Summary Report

**"Geetika\_MSML\_Project2\_Final.ipynb"**

# Results

**Feedforward Neural Network (FFN):**

After training a feedforward neural network with two hidden layers (e.g., 784 → 512 → 256 → 10 units) across 5 independent runs, the average test accuracy was approximately **97.89% over 5 runs**. Individual runs varied slightly, but all were close to this mark, indicating stable performance.

**Convolutional Neural Network (CNN):**

A CNN with two convolutional layers (32 and 64 filters) followed by two fully connected layers reliably achieved an average test accuracy of about **99.02% over 5 runs**. Compared to the FFN, the CNN consistently produced higher accuracy.

# Lessons Learned

**Architecture Matters:**

The CNN’s superior performance over the FFN confirms that convolutional structures are more effective for image data, allowing the model to learn meaningful features that generalize better than what a fully connected network can achieve.

**Parameter Tuning:**

Carefully tuning parameters—such as learning rate, batch size, and the number of epochs—was critical to achieving stable and high accuracy. For instance, using a learning rate of 0.001 with the Adam optimizer and training for 5 epochs provided a good balance between convergence speed and final accuracy.

**Regularization and Initialization:**

Although MNIST is relatively simple, proper weight initialization and basic regularization (like using ReLU activations and careful selection of network sizes) helped stabilize training and prevent overfitting.

In summary, CNNs outperformed FFNs on MNIST. Careful architecture design and parameter selection played a key role in achieving the target accuracy of over 95%, with the CNN exceeding this threshold by a comfortable margin.