Drug Discovery with Graph Neural Networks

Mihir S. Arya[(1)], Isabella Chittaro[(1)], Rinny Fan[(2)], George Simmons[(2)], Cristian Minoccheri[(3)], Kayvan Najarian[(3)]
(1) College of Engineering, University of Michigan, (2) College of Literature, Science, and the Arts, University of Michigan
(3) Computational Medicine and Bioinformatics, University of Michigan

MICHIGAN RESEARCH AND DISCOVERY SCHOLARS
UNIVERSITY OF MICHIGAN

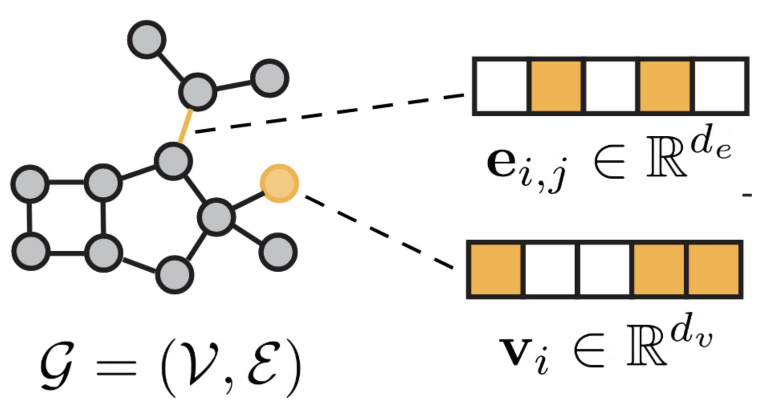
Abstract

The diminishing efficacy of antibiotics against increasingly resistant bacteria poses a significant public health challenge. Despite the escalating threat of antibiotic resistance, the development of new antibiotics is hindered by scientific, regulatory, and financial barriers. Recent literature suggests that antimicrobial polymers, despite their stochastic nature, may be viable alternatives to traditional peptides due to lower cost and production efficiency. To facilitate the process of exploring antimicrobial polymers, we investigate machine learning approaches for property prediction and molecule generation. This includes an attempt at introducing attention to a directed, weighted message passing neural network (MPNN) developed for polymers, and a comparative analysis on the property prediction accuracy of GCN and GAT model architectures. Our results from the comparative analysis suggest that GAT is a viable approach, and mean pooling yielding the highest accuracy. We further investigate the use of classifiers trained as energy models to generate novel data for research and development of antimicrobial polymers.

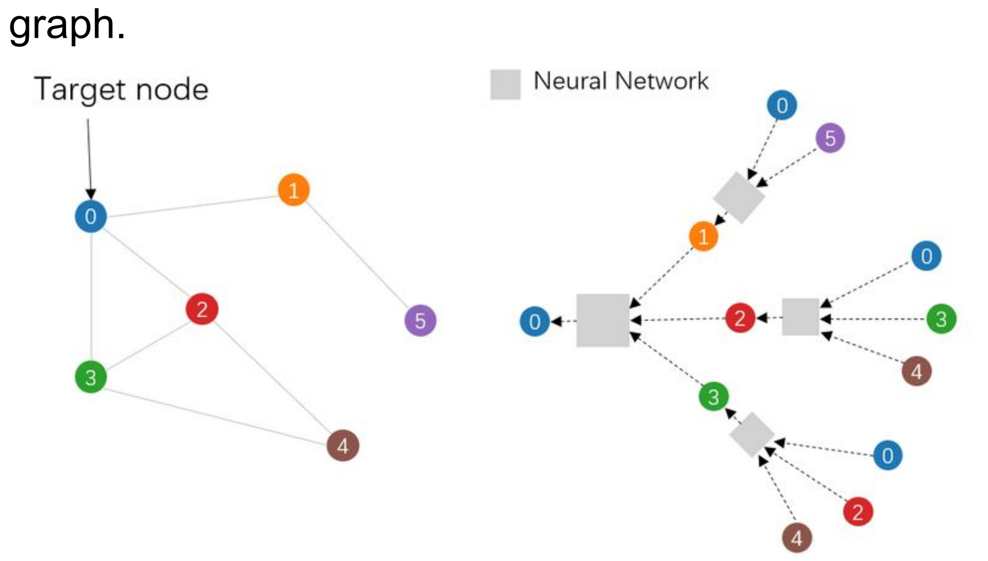
Introduction

Machine learning has enabled significant advances in science through rapid inference on large volumes of data. Among various deep learning methods, graph neural networks (GNNs) have emerged as a viable approach to performing inference on graph-structured data—a collection of nodes connected by edges.

The natural structure of molecules has lent itself to graph representations, where nodes represent atoms and edges represent chemical bonds. Notably, the Message Passing Neural Network (MPNN) framework for GNNs proposed by Gilmer et al. (2017)[1] has proven useful in molecular property prediction.



MPNNs aggregate information from neighborhoods of nodes through an iterative message-passing and aggregation procedure. For each node *v*, node embeddings of neighboring nodes *N(v)* are aggregated to update the node embedding of *v*. This produces a final feature vector *ŷ* that represents the whole graph.



(a) Input graph (b) Neighborhood aggregation

This exploratory study seeks to assess the efficacy of attention-based GNN architectures for molecular property prediction. By exploiting the unique advantages of GNNs for interpreting the graph-like structures of molecules, potentially in tandem with energy-based models for molecule generation, this study contributes to the broader understanding of how graph-inspired deep learning can propel innovations in drug discovery.

Methods

Consider an undirected graph *G* with node features x_v and edge features e_{vw}. Per Gilmer et al. (2017)[1], The message passing phase of an MPNN runs for *T* steps, during which the hidden states h_v^t at each node in the graph are updated based on messages m_v^{t+1} according to

$$m_v^{t+1} = \sum_{w \in N(v)} M_t(h_v^t, h_w^t, e_{vw}) \quad (1)$$

$$h_v^{t+1} = U_t(h_v^t, m_v^{t+1}) \quad (2)$$

where M_t is a learned message function and U_t is a learned vertex update function. *N(v)* denotes the neighbors of node *v* in graph *G*. The readout phase obtains a feature vector for the entire graph according to

$$\hat{y} = R(\{h_v^T \mid v \in G\}). \quad (3)$$

As described by Veličković et al. (2017)[2], Graph Attention Networks (GATs) introduce attention over a node's neighborhood to compute the importance of node *j*'s features to node *i*, represented by the attention coefficient

$$e_{ij} = a(\mathbf{W}\vec{h}_i, \mathbf{W}\vec{h}_j) \quad (4)$$

Normalized attention coefficients α_{ij} are used to compute the *i*th node's final output features $\vec{h}'_i$ according to the equation

$$\vec{h}'_i = \sigma\left(\sum_{j \in \mathcal{N}_i} \alpha_{ij} \mathbf{W} \vec{h}_j\right). \quad (5)$$

where *W* is a learned weight matrix, *σ* is a nonlinear activation function, and N_i is some neighborhood of node *i*.

Especially in the case of molecular input, pooling methods are critical for aggregating information on edges and nodes. Even slight variations in these data could result in very different structures. The two main types of pooling include max-pooling and mean-pooling.

Using PyTorch Geometric, we conducted a comparative analysis between the Graph Convolutional Network (GCN), as detailed by Kipf and Welling (2017)[3], and the Graph Attention Network (GAT), as described by Veličković et al. (2017)[2]. Each model was trained and evaluated on the Tox21_AhR dataset with a stratified train-test split over 50 epochs to assess classification task performance.

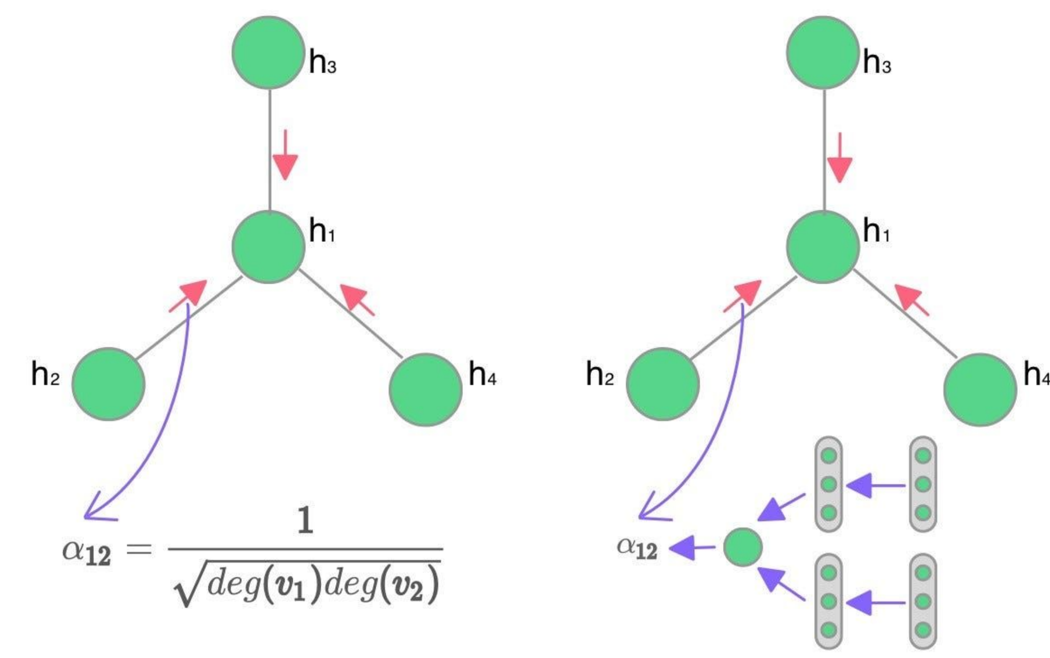
Results

Figure 1: Area under curve (AUC) results for different pooling types applied to two model architectures

Model Architecture		Pooling type		
		Max	**Mean**	**Add**
	GCN	0.8624	0.9241	0.6190
	GAT	0.7090	0.9379	0.6984

Model was trained over 50 epochs. Results suggest that GAT is viable approach to molecular property prediction.

Figure 2: Message passing neural networks (MPNNs)



GCN explicitly assigns non-parametric weight $\alpha_{ij} = \frac{1}{\sqrt{deg(v_i)deg(v_j)}}$, via the normalization function during neighborhood aggregation.

GAT implicitly captures the weight α_{ij}, via the attention mechanism, so that more important nodes receive higher weight during neighborhood aggregation.

Conclusions & Future Directions

Polymers are a new possible type of antimicrobial molecule, and our study investigated two ways in which machine learning can aid in their design: property prediction and molecule generation.

Property Prediction

In our exploration of how machine learning could be used for property prediction, we tried two approaches at designing a GNN classifier for polymers. These models are works in progress.

- Adding self attention to a directed GNN: we attempted to modify the Chemprop wd-MPNN developed by Aldeghi et al. (2022)[5]
- Adding direction to a graph attention network (GAT): We experimented with a GAT we built using PyTorch Geometric.
- Joint energy based framework: Another possibility is to use JEM by Grathwohl et al. (2020)[6], to train a classifier as an energy model.

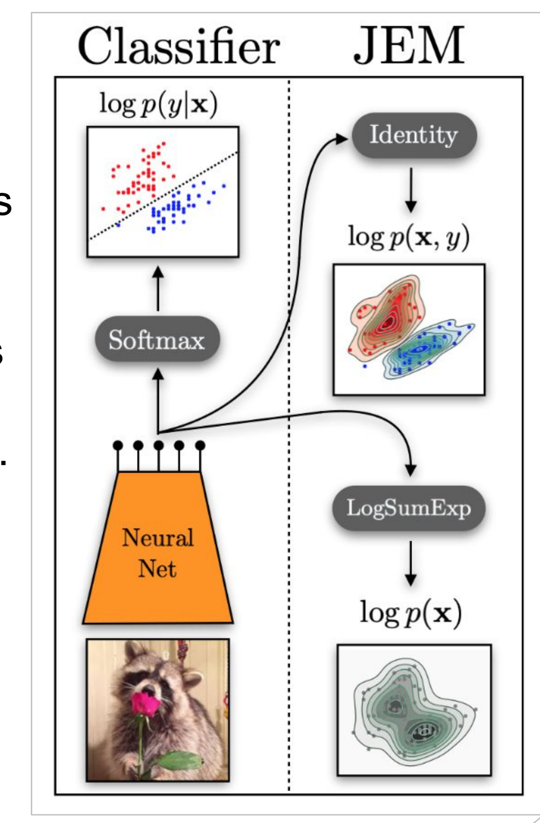
Molecule Generation

Training a classifier as an energy model means it can also be used to generate data.

- Grathwohl et al. (2020)[6] demonstrated this is possible using a convolutional neural network (CNN) that handled image input.
- A successful modification of JEM that allows it to handle molecular input could mean a new way of conducting molecule generation.

We experimented with a "toy version" of JEM and various PyTorch Geometric datasets, including Tox21. The work we did with JEM is a work in progress.

As of writing, using JEM for molecular input remains a largely unexplored topic that we consider an active research direction.



References

1. Gilmer J, Schoenholz SS, Riley PF, et al. (2017). Neural Message Passing for Quantum Chemistry.
2. Veličković P, Cucurull G, Casanova A, et al. (2018). Graph Attention Networks.
3. Thomas N. Kipf, & Max Welling. (2017). Semi-Supervised Classification with Graph Convolutional Networks.
4. Shaked Brody, Uri Alon, & Eran Yahav. (2022). How Attentive are Graph Attention Networks?.
5. Aldeghi, M., & Coley, C. W. (2022). A graph representation of molecular ensembles for polymer property prediction. Chemical science, 13(35), 10486–10498. https://doi.org/10.1039/d2sc02839e
6. Grathwohl W, Wang KC, Jacobsen JH, et al. (2020). Your Classifier is Secretly an Energy Based Model and You Should Treat it Like One.
7. GNNVis: A Visual Analytics Approach for Prediction Error Diagnosis of Graph Neural Networks - Scientific Figure on ResearchGate. Available from: https://shorturl.at/bhDGV
8. T035 - GNN-based molecular property prediction. Available from: https://shorturl.at/mMWZ3
9. What is Graph Attention Network? Available from: https://shorturl.at/hwBHQ