NEURAL SYSTEM COURSEWORK: Fashion MNIST Classification

P NO: P2903988

FULL NAME: RAYANI MANI RAGHAVENDRA

Abstract:

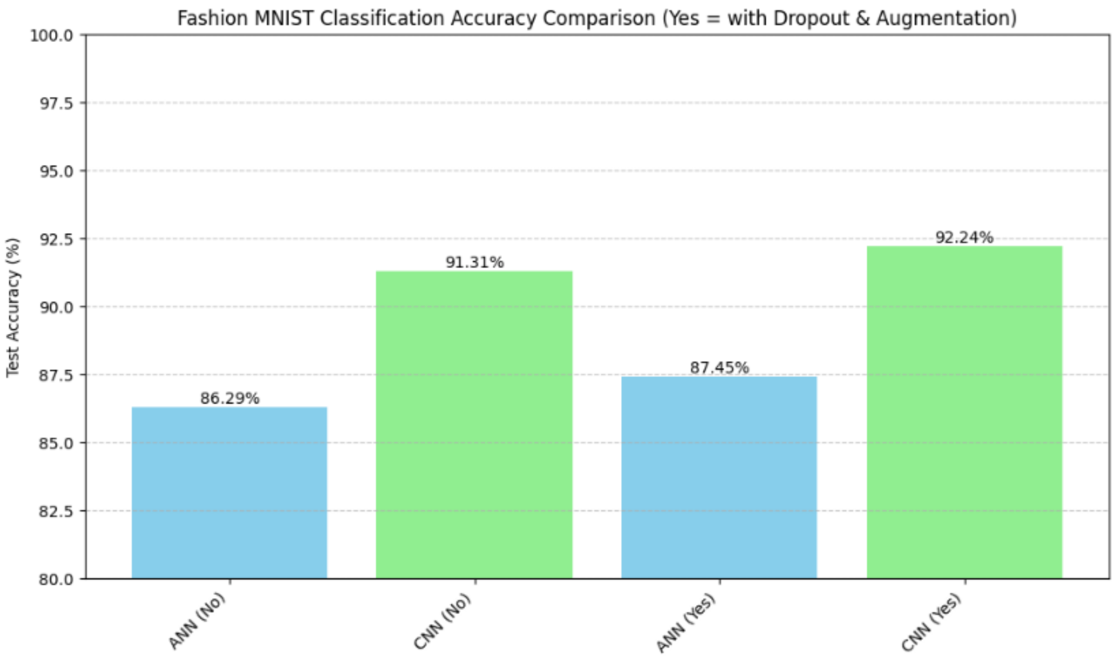
This document provides an in-depth study of the image classification on the Fashion MNIST dataset, consisting of 70,000 gray-scale images divided into 10 categories of fashion. The main goal was to create and assess two different neural network models: an Artificial Neural Network (ANN) and a Convolutional Neural Network (CNN); thus, the better one is chosen based on the performance in accurately distinguishing these fashion items.

Implementation Process:

1. The Fashion MNIST dataset was accessed using torchvision.datasets.FashionMNIST, and a transformation pipeline was applied to convert the images into tensors and normalize their pixel values. To ensure robust model training, the original training set was split into training and validation subsets using an 80/20 ratio. Data augmentation techniques such as random horizontal flipping and rotation were applied to the training data to improve generalization and reduce overfitting. Data loaders were then created for the training, validation, and test sets, each configured with a batch size of 64 to facilitate efficient mini-batch training.
2. The two model architectures have been implemented in different ways. An ANN model was created as a simple feedforward neural network composed of three linear layers followed by ReLU activation functions. Two dropout layers after the first two linear layers aided in reducing overfitting. On the other hand, the CNN model consisted of two convolutional layers max-pooling, ReLU, and connecting with each other. For the entire network's generalization, a dropout layer was placed before the fully connected layers. The different architecture features enabled the CNN to be more effective in identifying the spatial hierarchies in the image data than the ANN.
3. The Adam optimizer and CrossEntropyLoss function were used for training both models. Training lasted for five epochs, and a custom train_model function was responsible for the training loop. The function did the forward pass, computed the loss, backpropagated, and updated the optimizer.During the entire training, validation accuracy was the key player in observing the performance of the model as well as the signs of overfitting.
4. Models were evaluated with an evaluate function on a separate test set after training. This function calculated the overall classification accuracy and generated these confusion matrices to provide deeper insights into the performance of the models.

Results and Comparison:

The analysis findings illustrated that the CNN model was superior to the ANN model in all the situations. To be more precise, the ANN model without the application of dropout or augmentation received 86.29% accuracy, and the CNN model without any processing got 91.31%. The usage of dropout and data augmentation led to the ANN scoring 87.45%, and the CNN rising to 92.24%.



Observations:

1. Convolutional Neural Networks (CNNs) are considered the prime choice for image classification as they can learn complex and spatial features using the convolutional and pooling layers.
2. The use of dropout and data augmentation was a great advantage for both models in that they not only reduced overfitting but also helped to the increase in performance with new data.
3. The exceptional power of CNNs to distinguish between visually close classes was obvious, thus the case of shirts, pullovers, and coats.
4. Pervasive through the analysis of confusion matrix, clearly, CNNs had a lot fewer misclassifications than ANNs, especially among similar fashion categories.
5. The confusion matrices for ANNs displayed high off-diagonal values, which was a sign of frequent errors and the lack of understanding of the spatial features.
6. It's the convolutional layers in CNNs that primarily acted as the key barriers to class confusion and facilitators of model generalization.
7. In conclusion, CNNs combined with regularization and augmentation techniques came out as the strongest trend for image classification, and the confusion matrix results confirmed this.

Conclusion:

These results emphasize the dominance of CNNs in the field of image classification. The main reason for this is the CNN's use of convolutional and pooling layers that result in the learning of hierarchical representations, which in turn allow it to discover more relevant patterns compared to a plain feedforward ANN. Besides, dropout and data augmentation inclusion was advantageous for both models since it became possible to lessen overfitting and to increase the ability to correct the unseen data. The CNN was able to tell the difference between classes that were visually similar, such as shirts, pullovers, and coats, to a greater extent.