Configuration File For MLP

The Original Configuration:

Number of Epochs	**10**
Batch Size	**128**
Number of Neuron in a Layer	**512**
Number of Layers	**3**
Learning Rate	**0.001**
Activation Function	**[relu, softmax]**
Dropout Rate	**0.2**

The Original Loss and Accuracy:

Test Loss	**1.59**
Test Accuracy	**43.62%**

While modifying those parameters of the MLP model, here are the effect of changing parameters on loss and accuracy:

Number of Epochs	**20**	**40**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**	**10**
Batch Size	**128**	**128**	**32**	**256**	**128**	**128**	**128**	**128**	**128**	**128**	**128**	**128**
Number of Neurons	**512**	**512**	**512**	**512**	**256**	**1024**	**512**	**512**	**512**	**512**	**512**	**512**
Number of Layers	**3**	**3**	**3**	**3**	**3**	**3**	**4**	**3**	**3**	**3**	**3**	**3**
Learning Rate	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**	**0.001**	**0.0001**	**0.001**	**0.001**	**0.001**	**0.001**
Activation Function	**[relu, softmax]**	**[relu, softmax]**	**[relu, softmax]**	**[relu, softmax]**	**[relu, softmax]**	**[relu, softmax]**	**[relu, softmax]**	**[relu, softmax]**	**[tanh, softmax]**	**[relu, elu]**	**[relu, softmax]**	**[relu, softmax]**
Dropout	**0.2**	**0.2**	**0.2**	**0.2**	**0.2**	**0.2**	**0.2**	**0.2**	**0.2**	**0.2**	**0.1**	**0.5**

Rate												
Test Loss	**1.50**	**1.48**	**1.69**	**1.59**	**1.55**	**1.53**	**1.61**	**1.46**	**1.63**	**8.66**	**1.50**	**1.87**
Test Accuracy	**48.03%**	**48.95%**	**36.42%**	**42.47%**	**45.59%**	**47.55%**	**41.89%**	**46.98%**	**41.72%**	**29%**	**46.32%**	**35.99%**

Results:

1. Number of Epochs:
   One epoch stands for one forward pass and one backward pass of all the training examples. The more epochs, the better model performance;
2. Batch Size:
   Batch size is the number of training examples in one forward/backward pass. The higher the batch size, the more memory space we will need;
3. Number of Neurons:
   The more neurons in a layer, the better model performance;
4. Number of Layers:
   Higher than 3 layers, the model performance will decrease;
5. Learning Rate:
   The lower the learning rate, the better model performance;
6. Activation Function:
   Where the element-wise activation function is ‘relu’, and the last step activation function is ‘softmax’ resulted in a better model performance;
7. Dropout Rate:
   The lower dropout rate, the better model performance.

The Best Model Configuration During My Test:

Number of Epochs	**40**
Batch Size	**128**
Number of Neuron in a Layer	**1024**
Number of Layers	**3**
Learning Rate	**0.0001**
Activation Function	**[relu, softmax]**

Dropout Rate	**0.1**

The Loss and Accuracy With the Best Model:

Test Loss	**1.33**
Test Accuracy	**55.61%**