Graph Pooling Algorithm for Graph Classification Task

1. Introduction

This report details the implementation and evaluation of a graph pooling algorithm for graph classification tasks. The model architecture consists of multiple Graph Neural Networks (GNNs) followed by Down-Sampling & Pooling layers, and experiments are conducted on ENZYMES and D&D datasets.

2. Model Architecture and Parameters

- **Graph Neural Network (GNN) Layers:**
    - **Number of GNN layers**: 4
        - **GNN1 → GNN2 → Down-Sample & Pool1 → GNN3 → GNN4 → Down-Sample & Pool2 → Classification Head**
    - **Size of Hidden Layers**:
        - **GNN1 & GNN2**: The input dimension is determined by the node feature size of the dataset, and the hidden dimension is set to **64**.
        - **GNN3 & GNN4**: The output from the first down-sampling layer is **64**, and the hidden dimension for these GNN layers is increased to **128**.
- **Down-Sampling and Pooling Layers:**
    - **k-values**: These define the percentage of nodes retained in each down-sampling phase.
        - **k-values tested**: **90%**, **80%**, **60%**
- **Hierarchical Pooling**: This defines the number of clusters in the pooling layers.
    - **m-values**: **6** for the first pooling layer and **3** for the second pooling layer.
- **Classifier Layer:**
    - **Fully Connected (FC) Layers**:
        - The classifier consists of a series of fully connected layers:
            - **First FC layer**: **128** hidden units
            - **Second FC layer**: **100** hidden units
            - **Final output layer**: **50** hidden units connected to the number of classes(**6** for ENZYMES dataset, **2** for D&D).
- **Activation Function:**
    - **ReLU** (Rectified Linear Unit) activation is used after each GNN and fully connected layer to introduce non-linearity.
- **Dropout:**
    - **Dropout rate**: **0.4** for Enzymes and **0.3** for DD is applied after each GNN layer and pooling layer to prevent overfitting.

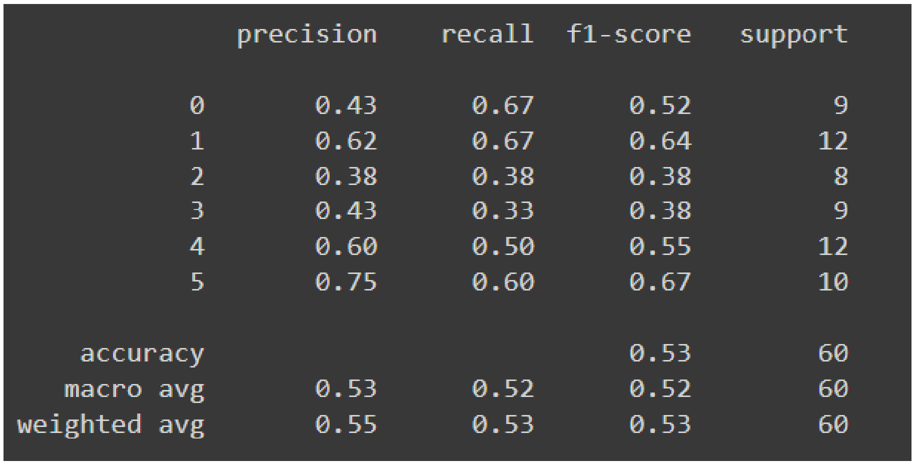
- **Batch Normalization:**
  - Batch normalization is applied after each GNN and fully connected layer to normalize activations.
- **Optimizer:**
  - **Adam** optimizer with a learning rate of **0.001**.
- **Epochs:**
  - **ENZYMES dataset**: **1000 epochs**
  - **D&D dataset**: **250 epochs**
- **Batch Size:**
  - **Batch size**: **20** for both datasets.
- **Learning Rate:**
  - Set to **0.001**.
- **Weight Initialization:**
  - Xavier initialization for the GNN layers with a gain factor adjusted for ReLU activation.

3. Results and Comparison

ENZYMES Dataset

k%	Validation Accuracy	Test Accuracy
90%	**0.65**	0.53
80%	0.58	**0.58**
60%	0.35	0.35

Table 1. Accuracy for enzymes dataset



```
              precision    recall  f1-score   support

           0       0.43      0.67      0.52         9
           1       0.62      0.67      0.64        12
           2       0.38      0.38      0.38         8
           3       0.43      0.33      0.38         9
           4       0.60      0.50      0.55        12
           5       0.75      0.60      0.67        10

    accuracy                           0.53        60
   macro avg       0.53      0.52      0.52        60
weighted avg       0.55      0.53      0.53        60
```

Fig 1. Classification Report (ENZYMES - k=90%):

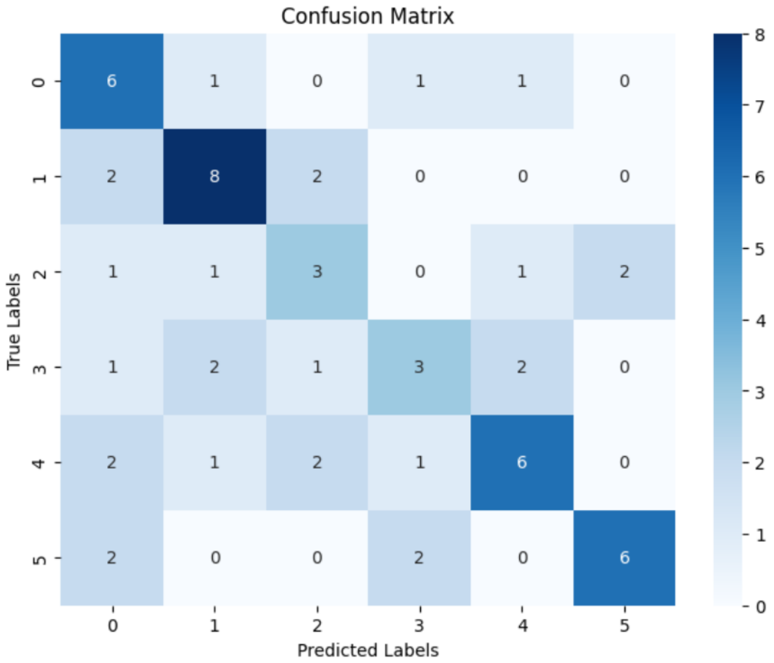


Fig 2. Confusion Matrix (ENZYMES - k=90%):

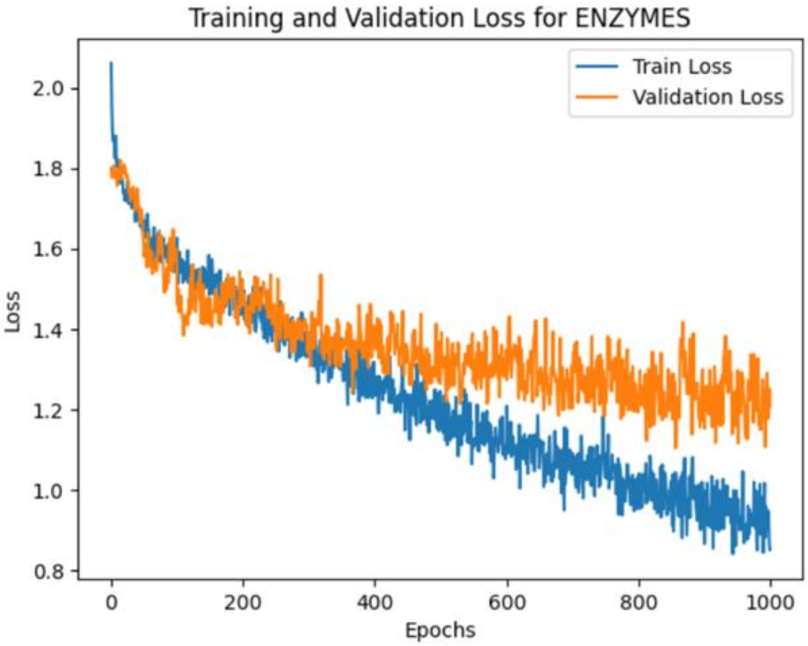


Fig 3. Training vs Validation Loss (ENZYMES - k=90%):

D&D Dataset

k%	Validation Accuracy	Test Accuracy
90%	0.84	**0.78**
80%	0.84	0.74
60%	**0.85**	**0.78**

Table 2. Accuracy for D&D Dataset

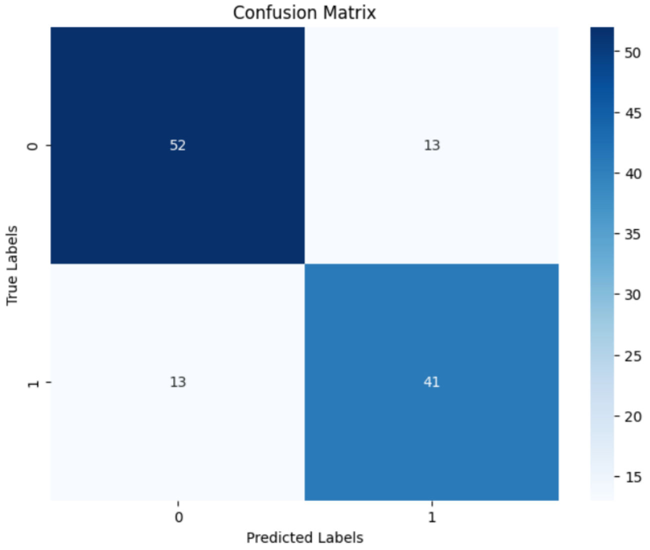
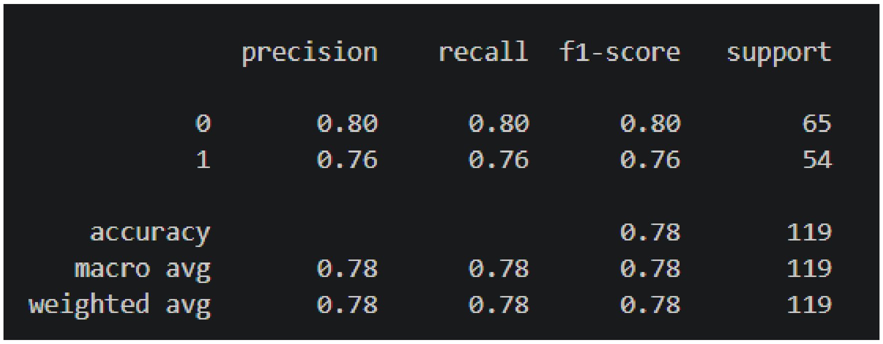


Fig 4. Confusion Matrix (D&D - k=90%):



```
              precision    recall  f1-score   support

           0       0.80      0.80      0.80        65
           1       0.76      0.76      0.76        54

    accuracy                           0.78       119
   macro avg       0.78      0.78      0.78       119
weighted avg       0.78      0.78      0.78       119
```

Fig 5. Classification Report (D&D - k=90%):

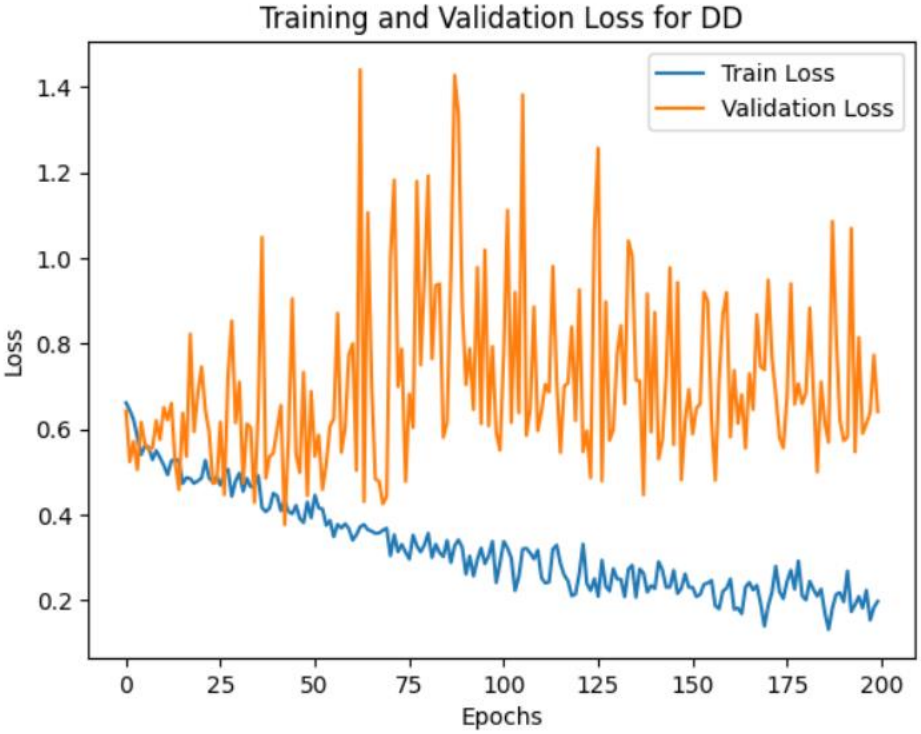


Fig 6. Training vs Validation Loss (D&D - k=90%):

4. Conclusion

This report presents the results of different experiments using various down-sampling percentages (k-values) and hierarchical pooling. The ENZYMES dataset achieved the highest test accuracy with k=80%, while the D&D dataset performed best with k=90%.