CS 250 Deep Learning

Report of Assignment 2 : Fashion-MNIST

Manohar Lal – 16896 (CSA)

Deep Neural Network:

Test accuracy of this model is 82.27% and training accuracy is 84.596%.

Each image input was 28 x 28 flattened into 784 pixels. The model I used had two hidden layers consisting of 200, 100 neurons and ReLu activation function. The optimizer used was Adam optimizer as it gave better accuracy. The output layer obviously had 10 outputs and used Softmax activation as these are standard practices.

Confusion Matrix:

[618, 0, 0, 331, 0, 0, 48, 0, 3, 0]
[14, 221, 0, 760, 0, 0, 5, 0, 0, 0]
[10, 0, 69, 399, 0, 0, 522, 0, 0, 0]
[8, 0, 0, 979, 0, 0, 15, 0, 0, 0]
[1, 0, 17, 350, 0, 0, 630, 0, 2, 0]
[1, 0, 0, 23, 0, 960, 8, 5, 3, 0]
[116, 0, 1, 311, 0, 0, 571, 0, 1, 0]
[1, 0, 0, 131, 0, 245, 0, 620, 3, 0]
[82, 0, 0, 90, 0, 13, 136, 1, 678, 0]
[0, 0, 0, 35, 0, 670, 1, 259, 11, 24]

Convolution Neural Network:

The CNN model I submitted produces accuracy of 90.88% in test set and 96.80% on training set.

For the convolutional layer I use a small filter size (3,3) and number of kernels is (32) followed by a max pooling layer. The filter maps is then flattened to provide features to the classifier.

I know that I will require an output layer with 10 nodes in order to predict the probability distribution of an image belonging to each of the 10 classes. This will also require the use of a softmax activation function. Between the feature extractor

and the output layer, I add a dense layer to interpret the features, in this case with 100 nodes.

All layers will use the ReLu activation function and the He weight initialization scheme.

I used stochastic gradient descent optimizer with a learning rate of 0.01 and a momentum of 0.9. The categorical cross-entropy loss function is used, as it is suitable for multi-class classification.