

Configuration	mean loss	mean accuracy
2 hidden layer, hidden units=64, loss function = categorical_crossentropy, activation function = relu, epochs = 8, batch_size = 512	1.0039258003234863	0.784060537815094
1 hidden layer	0.9131276607513428	0.790739119052887 (The best one)
3 hidden layer	1.0821723937988281	0.7653606534004211
hidden units = 32	1.1042351722717285	0.7622439861297607
hidden units = 128	0.9981028437614441	0.7996438145637512

