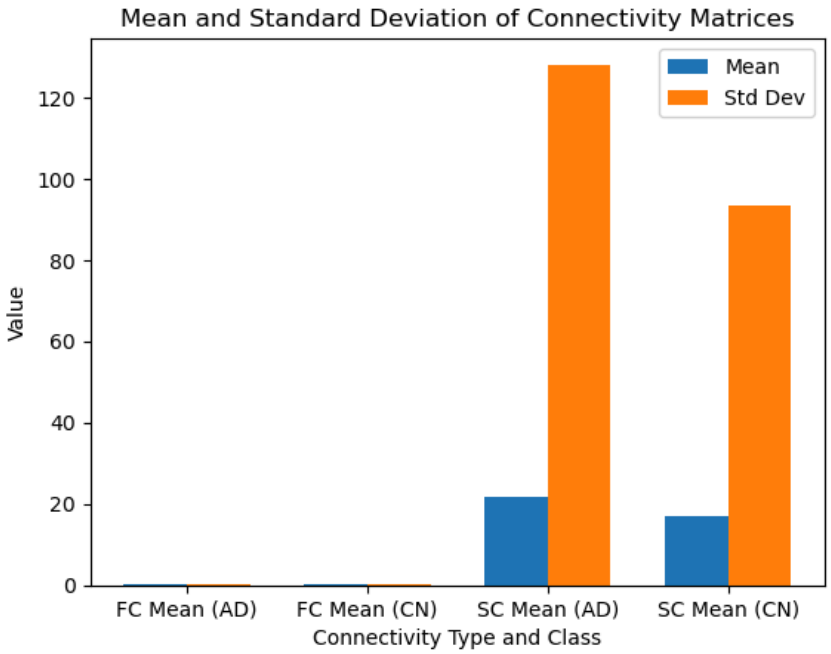
Name: Biraaj rout
UTA ID: 1002071886

Project 2(CSE 6389-001)

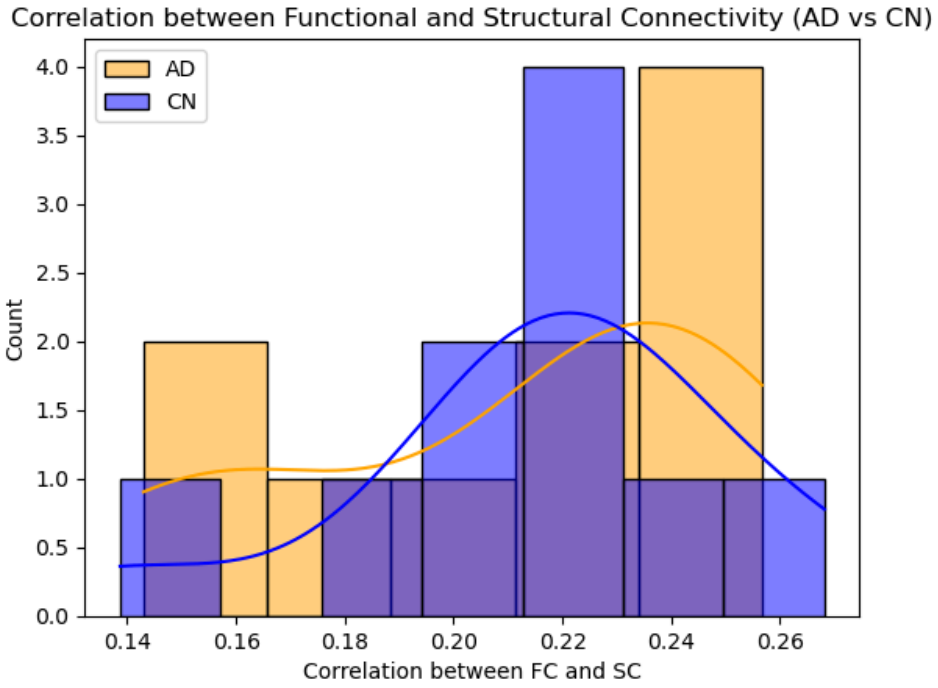
Dataset:

- The dataset consists of 10 samples for AD and 10 samples for CN so the classes are balanced for training.
- We used StratifiedKFold which is a cross-validation technique provided by scikit-learn that ensures each fold of the cross-validation split maintains the same proportion of each class label as in the entire dataset.
- AD is encoded as 1 and CN is encoded as 0 here for this project.
- The mean and std deviation of FC and SC matrices:



  - We did a **Z score normalization** to address this difference so that FC and SC are more comparable.
- The correlation between FC vs SC for AD vs CN:

Name: Biraaj Rout
UTA ID: 1002071886



- There is a significant overlap in the correlation distributions for AD and CN groups, with both centered around similar values (approximately 0.20-0.24).
- We did a t-test to see if there is a significant difference between both FC and SC for AD vs CN classification and found that there isn't as we got values **t-value = -0.2950, p-value = 0.7714**
- So GCN is a good approach here to find any complex differences for classification.

Model:

```
==========================================================================
Layer (type:depth-idx)                                   Param #
==========================================================================
├─GraphConvolution: 1-1                                  3,775
├─ReLU: 1-2                                              --
├─Dropout: 1-3                                           --
├─Linear: 1-4                                            52
==========================================================================
Total params: 3,827
Trainable params: 3,827
Non-trainable params: 0
==========================================================================
```

- We used a single Graph convolution layer as our dataset is small.
- Here our matrices for SC and FC are 150*150 in dimension. So GCN projects the 150 input features to a lower-dimensional space of 25 features.
- ReLU is applied as the activation function after the graph convolution layer to allow the model to learn complex patterns.

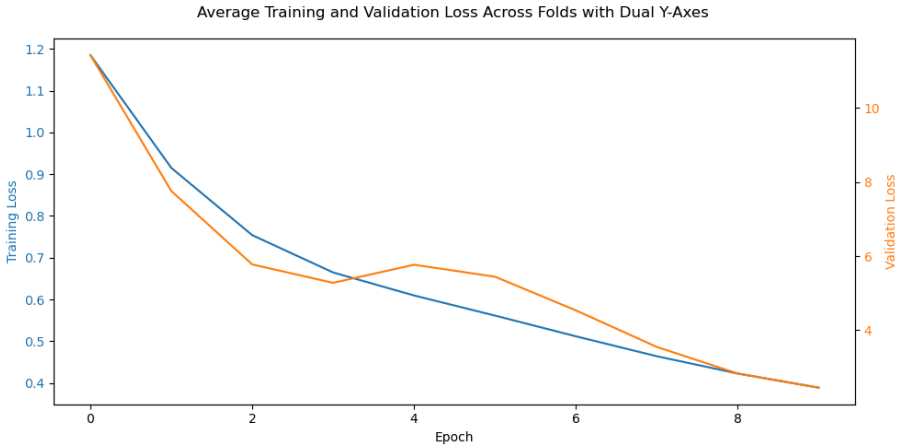
Name: Biraaj Rout
UTA ID: 1002071886

- **Dropout of 0.4**(optimal) [1] is used here to avoid overfitting. The training was done on 0.2, 0.3, 0.4, 0.5 and 0.6 dropout values.
- After graph convolution mean pooling [3] averages the feature representations across all nodes to create a single vector that represents the entire graph.
- The fully connected layer at the end is used to map the features down to two output units for binary classification, corresponding to the two classes: AD (Alzheimer's Disease) and CN (Cognitively Normal).
- The model has a total of **3,827 trainable parameters** which is relatively lightweight with respect to the small dataset we have.
- We used the **batch_size of 4** for training and validation as we have a very small dataset.
- Summary:
    - The GCN leverages the adjacency matrix (which represents the connectivity structure) to aggregate information from neighboring nodes.
    - With only 3,827 parameters, the model is relatively simple and less prone to overfitting.
    - The use of dropout helps with regularization, which further mitigates overfitting risks.

Training:

- We used k-fold cross-validation in model training. The K value is 4 here which means the model is split into 4 folds with a random sampling of **12 training inputs and 4 validation inputs**.
- The k-fold cross validation is used as the data set is small and we need to generalize the model to avoid overfitting.
- **gradient clipping** (torch.nn.utils.clip_grad_norm_(model.parameters(), max_norm=1.0)) is used to ensure that the gradient does not become too large during backpropagation. We avoid the issue of exploding gradient with this.
- We used an **ADAM optimizer** as it adjusts the learning rate dynamically and as our dataset is small we have to sure that the learning rate doesn't make the validation loss or training loss plateau at a certain point.
- The **learning rate** and **weight decay** [2] are set to **5e-4 and 5e-2** repectively.
- We used **10 epochs** to train because the dataset is small and we can see the validation loss becoming stable towards the 7th epoch.
- Average train and validation loss across all the folds.

Name: Biraaj Rout
UTA ID: 1002071886

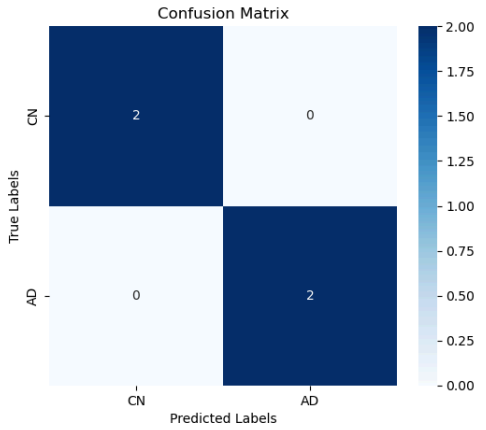


- Here we can see the training and validation loss decreases with each epoch which signifies model is learning well.
- There is a bit of fluctuation in validation loss as the dataset is small, so Improving it sometimes can lead to unstable results but since we have used L2 regularization the model doesn't overfit.
- Average validation results are as follows:
  - Accuracy: 0.6875
  - Precision: 0.6667
  - Recall: 0.8750
  - F1-Score: 0.7500
  - Balanced Accuracy: 0.6875
  - MCC: 0.3943
  - AUC-ROC: 0.6875
  - Confusion Matrix: 1.0000
  - Validation Loss: 2.4512
- Here we see there is high recall which is important for medical results and also the accuracy is low because mostly the model is able to predict ¾ cases which is good in terms of small dataset. Also validation loss is low which implies model trains well.
- After training is done for each fold we save the model as .pth file and test with the best model that we get after k-fold cross validation.

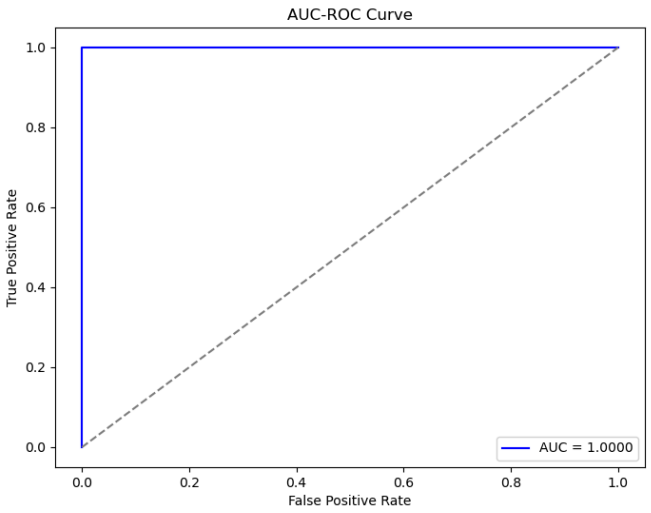
Evaluation:

- We referred to the paper "*A Convolutional Neural Network and Graph Convolutional Network Based Framework for AD Classification*" [4] for evaluation.
- We used the test dataset to evaluate our model which had 4 samples 2 from AD and 2 from CN.
- Evaluation metrics for test set:
  - **Confusion** Matrix:
    - TP: 2, TN: 2, FP: 0, FN: 0
    - All predictions are correct with this.

Name: Biraaj Rout
UTA ID: 1002071886



- **Accuracy**: 100.00%
- **Precision**: 100.00%
- **Recall (Sensitivity)**: 100.00%
  - Recall (sensitivity) shows that the model successfully identified all actual positive cases.
- **F1-Score**: 1.00
  - indicates a perfect balance between precision and recall.
- **Balanced Accuracy**: 1.00
  - meaning the model performs perfectly across both classes (AD vs CN).
- **Matthews Correlation Coefficient (MCC)**: 1.00
  - MCC is 1, which indicates perfect prediction capability
- **AUC-ROC**: 1.00
  - The Area Under the Curve is 1, indicating perfect classification with no errors.



Name: Biraaj Rout
UTA ID: 1002071886

Conclusion:

- This project aimed to develop a **Graph Convolutional Network (GCN)** model to distinguish between Alzheimer's Disease (AD) patients and Cognitively Normal (CN) individuals using functional and structural brain connectivity data. The dataset comprised connectivity matrices derived from brain scans, organized in folders representing AD and CN classes. We applied **Z-score normalization** to standardize the connectivity matrices and used **4-fold cross-validation** to evaluate the model's robustness. The results indicate that the GCN model was highly effective in distinguishing between AD and CN individuals using the functional and structural connectivity data provided. Achieving perfect scores on the test set implies that the model was able to capture relevant features in the connectivity matrices that correlate strongly with Alzheimer's Disease.However, it is essential to consider that these results may be due to specific characteristics of the dataset as it is small and has similar samples for AD and CN. In real-world applications, achieving such perfect performance is rare, especially on new, unseen data.

References:

1. https://doi.org/10.1145/3487553.3524725
2. Regularization graph convolutional networks with data augmentation - ScienceDirect
3. https://arxiv.org/pdf/2004.03519
4. https://pmc.ncbi.nlm.nih.gov/articles/PMC9961367/