NS EXECUTIVE SUMMARY

NAME: SOWMYA MANNAM ID: P2939005

INTRODUCTION

On the very first page of the coursework, the writer gives an insight into the research work on the use of Artificial Neural Networks (ANNs) and Convolutional Neural Networks (CNNs) for the classification of images through the Fashion-MNIST dataset. The Fashion-MNIST dataset consists of ten different fashion items represented by a collection of 28x28 grayscale images. The objective is to develop models that will be able to classify these images with high accuracy. The entire process is meanwhile going on in the notebook which consists of data loading to model definition and training, through evaluation to comparing the two model architectures.

MODELS

The subsequent are the two dominant models' structures that are carried out and evaluated:
Artificial Neural Network (ANN):

- A simple feedforward neural network with three linear layers.
- The input layer accepts the flattened 28x28 image (784 features).
- Two hidden layers of 128 and 64 neurons respectively, with ReLU activation functions.
- The output layer is composed of 10 neurons, with one for every fashion class among the 10 classes.
- To prevent overfitting, a Dropout layer with a 0.5 dropout rate has been introduced after the second hidden layer of the Improved ANN.

Convolutional Neural Network (CNN):

- The idea of maxing downpooling refers to containing only the maximum value in the used filter shape over all the values in the convoluted filter shape.
- The architecture has two convolutional layers of 32 and 64 output channels respectively using 3x3 kernels and ReLU activation.
- Max pooling operation is performed after every convolutional layer to decrease the size of the feature maps.
- Then the feature maps will be turned into a 1D vector by the flatten layer to be fed into the fully connected layers.
- Two fully connected layers are used for classification purposes.
- An Enhanced CNN is also applied with sequential blocks for better architecture and a Dropout layer of a 0.5 dropout rate in the classifier section to fight overfitting. The size of the input to the first linear layer in the upgraded CNN is derived from the output sizes of the convolutional and pooling layers (64 * 7 * 7).

TRAINING AND TESTING

- **Dataset:** The Fashion-MNIST dataset is employed, divided into training, validation, and test sets. Data augmentation (random horizontal flips and rotations) is performed on the training data of the better models to multiply the data, thus leading to better generalization.
- **Loss Function:** The loss function is nn.CrossEntropyLoss which is the best for multi-class classification problems.
- **Optimizer:** The Adam optimizer is the one that has a learning rate of 0.001 and is used for the model weights update during the training process.
- **Training Process:** Each of the models is trained for a pre-defined number of epochs (10 for the initial, 15 for the improved models). At the end of each epoch, the training loss and validation accuracy are monitored to track the learning process and to uncover possible overfitting.
- **Testing Process:** The models after training are assessed on the independent test set to measure their performance on data that has not been seen before.

RESULTS

The following test accuracies were obtained for each model:

- Initial ANN Test Accuracy: 86.85 %
- Initial CNN Test Accuracy: 90.41 %
- Improved ANN Test Accuracy: 86.33 %
- Improved CNN Test Accuracy: 91.36 %

EVALUATION AND CONCLUSION

According to the test accuracies, the CNN models reliably exceeded the performance of the ANN models. The Improved CNN reached a peak test accuracy of 91.36%, which shows that the convolutional architecture is very effective for this image classification task. The training with data augmentation and the dropout layers in the improved models have also contributed to the generalization and have thus made the models perform better on the test set, especially for the CNN model.The confusion matrices that were created for each model serve as a diagram that shows the classification performance of the classification of each class providing a basis for the comparison of the correctly classified fashion items and those mostly confused with others. The confusion matrices have revealed the following points of interest:

- The two models get right the easily distinguishable items such as 'Trouser', 'Sandal', and 'Bag'.
- The visually similar classes 'T-shirt/top', 'Pullover', and 'Shirt' are more likely to be confused with one another by both models.
- The improved CNN has reduced the misclassifications of these similar classes compared to the initial models and the ANN models.

The result is that the Improved CNN is a very good performer on the Fashion-MNIST dataset and excels at the challenging task of separating visually similar fashion items.