機器學習概論 Introduction to Machine Learning
Assignment 3 – Neural Network
Department: 理學院學士班
Student ID: 108020017 Name: 黃珮綺
-

1. Problems I encountered

Sometimes I got troubled with the indexing of an array and found it hard to tell what functions to use. Take the ReLU function in activation forward for example, here I use np.maximum() instead of np.max(). Although the two function literally seem to be identical, they work in a slightly different way. The former returns the maximum value of the single input array, while the latter takes two arrays and returns the element-wise maximum.

To solve this kind of problems, I often searched on the internet to find what should I do to better write the code. Last but not least, I also discuss with classmates and figure out how to simplify our code together (Of course, we prevent plagiarism!).

2. The structure of my binary and multi-class classifiers

For the binary classifier, I use four layers with dimension 30 for the input, 15 and 10 for the two hidden layers, and 1 for the output. The activation functions are sigmoid function for the last layer and ReLU functions for the rest of the layers. It will iterate 3001 times with learning rate 0.1.

For the multi-class classifiers, I flattened 28x28 pixels to an array with 784 elements for each data, hence the dimension of the first layer of the classifier is 784. The dimensions of the two following layers are 100 and 50, and the final output will be of dimension 10. The activation functions are two ReLU functions and a softmax function. It will iterate 151 times with learning rate 0.08, and the batch size is 70.

3. Effort I put to improve my model

I’ve tried to iterate on the dimensions of each layers in the basic part (The source code is not included in the submitted file). But later on, I found that the maximum accuracy is limited to 0.97, no matter how much dimensions I use in the model. I again checked my code, and then modified the min-max scaling of X. I originally normalized them based on the global maximum and minimum of X, however replaced them with the maximum and minimum of each feature respectively. Finally, the accuracy could reach 0.98. As for the advanced part, I manually adjusted the dimensions and number of iterations to get higher accuracy.