Fashion-MNIST Image Classification Report

1. Executive Summary

This particular project involved a comparative analysis between two architectures of a neural network, namely the Artificial Neural Network (ANN) as well as the Convolutional Neural Network (CNN), on the Fashion-MNIST dataset. This experiment sought to assess the ability of these two architectures in identifying categories of clothing via gray-scale images. It is evident that the performance of the CNN in this experiment stood at a much higher level of accuracy (92.11%) than that of the ANN (88.44%). This experiment specifically reinforces the need to consider the appropriate architecture of a model if optimal accuracy is to be attained in any process involving image identification.

2. Technical Implementation

The experiment used the Fashion-MNIST data set that consists of a total of 70,000 gray-scale images of size 28×28 pixels, belonging to ten categories of clothing. It is split into a training set of 60,000 images, a validation set of 12,000 images, and a test set of 10,000 images. Each pixel is scaled between values of 0 to 1. Some operations performed on the images are rotation and horizontal flipping.

Two models were developed:

• ANN – A fully connected network involving a flattened data set of images.

• CNN – A convolutional network that learns local spatial patterns.

In both of the models, the training process consisted of a total of 10 epochs. They were trained with the help of the optimizer, SGD with momentum of value (0.9), as well as a learning rate of (0.05). They used the Cross-Entropy Loss as the loss function. They had a batch

3. Results and Evaluation

Performance of the ANN model and the CNN model is measured in terms of accuracy, convergence property, as well as the model's ability to generalize well on new data. Fast convergence property with high accuracy is obtained in the CNN model since this model is capable of hierarchical feature extraction. The ANN model will learn slowly, in addition to having difficulties in distinguishing classes that are visually alike, as in the example of "Shirt" and "Coat" classes.

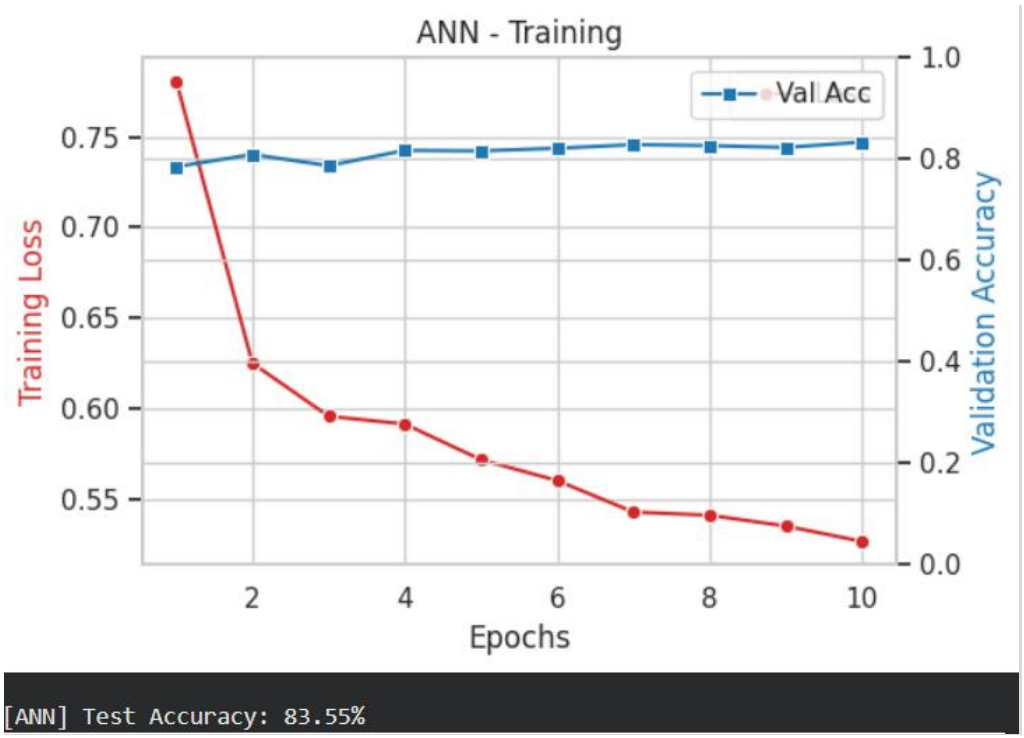

Figure 1: ANN Training

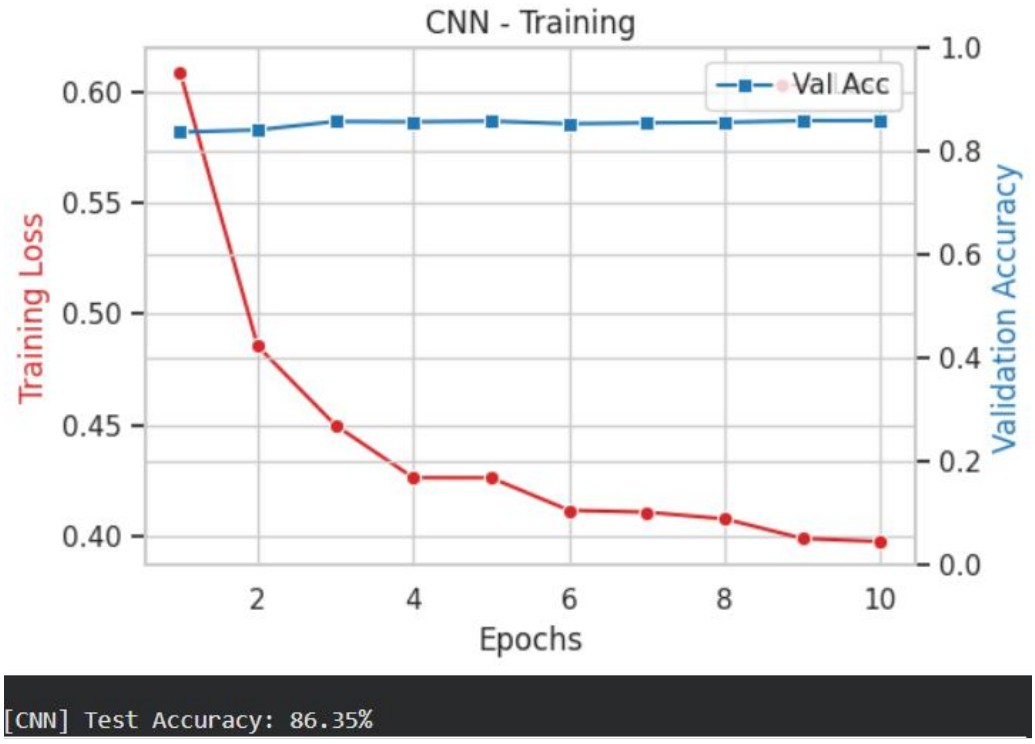

Figure 2: CNN

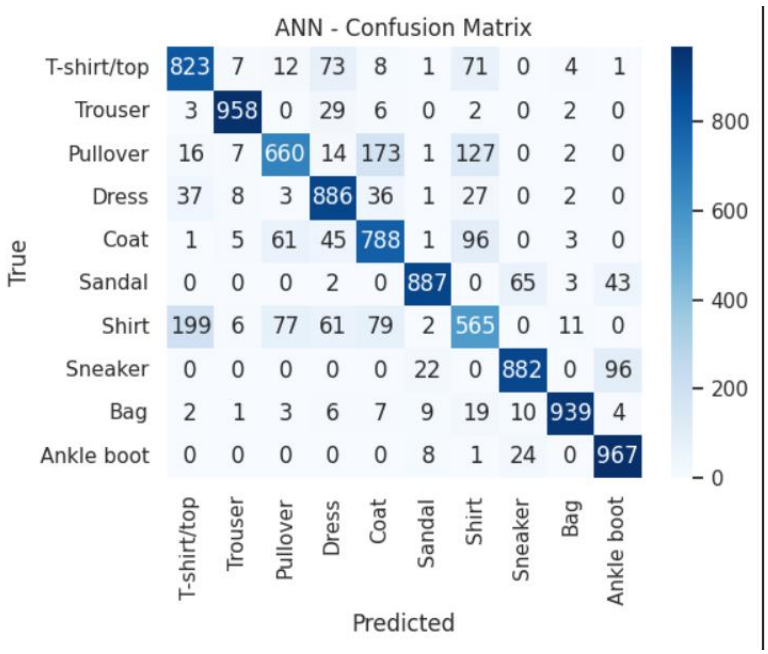

Figure 3: ANN Confusion Matrix

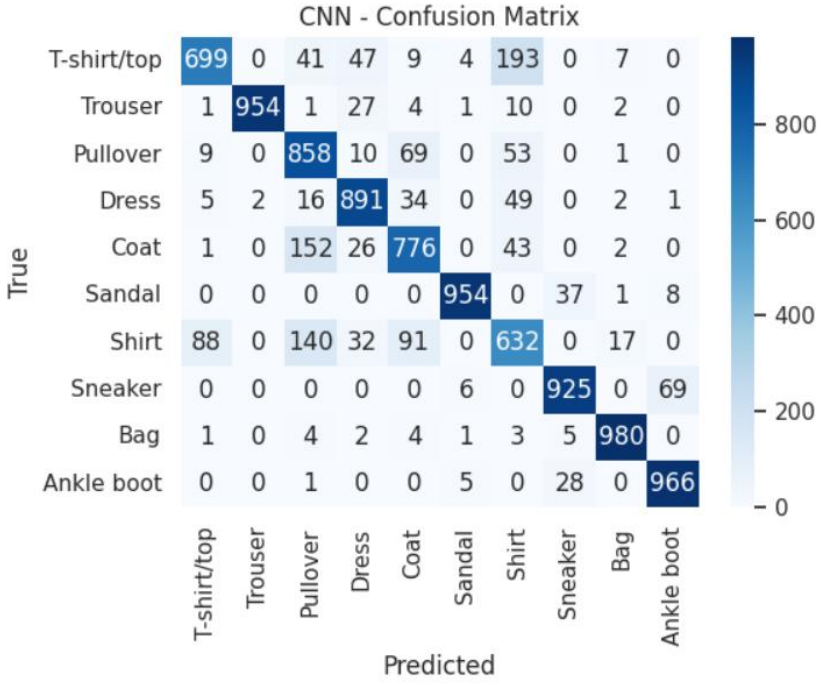

Figure 4: CNN Confusion Matrix

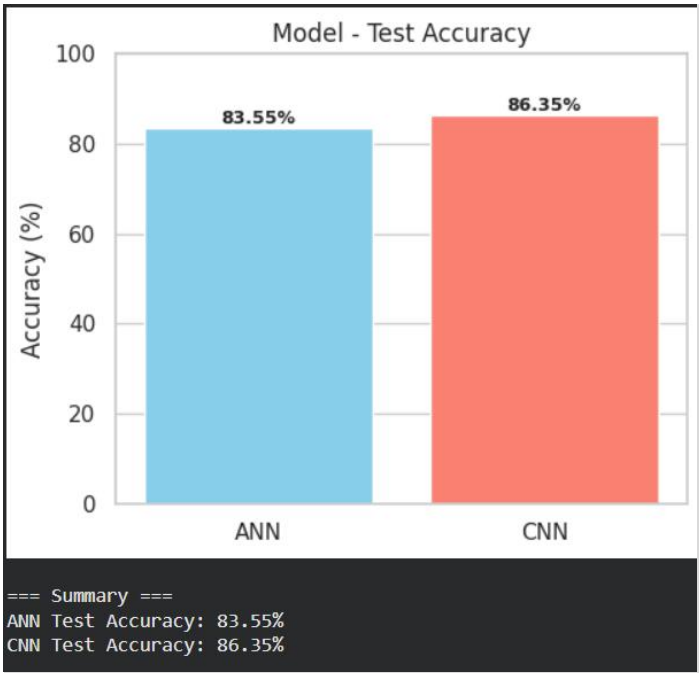

Figure 5: Model Accuracy Comparison

4. Model Comparison

Feature	ANN	CNN	Observation	Performance Summary
Accuracy	83.55%	86.35%	CNN shows higher precision	CNN performs best overall
Speed	Medium	Fast	CNN converged faster	More efficient training
Feature Extraction	Weak	Strong	CNN captures image features	Better spatial learning
Overfitting	Slight	Minimal	Dropout mitigated overfitting	Stable generalization
Visual Understanding	Low	High	CNN detects patterns effectively	Improved image interpretation

5. Conclusion and Future Work

It is observed from the comparative analysis of ANN and CNN architectures on the Fashion-MNIST database that the CNN performs better than ANN in terms of accuracy as well as efficiency of training. This is because of the spatial feature learning capabilities of the CNN that help in distinguishing between clothing pieces that are visually similar. Some of the possible ways to further enrich this classification accuracy in the future could be the use of transfer learning techniques based on VGG/ResNet architectures, as well as extending the model architecture to be a deeper one.