

AML Assignment-1

Name	ID
Ananya Sinha	MDS202307
Divyanshi Kumari	MDS202322
Rohit Roy	MDS202340

Question 1

FashionMNIST - CNN

Mapped all the given labels to these three: Clothes, Shoes, and Others.

We then created a CNN architecture which consists of three convolutional layers, each followed by ReLU activation and max pooling, which help extract features from the input images. A dropout layer is included to prevent overfitting. The model then has three fully connected layers, where the final layer outputs class scores for three categories using a log-softmax activation function.

This architecture gives an accuracy of 99.55% for the test dataset and an accuracy of 99.26% for scrambled images.

Question 2

ResNet-18 Implemented from Scratch

We initialize a ResNet-18 model without pre-trained weights and modify its final layer to match the number of output classes in the dataset.

A cross-entropy loss function is defined for multi-class classification, and an Adam optimizer is initialized with a learning rate of 0.001. Additionally, a learning rate scheduler is set up to decrease the learning rate by a factor of 0.1 every 5 epochs.

This model obtained an accuracy of 62.72% on the test set.

Fine-tuning on Pre-trained ResNet-18

We initialized a ResNet-18 model with pre-trained weights from the ImageNet dataset and modified its final layer to match the number of classes in the dataset. Initially, all parameters are set to non-trainable to prevent updating the weights during training. Selected parameters of the final fully connected layer, and the last convolutional block are unfrozen to allow for fine-tuning. An Adam optimizer is defined specifically for the trainable parameters, along with an initial learning rate of 0.001 and a learning rate scheduler that reduces the learning rate every 5 epochs by a factor of 0.1.

This model obtained an accuracy of 66.27% on the test set.

Visualising the Difference between the two Models

The function visualizes the predictions from two models on images from a test data loader. This allows us to visually compare the prediction probabilities of both models.