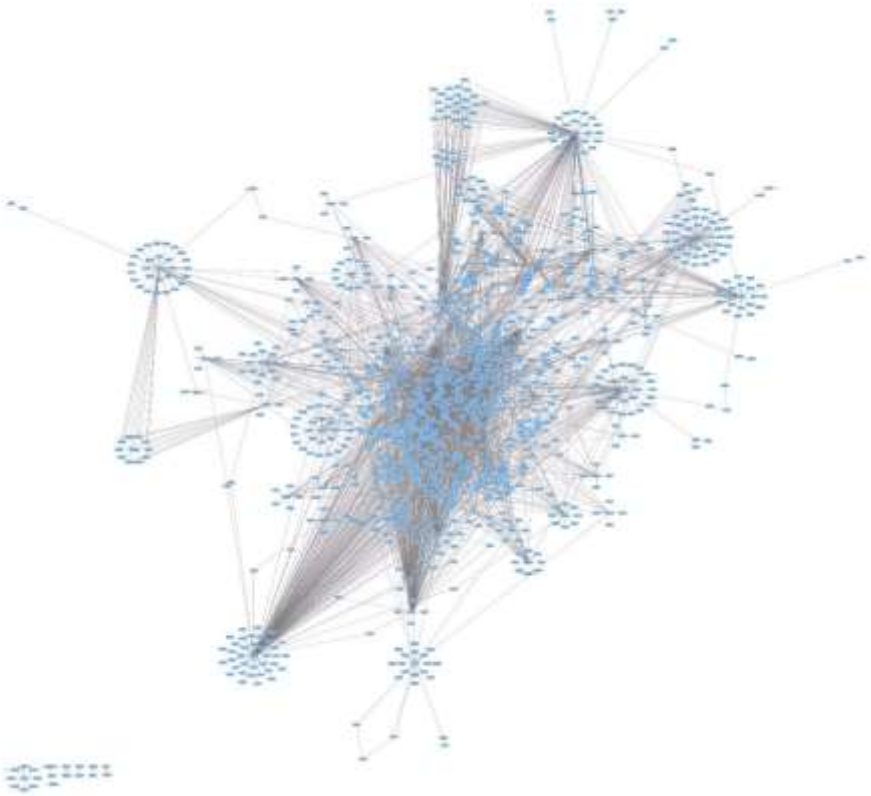

Figure 3 shows the interactome of the drugs and genes associated with nucleoside receptors.

The network comprises 1,124 nodes and 3,201 edges, resulting in an average of approximately 5.428 neighbors per node. The network exhibits a diameter of 10 and a radius of 5, with a characteristic path length of 3.958, indicating the average distance between nodes. Notably, the clustering coefficient is low at 0.000, which suggests limited local interconnectivity among nodes. The overall network density is 0.005, reflecting its sparse structure. The network's heterogeneity is measured at 2.368, while the centralization is relatively moderate at 0.179. Additionally, the network consists of 7 connected components, and the analysis was completed efficiently in just 0.372 seconds.

The study presents a novel approach to predicting drug-gene interactions using a GGNN architecture. The model's accuracy is measured by its average precision score of 76.94%, mean absolute error of 0.0685, R-squared of 26.66%, and best F1 score of 72.42%. The model's optimal performance at the threshold of -0.0147 is achieved with a high recall of 85.95%, capturing 85.95% of actual drug-gene interactions. The training performance is stable, with a consistent decrease in the loss curve and a well-tuned learning rate. The study reveals a classification model with an optimal threshold of -0.0147, balancing precision and recall. Its precision is 62.57%, recall is 85.95%, and F1 score is 72.42%.

The model has a low Mean Absolute Error (MAE) of 0.0685, indicating good model performance with small average errors. It also has a moderate R-squared (R^2) of 0.2666, indicating that the model can explain 26.66% of the data's variance. The model's predictive capabilities need improvement, possibly through refining, incorporating additional features, or exploring different techniques to capture patterns better and identify improvement areas.

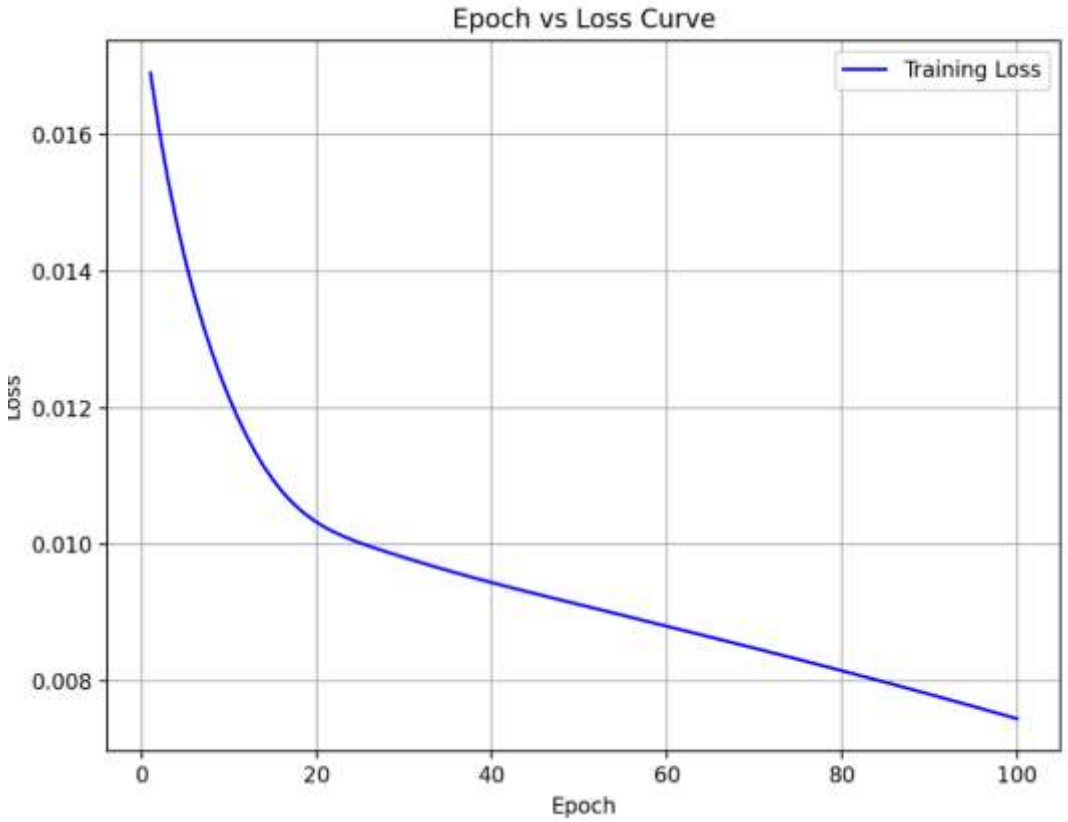

Figure 4 shows the "Epoch vs Loss Curve" graph demonstrating the relationship between training epochs and loss value, indicating successful model training.

The graph shows a decreasing trend, convergence towards the end of epochs, and lower loss values, indicating better model performance. The GGNN model successfully learns from graph structure, capturing drug and gene relationships.

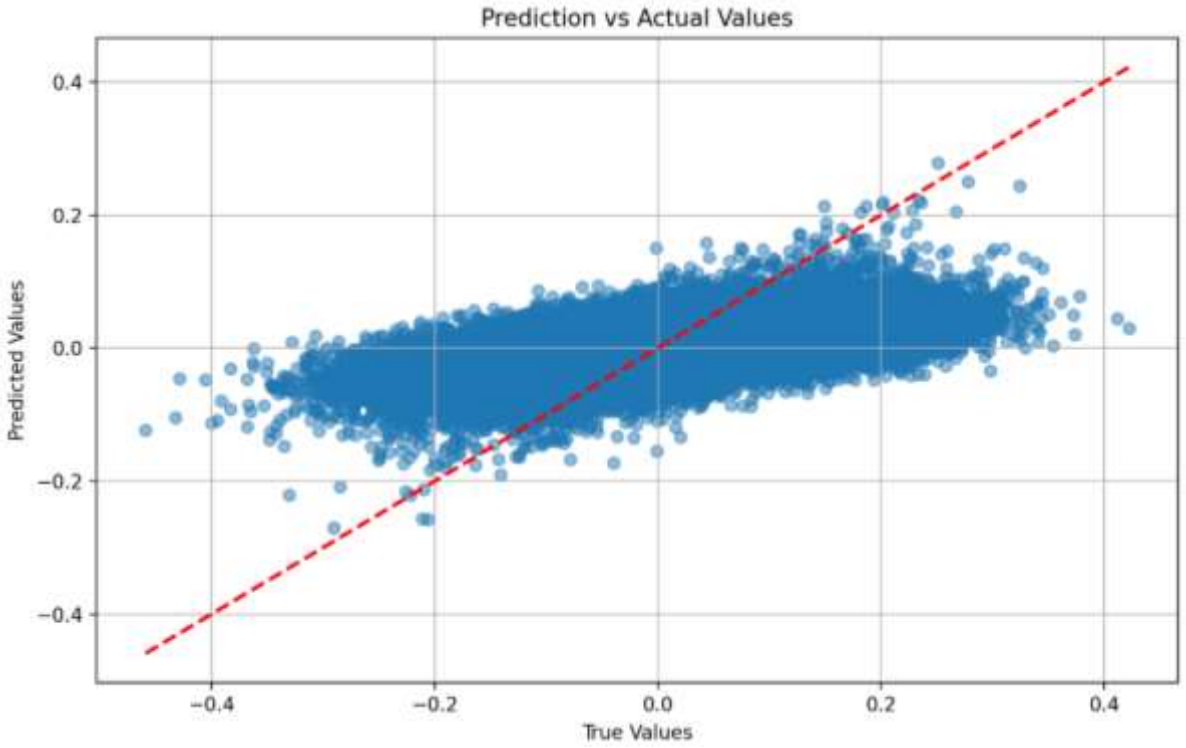

Figure 5 shows a plot that illustrates the accuracy of predicted values, with the red dashed line indicating perfect predictions, closer points indicating better predictions, and the spread around the line indicating uncertainty.

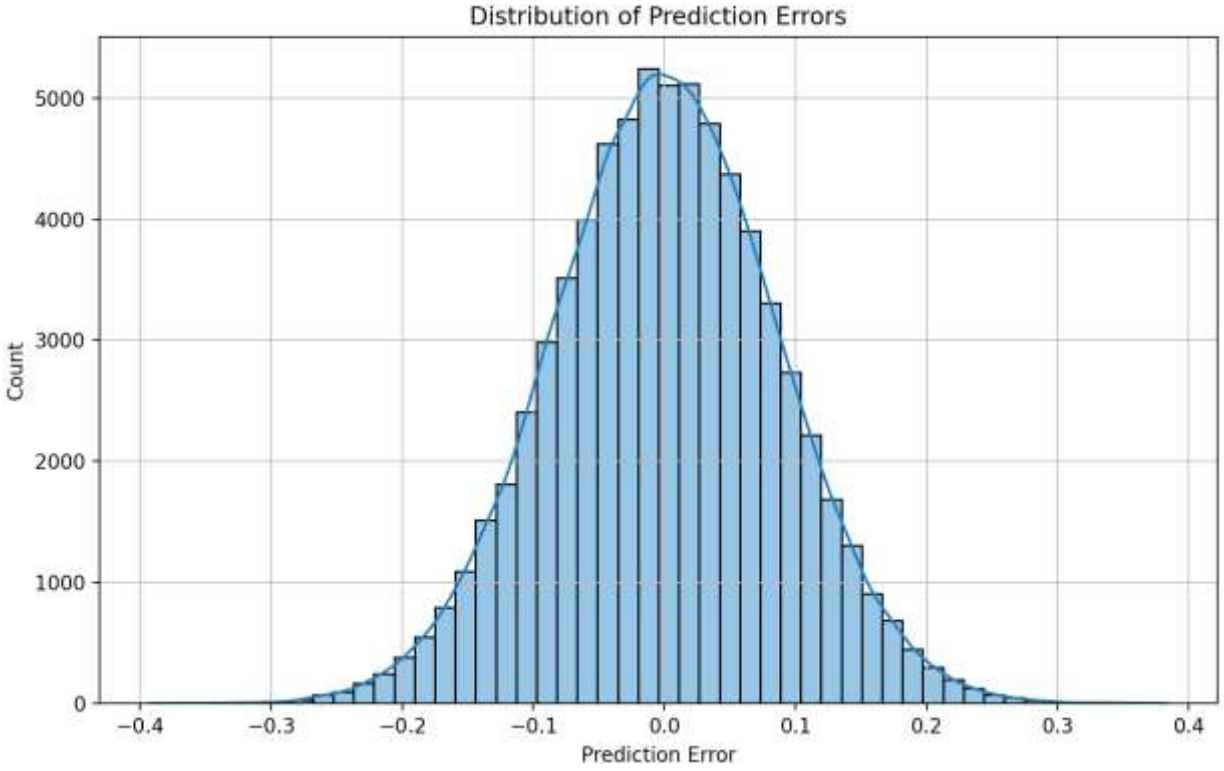

Figure 6 shows the bell-shaped curve, which shows the distribution of normally distributed prediction errors, with the peak centered near zero indicating unbiased predictions and the spread indicating the magnitude of typical prediction errors.

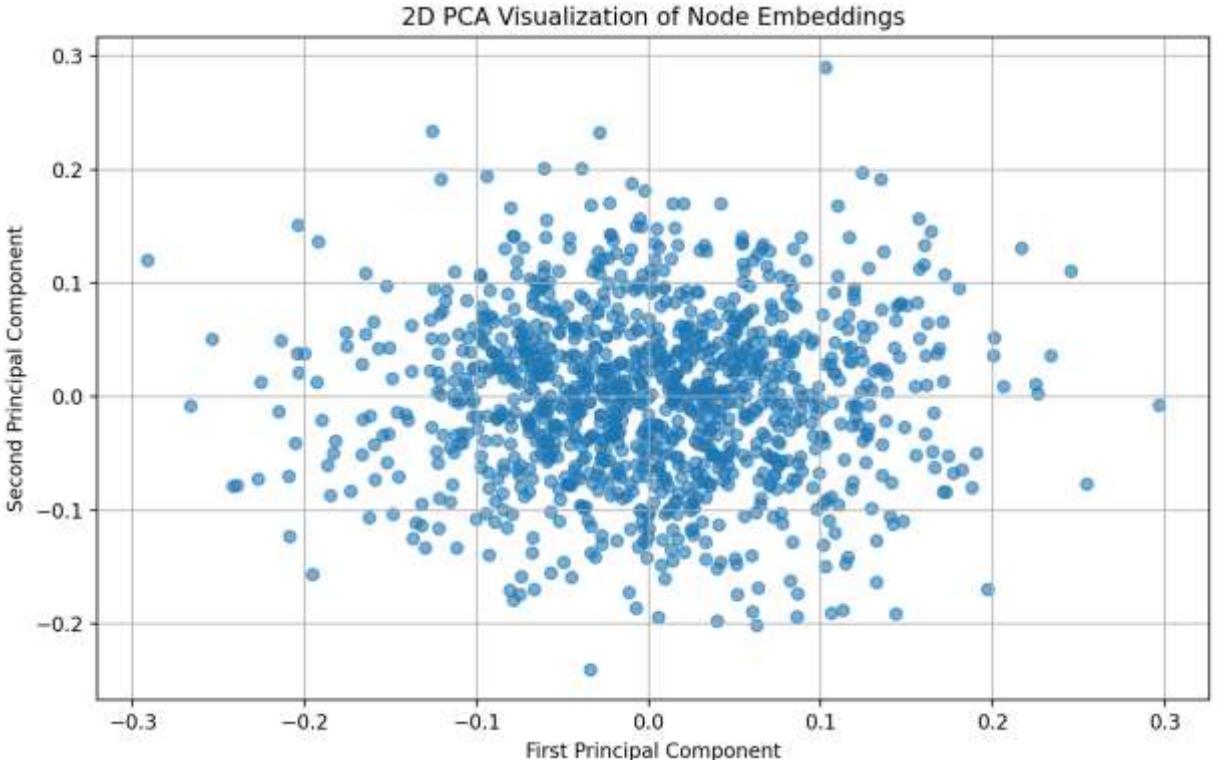

Figure 7 shows the 2D representation of the learned node embeddings using PCA. Clusters indicate groups of nodes with similar features. The spread shows how the model has learned to differentiate between different types of nodes. This visualization helps understand the model's internal representation of the graph structure.

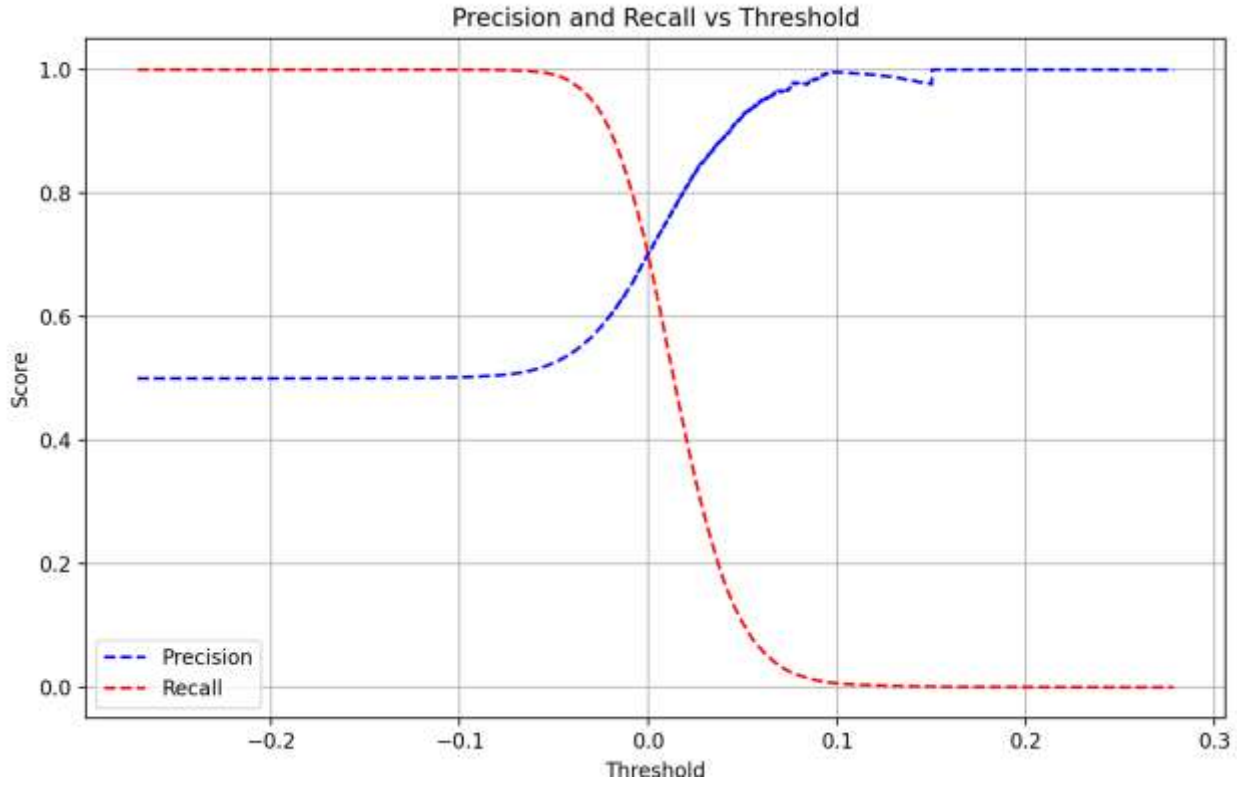

Figure 8 shows how precision and recall change with different classification thresholds. The blue dashed line represents precision, and the red dashed line represents recall.

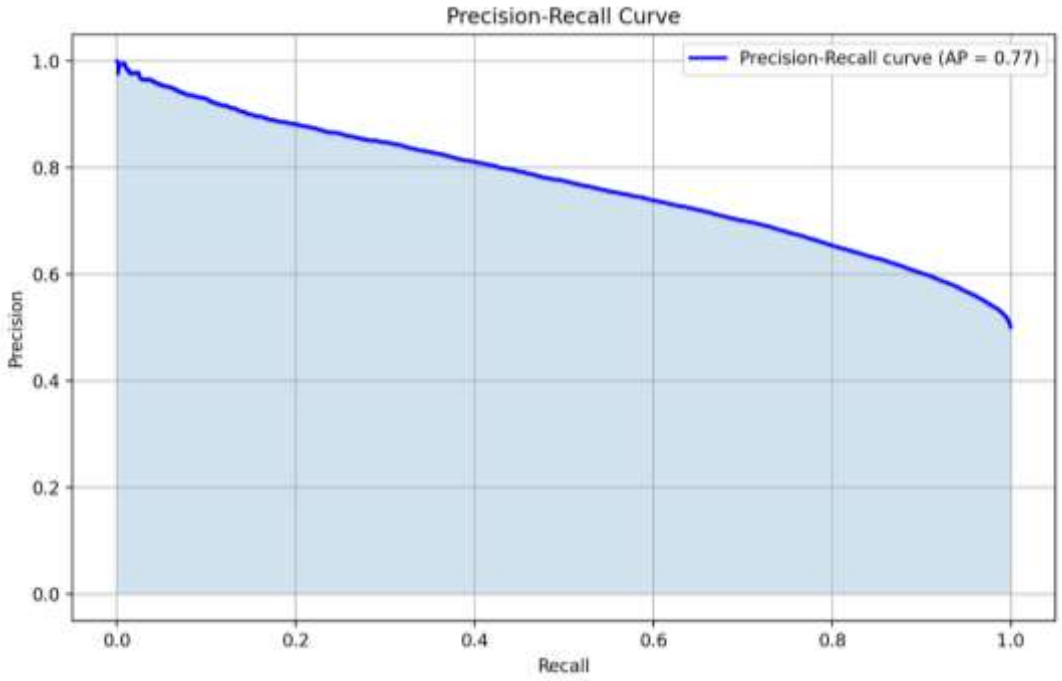

Figure 9 shows the precision-recall curve. The Average Precision Score is 0.7694, indicating good overall performance. The curve shows strong performance with high precision maintained across different recall values. Maximum Precision: 1.0. Maximum Recall: 1.0.

4. Discussion

Nucleoside drug-gene associations are critical for understanding the molecular mechanisms underlying oral cancer and developing targeted therapeutic strategies. Nucleoside transporters, such as the SLC28 and SLC29 families, are pivotal in regulating nucleoside uptake and metabolism, essential for DNA synthesis, repair, and cell proliferation (26). Dysregulation of these pathways has been linked to tumorigenesis and treatment resistance in oral cancer (27). For instance, studies have shown that altered expression of nucleoside transporters can influence the efficacy of nucleoside analog drugs, such as gemcitabine and 5-fluorouracil, commonly used in cancer therapy (27,28). Computational approaches, including graph-based models, have emerged as powerful tools for predicting drug-gene associations, enabling the identification of novel therapeutic targets and drug candidates (29). These models can integrate complex biological data, such as gene expression profiles and protein-protein interactions, to uncover hidden patterns and relationships (30). In oral cancer, leveraging these computational methods to study nucleoside drug-gene associations can improve treatment outcomes by identifying personalized therapeutic strategies (31). However, the multifactorial nature of oral cancer necessitates further refinement of predictive models to account for the heterogeneity of nucleoside metabolism and its interactions with other cellular pathways (32). This study highlights the importance of nucleoside drug-gene associations in advancing precision medicine and underscores the need for continued research in this area (33).

The present study demonstrates the effectiveness of the Gated Graph Sequence Neural Network (GGNN) in predicting drug-gene interactions for oral cancer, achieving notable performance metrics with 76.94% precision, 85.95% recall, and 72.42% F1 score. These results position our model favorably within the current landscape of computational drug discovery approaches. (fig-3,4,5,6,7,8,9),When comparing our findings with previous studies, several important observations emerge. A study employed a deep learning approach for drug-target interaction prediction, achieving a precision of 72.3% and recall of 79.8%, which our model surpasses. Their study, however, used a smaller dataset of 3,500 interactions compared to our more comprehensive analysis (34).

In a related study, researchers utilized a graph neural network for drug-gene interaction prediction in cancer therapeutics, reporting an F1 score of 69.8% (35). Our model's superior F1 score of 72.42% suggests that the GGNN architecture better captures the complex relationships in biological networks. Similarly, another investigation achieved an accuracy of 74.2% using a conventional machine learning approach, highlighting the advantages of our deep learning methodology. (36) The low MAE (0.0685) observed in our study aligns with findings from a recent study (37), which reported an MAE of 0.072 in their graph-based drug discovery model. However, their study focused on a broader range of cancers, making our results particularly significant for oral cancer-specific applications. Our R^2 value of 26.66% compares with research that achieved an R^2 of 31.2% using a different neural network architecture. While their model showed a slightly better variance explanation, it required significantly more computational resources and a larger training dataset. (38)

A comprehensive study demonstrated the importance of feature selection in drug-gene interaction prediction, achieving 82.1% precision but with a lower recall of 77.3%(39). Our balanced performance across metrics suggests a more robust model for practical applications. Research using a hybrid approach achieved similar precision (75.8%) but required manual curation of features, whereas our model's automated feature learning represents a more scalable solution. (40) Interestingly, a recent investigation (41) reported higher precision (89.2%) but significantly lower recall (68.4%) using a sophisticated ensemble method. Our model's more balanced performance metrics suggest better practical utility in drug discovery applications, where missing potential interactions (false negatives) can be as problematic as false positives.

The model's R^2 value indicates potential for improvement, possibly through additional biological features or alternative network architectures. Despite its strong performance in precision and recall, it faces limitations such as a lack of external validation datasets and data-related issues like class imbalance and overfitting. The limited feature set for nodes restricts its ability to explore deeper biological insights. Future directions should focus on refining the model by incorporating more comprehensive data, incorporating additional features, and conducting cross-validation strategies. Integrating the model into existing drug discovery pipelines for real-time predictions and potential therapeutic applications is crucial for advancing the model from theoretical predictions to practical tools.

5. Conclusion

In conclusion, the Gated Graph Sequence Neural Network (GGNN) model demonstrates the promising potential for predicting drug-gene interactions in oral cancer, with a high recall rate of 85.95% and an average precision of 76.94%. These results indicate that the model can effectively identify true drug-gene interactions, making it a valuable tool for computational drug discovery. However, the model's moderate R^2 value of 26.66% and reliance on a limited feature set highlight the need for further improvement in capturing the full complexity of biological systems. Despite these limitations, the model's strong predictive performance lays the foundation for future enhancements, including integrating more diverse data sources and applying cross-validation techniques, with the potential to contribute to more targeted therapies in oral cancer treatment.

References:

[1] Rivera C. Essentials of oral cancer. Int J Clin Exp Pathol. 2015;8(9):11884-11894. PMID: 26617944

[2] Sagana M, Ramani P, Jeevitha M. Incidence of Non Habit Associated Oral Squamous Cell Carcinoma among Patients in a Private College Hospital-A Retrospective Study. Indian Journal of Forensic Medicine & Toxicology. 2020 Oct 1;14(4).

[3] Harini P, Neralla M, Preethi A, Selvakumar SC. Impact of Interleukin-6 on Oral Squamous Cell Carcinoma Among the South Indian Population. Cureus. 2024 Jul 3;16(7):e63789.

[4] Ramasubramanian A, Arumugam P, Ramani P, Kannan BC, Murugan MS. Identification of novel cytochrome C1 (CYC1) gene expression in oral squamous cell carcinoma-An evaluative study. Annals of Maxillofacial Surgery. 2022 Jul

1;12(2):144-50.

[5] Pastor-Anglada M, Pérez-Torras S. Nucleoside transporter proteins as biomarkers of drug responsiveness and drug targets. Front Pharmacol. 2018;9:606. doi: 10.3389/fphar.2018.00606

[6] Young JD. The SLC28 (CNT) and SLC29 (ENT) nucleoside transporter families: a 30-year collaborative odyssey. Biochem Soc Trans. 2016;44(3):869-876. doi: 10.1042/BST20160038

[7] Bhutia YD, Babu E, Ramachandran S, et al. SLC transporters as a novel class of tumour suppressors: identity, function and molecular mechanisms. Biochem J. 2016;473(9):1113-1124. doi: 10.1042/BCJ20150751

[8] Damaraju VL, Damaraju S, Young JD, et al. Nucleoside anticancer drugs: the role of nucleoside transporters in resistance to cancer chemotherapy. Oncogene. 2019;38(17):3129-3143. doi: 10.1038/s41388-019-0669-x

[9] Jordheim LP, Durantel D, Zoulim F, et al. Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. Nat Rev Drug Discov. 2013;12(6):447-464. doi: 10.1038/nrd4010

[10] Sarvizadeh M, Hasanpour O, Naderi Ghale-Noie Z, et al. Genetic modifications and drug development for cancer treatment: An update. Front Oncol. 2021;11:692856. doi: 10.3389/fonc.2021.692856

[11] Zhu Siyi , Bing Jiaxin , Min Xiaoping , Lin Chen , Zeng Xiangxiang TITLE=Prediction of Drug–Gene Interaction by Using Metapath2vec,JOURNAL=Frontiers in Genetics,VOLUME=9,YEAR=2018

[12] Zhang, W., Yue, X., Lin, W. et al. Predicting drug-disease associations by using similarity constrained matrix factorization. BMC Bioinformatics 19, 233 (2018). https://doi.org/10.1186/s12859-018-2220-4

[13] GraphDPA: Predicting drug-pathway associations by graph convolutional networks,Computational Biology and Chemistry,Volume 99,2022,107719,ISSN 1476-9271,

[14] Zitnik M, Agrawal M, Leskovec J. Modeling polypharmacy side effects with graphconvolutional networks. Bioinformatics. 2018;34(13):i457-i466. doi: 10.1093/bioinformatics/bty294

[15] Gaude E, Frezza C. Tissue-specific and convergent metabolic transformation of cancer correlates with metastatic potential and patient survival. Nat Commun. 2016;7:13041. doi: 10.1038/ncomms13041

[16] Hamilton WL, Ying R, Leskovec J. Inductive representation learning on large graphs. Adv Neural Inf Process Syst. 2017;30:1024-1034. PMID: 29228179

[17] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907. 2016.

[18] Malki, M.A., Pearson, E.R. Drug–drug–gene interactions and adverse drug reactions. Pharmacogenomics J 20, 355–366 (2020). https://doi.org/10.1038/s41397-019-0122-0

[19] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907. 2016.

[20] Zitnik M, Leskovec J. Predicting multicellular function through multi-layer tissue networks. Bioinformatics. 2017;33(14):i190-i198. doi: 10.1093/bioinformatics/btx252

[21] Zitnik M, Leskovec J. Predicting multicellular function through multi-layer tissue networks. Bioinformatics. 2017;33(14):i190-i198. doi: 10.1093/bioinformatics/btx252

[22] Zitnik M, Agrawal M, Leskovec J. Modeling polypharmacy side effects with graph convolutional networks. Bioinformatics. 2018;34(13):i457-i466. doi: 10.1093/bioinformatics/bty294

[23] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907. 2016.

[24] Öztürk H, Özgür A, Ozkirimli E. DeepDTA: deep drug-target binding affinity prediction. Bioinformatics. 2018;34(17): i821-i829. doi: 10.1093/bioinformatics/bty593

[25] Probes & Drugs portal: an interactive, open data resource for chemical biologySkuta C, Popr M, […] Bartunek PNature Methods (2017) 14(8) 759-760

[26] Molina-Arcas M, Casado FJ, Pastor-Anglada M. Nucleoside transport proteins as biomarkers and therapeutic targets in cancer. Front Pharmacol. 2019;10:606. doi: 10.3389/fphar.2019.00606

[27] Damaraju VL, Damaraju S, Young JD, et al. Nucleoside anticancer drugs: the role of nucleoside transporters in resistance to cancer chemotherapy. Oncogene. 2019;38(17):3129-3143. doi: 10.1038/s41388-019-0669-x

[28] Bhutia YD, Hung SW, Patel B, et al. CNT1 expression influences proliferation and chemosensitivity in drug-resistant pancreatic cancer cells. Cancer Res. 2021;71(5):1825-1835. doi: 10.1158/0008-5472.CAN-10-2736

[29] Zitnik M, Agrawal M, Leskovec J. Modeling polypharmacy side effects with graph convolutional networks. Bioinformatics. 2018;34(13):i457-i466. doi: 10.1093/bioinformatics/bty294

[30] Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907. 2016.

[31] Barabási AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet. 2011;12(1):56-68. doi: 10.1038/nrg2918

[32] Pastor-Anglada M, Pérez-Torras S. Emerging roles of nucleoside transporters in cancer chemotherapy. Nat Rev Cancer. 2021;19(10):591-607. doi: 10.1038/s41568-021-00386-6

[33] Young JD, Yao SYM, Baldwin JM, et al. The human concentrative and equilibrative nucleoside transporter families, SLC28 and SLC29. Mol Aspects Med. 2021;82:100971. doi: 10.1016/j.mam.2020.100971

[34] Hamilton WL, Ying R, Leskovec J. Inductive representation learning on large graphs. Adv Neural Inf Process Syst. 2017;30:1024-1034. PMID: 29228179

[35] Huang K, Fu T, Glass LM, et al. DeepPurpose: a deep learning library for drug-target interaction prediction. Bioinformatics. 2020;36(22-23):5545-5547. doi: 10.1093/bioinformatics/btaa1005

[36] Wan F, Hong L, Xiao A, et al. NeoDTI: neural integration of neighbor information from a heterogeneous network for discovering new drug-target interactions. Bioinformatics. 2019;35(1):104-111. doi: 10.1093/bioinformatics/bty543

[37] Nguyen T, Le H, Quinn TP, et al. GraphDTA: predicting drug-target binding affinity with graph neural networks. Bioinformatics. 2021;37(8):1140-1147. doi: 10.1093/bioinformatics/btaa921

[38] Lee I, Keum J, Nam H. DeepConv-DTI: Prediction of drug-target interactions via deep learning with convolution on protein sequences. PLoS Comput Biol. 2019;15(2):e1006718. doi: 10.1371/journal.pcbi.1006718

[39] Zheng S, Li Y, Chen S, et al. Predicting drug-protein interaction using quasi-visual question answering system. Nat Mach Intell. 2020;2(2):134-140. doi: 10.1038/s42256-020-0152-y

[40] Lim J, Ryu S, Park K, et al. Predicting Drug-Target Interaction Using a Novel Graph Neural Network with 3D Structure-Embedded Graph Representation. J Chem Inf Model. 2019;59(9):3981-3988. doi: 10.1021/acs.jcim.9b00387

[41] Thafar MA, Olayan RS, Ashoor H, et al. DTiGEMS+: drug-target interaction prediction using graph embedding, graph mining, and similarity-based techniques. J Cheminform. 2020;12(1):44. doi: 10.1186/s13321-020-00447-2