### Deliverable 1: Executive Summary

**Executive Summary: ANN vs. CNN Performance on Fashion-MNIST**

**Introduction**

This project implements two classes of neural network models, which include an Artificial Neural Network (ANN) and a Convolutional Neural Network (CNN), and applied them to the Fashion-MNIST dataset. It aimed at establishing which type of model performs best and to estimate the performance architectural benefits of a spatially-aware network (CNN) as compared to a simple dense network (ANN) on a typical computer vision problem. The scientific community should find this comparison noteworthy because it offers an easy to understand practical example of why image data cannot be processed by any standard architecture but rather requires specialized architectures such as CNNs, instead of the theoretical benefits that such architectures may offer, to actual performance measures.

**Experimental Methodology**

All the designs were measured based on their efficiency by implementing, training, and evaluating the models in PyTorch. Fashion-MNIST images were loaded and normalized, Fashion-MNIST is an image dataset of 70,000 28x28 grayscale images of clothing. The 60,000 training images were divided into a 50,000-image training image and a 10,000-image validation image. Two layer ANN (MLP) (128, 64 units) and two layer CNN (32, 64 filters) were constructed. The two models were trained on 10 epochs with the Adam optimizer and Cross-Entropy Loss, and all the calculations were run faster on a CUDA-based GPU. The results of the performance were gathered and were in form of loss/accuracy curves and confusion matrices which are provided in the Appendix.

**Results**

The CNN model performed the most successfully on the Fashion-MNIST dataset with the ultimate test accuracy of 91.22%. ANN model was extremely inferior, with the ultimate test accuracy of just 87.59%. The CNN was also advanced in that its convolutional layers are strongly designed so that they can extract the spatial hierarchy and local structure (edges, textures) in the images, which the "flat" ANN architecture does not have at all. One of the significant mistakes made by both models was a mix up between the similar classes on their appearance. As an example, the ANN falsely classified 148 "Shirts" as "T-shirt/top" and the better CNN falsely classified 93. The complete performance graphs and confusion matrices can be found in the Appendix.

**Conclusion**

Using the difference of final accuracy of 3.63 percent, one can say that CNN is clearly the better and more suitable architecture in this case of image classification. To further enhance the work of CNN, it should be outlined that in the future, one should work on the introduction of data augmentation (e.g., random flips and rotations) to minimize overfitting and consider more complex and efficient architectures (e.g., ResNet) that can learn more distinguishing features.

**References**

* PyTorch Documentation
* Fashion-MNIST Dataset