Comparing Graph Kernels and Graph Neural Networks for Brain Disease Classification
Midterm Presentation

Erica Choi\textsuperscript{1}    Sally Smith\textsuperscript{2}    Ethan Young\textsuperscript{3}

\textsuperscript{1}Columbia University
\textsuperscript{2}Georgia Institute of Technology
\textsuperscript{3}University of California - Los Angeles

Emory REU/RET, 2022
Models Meet Data

Figure 1: Example fMRI brain scan.

Can you tell if this brain is diseased or not?

1Center for Functional MRI - UC San Diego
Goals & Motivation

- Accurately classify patients as diseased or healthy
- Improve upon existing graph neural network performance by developing novel architectures
- Contribute to computational neuroscience literature by improving models that could eventually be used for mental illness diagnosis
Datasets

- We are working with 2 datasets, each classifying HIV and BP (bipolar disorder).
- Each dataset consists of DTI scans, FMRI scans, and classification labels (diseased (-1)/ non-diseased (1)).
- Both datasets have been cleaned for us and consist of less than 100 patients.
Datasets

- The DTI and FMRI brain scans of each patient $i$ are represented as weighted adjacency matrices $W_i \in \mathbb{R}^{M \times M}$.
- FMRI scans are considered to be more robust than DTI scans, so our experiments prioritize working with them.
- Nodes in the brain network represent regions of interest (ROI), and edge links between nodes indicate the strength of the connection between ROI’s.
Data Preprocessing

- For our data, we implemented a rounding scheme to remove edge weights and sparsify the adjacency matrices.

- We have: $A_{ij} = \begin{cases} 1 & \text{if } A_{ij} \geq \alpha \\ 0 & \text{otherwise} \end{cases}$, where $A_{ij}$ is the $ij$-th entry of the adjacency matrix $A$ and $\alpha \in (0, 1)$ is our rounding threshold.
Data Preprocessing

- We further manipulate the data to obtain a list of (graph) objects that can be used with the Python packages GraKel and PyG.

- GraKel’s functions and classes implement efficient computations of graph kernels to be used for tasks such as classification.

- PyG (PyTorch Geometric) builds on PyTorch and streamlines the implementation of graph neural network (GNN) pipelines.
The standard graph classification task considers the problem of classifying graphs into two or more categories.

The goal is to learn a model that maps graphs in the set of graphs $G$ to a set of labels $Y$. 
Graph kernels\textsuperscript{2} are popular in graph-based learning and have applications in many fields because their computation boils down to an inner product.

Our goal is to compute graph kernels and plug them into a kernelized learning algorithm to benchmark their performance on our datasets.

\textsuperscript{2}Yanardag and Vishwanathan, “Deep Graph Kernels”, 2015
Intuitively, this counts the frequency of size-$k$ subgraphs and compares that between two graphs.

This kernel is defined as $\mathcal{K}_{GK}(\mathcal{G}, \mathcal{G}') = \langle f^\mathcal{G}, f^{\mathcal{G}'} \rangle$. 

Intuitively, this kernel compares the number of shared subtrees between two graphs. The W-L kernel is defined as $\mathcal{K}_{WL}(G, G') = \langle 1^G, 1^{G'} \rangle$. 
Graph Neural Networks (GNN’s)

- GNN’s combine node features and graph structures to perform specific prediction tasks
- A generic framework of GNN:
  - computing the representation of each node
  - applying a pooling strategy to obtain the graph representation
  - Multilayer perceptron (MLP) can be applied to make predictions
GAT and GCN

- Graph Attention Network (GAT) is a type of Convolutional Neural Network that operates on graphs.
- Graph Convolutional Network (GCN) is a special case of GAT’s with attention fully determined by graph structure alone, without node features.
BrainGB

- **BrainGB**: A Benchmark for Brain Network Analysis with Graph Neural Networks$^3$
- Measures accuracy, F1-score, and AUC of different parameters
  - Node feature construction
  - Message passing mechanisms
  - Pooling Strategies

---

$^3$Cui et al., “BrainGB: A Benchmark for Brain Network Analysis with Graph Neural Networks”, 2022
BrainGB - Node Feature Construction

- Natural node features are usually not available in brain network analysis
- Connection profile
BrainGB - Message Passing Mechanisms

- Message vector

\[ m^l_i = \sum_{j \in \mathcal{N}_i} m_{ij} = \sum_{j \in \mathcal{N}_i} M_l(h^l_i, h^l_j, w_{ij}) \]

\[ h^{l+1}_i = U_l(h^l_i, m^l_i) \]

- Node concat

\[ m_{ij} = MLP(h_i \| h_j) \]
BrainGB - Pooling Strategies

- Pooling strategy

\[ g_n = R(\{h_i \mid v_i \in G_n\}) \]

- Concat pooling

\[ g_n = \|_{k=1}^{M} h_i = h_1 \| h_2 \| \ldots \| h_k \]
Support Vector Machines (SVM)

- SVM is a supervised learning model that maps training data to points in Euclidean space, then separates them with a hyperplane.

- Because of the small number of observations, we averaged classification accuracy over 20 different train-test splits to get a handle on how well SVM is performing.
## Results - BrainGB

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Accuracy</th>
<th>F1</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV-GCN</td>
<td>$51.43 \pm 17.73$</td>
<td>$50.61 \pm 12.87$</td>
<td>$49.23 \pm 17.97$</td>
</tr>
<tr>
<td>BP-GCN</td>
<td>$61.74 \pm 11.15$</td>
<td>$65.72 \pm 7.84$</td>
<td>$61.06 \pm 11.24$</td>
</tr>
<tr>
<td>HIV-GAT</td>
<td>$57.14 \pm 12.78$</td>
<td>$59.18 \pm 21.87$</td>
<td>$51.43 \pm 18.00$</td>
</tr>
<tr>
<td>BP-GAT</td>
<td>$55.63 \pm 9.52$</td>
<td>$59.03 \pm 9.54$</td>
<td>$55.49 \pm 9.51$</td>
</tr>
</tbody>
</table>
## Results - SVC (W-L)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Threshold = 0.5</th>
<th>Optimal Threshold*</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV-dti (0.85*)</td>
<td>0.40±0.18</td>
<td>0.56±0.18</td>
</tr>
<tr>
<td>BP-dti (0.5*)</td>
<td>0.52±0.14</td>
<td>0.52±0.14</td>
</tr>
<tr>
<td>HIV-fmri (0.2*)</td>
<td>0.58±0.20</td>
<td>0.65±0.17</td>
</tr>
<tr>
<td>BP-fmri (0.2*)</td>
<td>0.53±0.13</td>
<td>0.57±0.14</td>
</tr>
</tbody>
</table>
## Results - SVC (GS)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Threshold $= 0.5$</th>
<th>W-L Optimal Threshold $^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV-dti (0.85*)</td>
<td>0.54±0.26</td>
<td>0.47±0.15</td>
</tr>
<tr>
<td>BP-dti (0.5*)</td>
<td>0.48±0.17</td>
<td>0.46±0.15</td>
</tr>
<tr>
<td>HIV-fmri (0.2*)</td>
<td>0.30±0.14</td>
<td>0.30±0.14</td>
</tr>
<tr>
<td>BP-fmri (0.2*)</td>
<td>0.50±0.16</td>
<td>0.50±0.16</td>
</tr>
</tbody>
</table>
Challenges

- Limited data (<100 patients in each dataset)
  - Consequently, we need to feed the datasets through our models more, which increases computation time.
- Ethical considerations unique to the field of neuroscience
BrainGB - Limitations

- GNN’s are usually shallow; deep GNN’s are still an active area of research.
- For brain networks, what kinds of graph structures are effective beyond the pairwise connections are still unknown.
Graph Kernel GNN’s

- Kernel SVC basically classifies at random; however, we can still leverage some notion of higher-order information given by kernels in GNN’s.
- Implement and establish benchmarks with different GNN architectures, such as the one proposed by Morris et al.\(^4\)
- Develop novel GNN architectures that incorporate graph kernels; this is motivated by work done by Feng et al.\(^5\)

---

\(^4\) Morris et al., “Weisfeiler and Leman Go Neural: Higher-order Graph Neural Networks”, 2019

\(^5\) Feng et al., “KerGNNs: Interpretable Graph Neural Networks with Graph Kernels”, 2022
Our goal is to improve brain disease classification models.

While we are limited by factors such as accessibility of datasets, we are working around the issues we are facing.

Our next step is to work on combining graph kernels with GNNs.
Acknowledgements & Contact Info

- Thank you to Emory for bringing us to this research site, and special thanks to our mentor Dr. Carl Yang for his guidance and 24 hour email service.
- Thank you to Hejie for her assistance with using BrainGB.
- Thank YOU all for being here and listening to our talk!

- If you have any further questions, feel free to contact us!
  - Erica Choi - erica.c@columbia.edu
  - Sally Smith - sallysmith@gatech.edu
  - Ethan Young - young.j.ethan@gmail.com
References

- Feng, Aosong et al. “KerGNNs: Interpretable Graph Neural Networks with Graph Kernels”. In: *AAAI*. 2022.
- Morris, Christopher et al. “Weisfeiler and Leman Go Neural: Higher-order Graph Neural Networks”. In: *2019*.