# Solution to the Open Question, Homework 2

Guangyu Li, Xiaohan Wang

January 10, 2023

**b.** Before the experiment, data normalization is applied to the dataset for training efficiency. Here, we use the Min-Max normalization method, which standardizes features by removing the minimal values and scaling to the range of data. Particularly, for calculating the normalized data, the formula is

$$x_j \leftarrow \frac{x_j - \min(x_j)}{\max(x_j) - \min(x_j)}$$

The first 4 samples of the training data is given by

```
>>>x_train[:4,:], y_train[:4]
(array([[0.29091901, 0.32840185],
        [0.18739645, 0.07590681],
        [0.41890954, 0.15372418],
        [0.81728824, 0.98426067]]),
 array([0., 0., 0., 0.]))
```

**c.** The MLP model (Figure 1) for binary classification consists of four layers, with two hidden layers, each made up of 10 neurons, and an output layers of two neurons indicating the 1 or 0 classes. The input layer has two nodes, as the training data contains two features `X_1`, `X_2`.

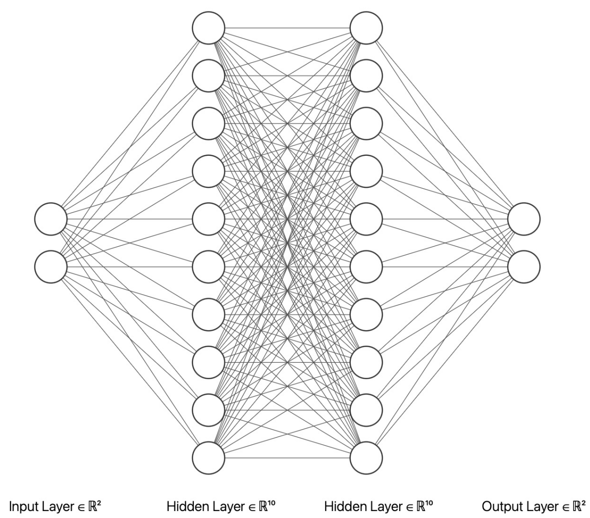


Figure 1: MLP model structure

**d.** For both hidden layer, we use `ReLU` as activation function. In addition, we use softmax as loss function in the output layer since it is a classification task.

**e.** Parameter values are initialized using Gaussian distribution.

**f.** MLP network implementation. Please refer to our code in `MLP.ipynb`.

**g.** In our implementation, several hyperparameters can be adjusted, such as the number of epochs, learning rate, batch size, the weight scale of each neural, and the size of each hidden layer.

**h.** Training loss.

Stopping criteria: We set 100 train epochs, in which we go through the whole train dataset with a batch size 4.

**i.** The final accuracy of our model is 0.7.

**Workload Distribution** Xiaohan Wang is responsible for Question 1 and 2. Guangyu Li is responsible for Question 3, 4, and 5.