MNIST Dataset

Ahmed Ahmed Abdelhadi

Models Architectures

SimpleNN (Fully Connected Neural Network):

Model Architecture:

- Input Layer: 28x28 pixels (MNIST image size)
- Fully Connected Layer 1: 28x28 (input) -> 128 (hidden)
- ReLU Activation Function
- Fully Connected Layer 2: 128 (hidden) -> 10 (output, 10 classes)

Training:

- Loss Function: CrossEntropyLoss
- Optimizer: Adam
- Learning Rate: 0.001
- Batch Size: 64
- Epochs: 10

Training Output:

```
Epoch [1/10], Loss: 0.1482
Epoch [2/10], Loss: 0.1372
Epoch [3/10], Loss: 0.0998
Epoch [4/10], Loss: 0.0473
Epoch [5/10], Loss: 0.1818
Epoch [6/10], Loss: 0.0381
Epoch [7/10], Loss: 0.3520
Epoch [8/10], Loss: 0.3133
Epoch [9/10], Loss: 0.1703
Epoch [10/10], Loss: 0.0711
```

Test Accuracy: 97.44%

CNN (Convolutional Neural Network)

Model Architecture:

- Convolutional Layer 1: 1 channel (grayscale) -> 32 channels, kernel size 3x3, ReLU activation
- Max Pooling Layer 1: 2x2
- Convolutional Layer 2: 32 channels -> 64 channels, kernel size 3x3, ReLU activation
- Max Pooling Layer 2: 2x2
- Flatten Layer
- Fully Connected Layer 1: 64 * 7 * 7 (flattened) -> 128, ReLU activation
- Fully Connected Layer 2: 128 -> 10 (output, 10 classes)

Training:

- Loss Function: CrossEntropyLoss
- Optimizer: Adam
- Learning Rate: 0.001
- Batch Size: 64
- Epochs: 10

Training Output:

```
Epoch [1/10], Loss: 0.1155
Epoch [2/10], Loss: 0.0012
Epoch [3/10], Loss: 0.0044
Epoch [4/10], Loss: 0.0145
Epoch [5/10], Loss: 0.0017
Epoch [6/10], Loss: 0.0096
Epoch [7/10], Loss: 0.0005
Epoch [8/10], Loss: 0.0026
Epoch [9/10], Loss: 0.0000
Epoch [10/10], Loss: 0.0137
```

Test Accuracy: 98.86%

Technical Report:

- **SimpleNN:**
  - The fully connected neural network achieved a test accuracy of 98.86%, indicating good performance on the MNIST dataset.
  - The loss decreased consistently during training, suggesting effective learning.

- **CNN:**
  - The convolutional neural network also achieved a test accuracy of 98.86%, matching the performance of the fully connected network.
  - The convolutional layers, followed by max-pooling, proved effective in capturing hierarchical features from the images.

- **Comparison:**
  - Both models performed equally well on the MNIST dataset, achieving high accuracy.
  - The CNN might offer advantages in handling spatial features due to its convolutional layers, making it more suitable for image-related tasks.

- **Considerations:**
  - The test accuracy reached a high value, suggesting that the models generalize well to unseen data.
  - Fine-tuning and hyperparameter adjustments could further optimize the models, but the current results are already promising.

In conclusion, both models demonstrated strong performance on the MNIST dataset, showcasing the effectiveness of neural networks in image classification tasks.