

Machine Learning Final project

In this project, we have a neural network classification problem using MNIST fashion dataset using a committee of three different neural networks. Neural network itself does not work very well because it has some limitations, for example it is based on how the data is initialized and also the chosen architecture. Taking 28×28 images and outputs to which class belong. By changing the architecture and the hyperparameters of the neural networks, we construct the committee and make sure the variance is not very high. We use Keras, a high-level Deep Learning API that makes it easy to build, train, evaluate and execute all sorts of neural networks. After loading MNIST dataset, and since we are going to train the neural network using Gradient Descent, we must scale the input features.

Splitting the data:

All except the last 5000 records are chosen to validate after normalize them by dividing them by 255, and the first 5000 records are for the test.

Reshape the train, valid and test sets to (55000,784), (5000,784) and (10000,784) respectively.

The committee analysis:

Before deciding this committee, I used 2 different once but I kept getting very low average accuracy.

In the first model of the neural network, I added 2 dense layers with dropout. I used the activation function *relu* in the first one and *softmax* in the second. With *adam* optimizer and 40 epochs, the network validation accuracy fluctuated back and forth a little bit between %90 and %89 and the resulting accuracy is %86.

In the second neural network, 4 dense layers were added with *relu* activation function in three of them and *softmax* in the last one. With *sgd* optimizer and 50 epochs, the network slowly progress towards higher validation accuracy of %89. The resulting accuracy is %85.

In the last neural network, 2 dense layers without dropout were added with *relu* activation function in the first one and *softmax* in the other one. Using *adam* optimizer and 100 epochs, the validation accuracy increases from %85 to %89 and the resulting accuracy is %80.

The average accuracy for the three neural networks is as high as %84 since each one of them is more than %80.