Assignment-2: Three Hidden Layer Neural Network for Multi-Class Classification

1. Dataset Generation

A synthetic dataset with five distinct classes was generated. Each class had 2000 samples, created using Gaussian distributions centered at different points in 2D space. This ensures linearly non-separable but visually distinguishable clusters. Labels were one-hot encoded using sklearn's OneHotEncoder.

2. Neural Network Architecture and Implementation

The neural network was implemented from scratch using NumPy. It contains:
a. Input Layer: 2 neurons (matching input features)
b. Hidden Layers: 3 layers with 13 neurons each using sigmoid activation
c. Output Layer: 5 neurons for 5-class classification

Weights are initialized randomly. The activation function used is sigmoid for both hidden and output layers. Loss function used is Mean Squared Error (MSE), and weights are updated using manual gradient descent.

3. Training and Evaluation

The dataset was split into 60% training, 20% validation, and 20% testing. Model was trained for 750 epochs with a learning rate of 0.0001. Accuracy and loss steadily improved across epochs without overfitting.

Sample Progress:
Epoch 0: Train Accuracy = 20.80%, Validation Accuracy = 20.75%
Epoch 400: Train Accuracy = 78.78%, Validation Accuracy = 76.85%
Epoch 750: Train Accuracy = 84.25%, Validation Accuracy = 82.50%

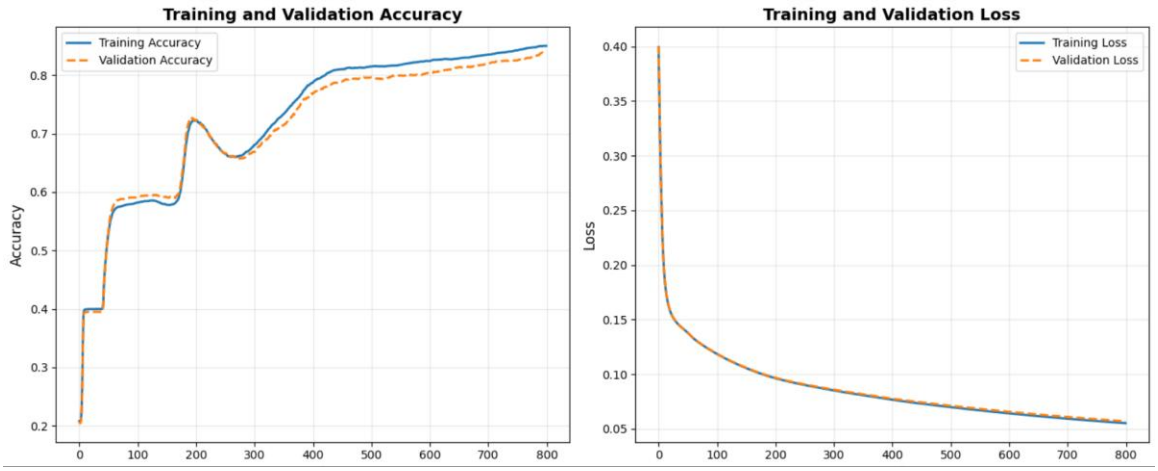

Figure 1: Training and Validation Accuracy/Loss

4. Test Results and Confusion Matrix

Final model was evaluated on 2000 test samples.

Classification Report:

Class 0: Precision = 0.84, Recall = 0.92, F1 = 0.88
Class 1: Precision = 0.94, Recall = 1.00, F1 = 0.97
Class 2: Precision = 0.97, Recall = 0.37, F1 = 0.54
Class 3: Precision = 0.96, Recall = 1.00, F1 = 0.98
Class 4: Precision = 0.71, Recall = 0.99, F1 = 0.83

Overall Test Accuracy = 86%
Macro F1-score = 0.84

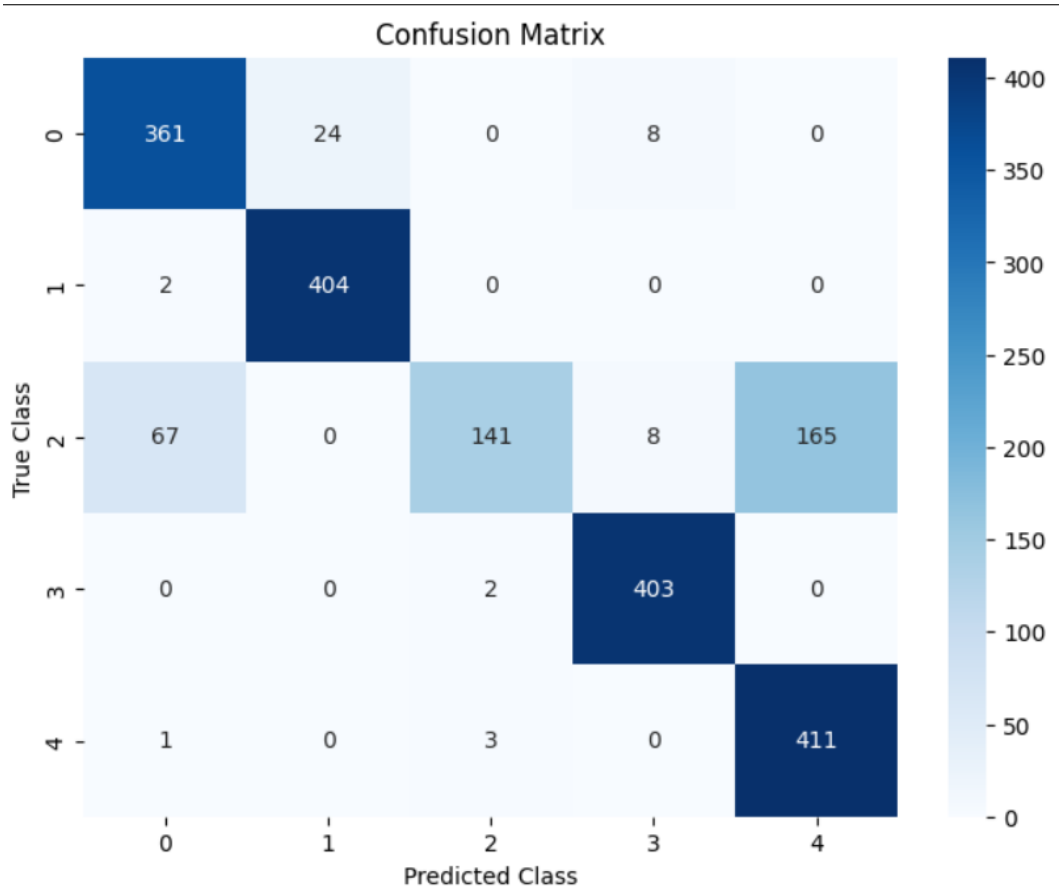

Figure 2: Confusion Matrix

5. ROC Curve for Each Class

One-vs-rest ROC curves were generated for all classes. The AUC values are:

Class 0: AUC = 0.99
Class 1: AUC = 1.00
Class 2: AUC = 0.92
Class 3: AUC = 1.00
Class 4: AUC = 1.00

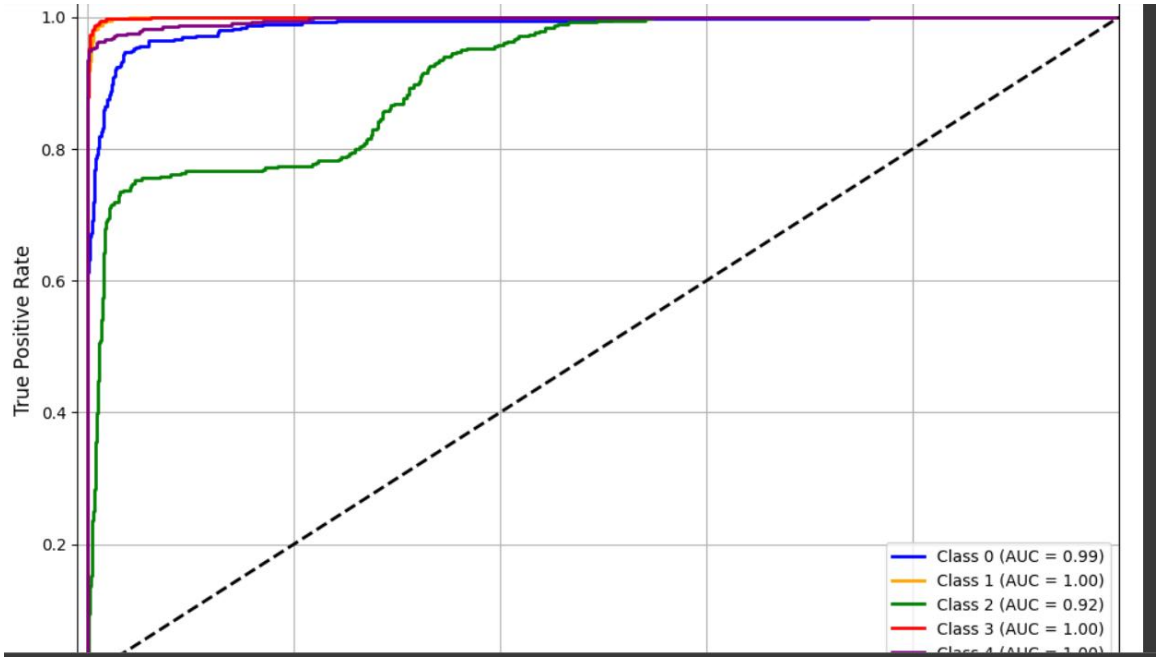

Figure 3: ROC Curve

6. Observations and Insights

a. Training and validation curves are close, indicating low overfitting.
b. Class 2 had a significantly lower recall, suggesting overlap with other classes.
c. ROC curves confirm strong classification power for most classes except class 2.

7. Conclusion and Future Scope

The neural network effectively classified the synthetic multi-class dataset. With proper tuning, 86% test accuracy was achieved. In future, using ReLU for hidden layers, softmax activation with cross-entropy loss, dropout regularization, and adaptive optimizers like Adam may further improve performance.