XI'AN JIAOTONG-LIVERPOOL UNIVERSITY
西 交 利 物 浦 大 学

YEAR 4
COURSE WORK SUBMISSION

Name	*Yuxuan*	*Wu*
ID Number	*1716309*	
Programme	*INT301*	
Module Title	*Biocomputation*	
Module Code	*INT303*	
Assignment Title	*Lab1*	
Submission Deadline	*Time of submission:12.20.2020*	
Lecturer Responsible		

I certify that:

- I have read and understood the University's definitions of COLLUSION and PLAGIARISM (available in the Student Handbook of Xi'an Jiaotong-Liverpool University).

With reference to these definitions, I certify that:

- I have not colluded with any other student in the preparation and production of this work;
- this document has been written solely by me and in my own words except where I have clearly indicated and acknowledged that I have quoted or used figures from published or unpublished sources (including the web);
- where appropriate, I have provided an honest statement of the contributions made to my work by other people including technical and other support staff.

I understand that unauthorised collusion and the incorporation of material from other works without acknowledgement (plagiarism) are serious disciplinary offences.

Signature …….…吴宇轩……………………… Date …………………*11.26.2020*……………………

For Academic Office use:	Date Received	Days Late	Penalty

Introduction

Multilayer Perceptrons (MLPs), one of the classical types of artificial neural networks, are widely used in classification prediction problems where inputs are assigned into classes or labels [1]. The simple network is composed of three components: input layer, hidden layer and output layer, and with the back-propagation training algorithms, the model could automatically update each weight until convergence. Hence, in this project, I would design a simple MLP model with sigmoid activation function to conduct the vehicle logo classification. Additionally, I tuned and returned the hyperparameters to achieve the minimum mean squared error (MSE) model and used the confusion matrix to compare the best MLP model.

Methodology

The provided dataset contains 354 samples with 168 dimensions and the corresponding labels with six categories. For the purpose of training and testing, I split the dataset into a training set (80%) and a testing set (20%). In addition, min-max normalization was followed to minimize the potential outliers' influence. In this experiment, the MLP model contains three hidden layers with sigmoid activation function in both hidden and output layers (Figure 1). Generally, hyperparameters like the number of hidden units, learning rate and momentum could have a significant influence on the model performance. Therefore, I used the grid search method to loop through the hyperparameters mentioned previously and return the hyperparameters which could minimize the model's MSE value. The testing sample was then fed into the model to simulate the prediction results and undergo the reverse normalization process to return the predicted labels or categories. The density plots of testing labels and predicted labels were visualized to illustrate the distribution and the confusion matrix was applied to evaluate the prediction performance. Moreover, I compared the effect of different numbers of hidden units by fixing the learning rate and momentum, meanwhile, the effect of learning rate and momentum was assessed by fixing the hidden units.

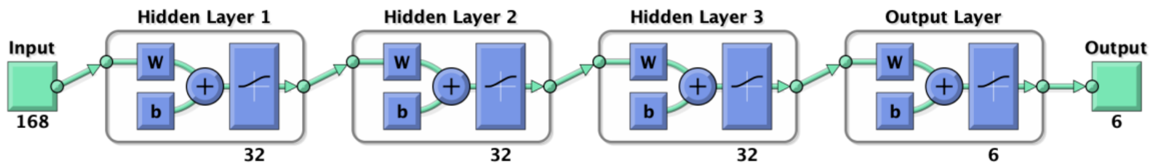
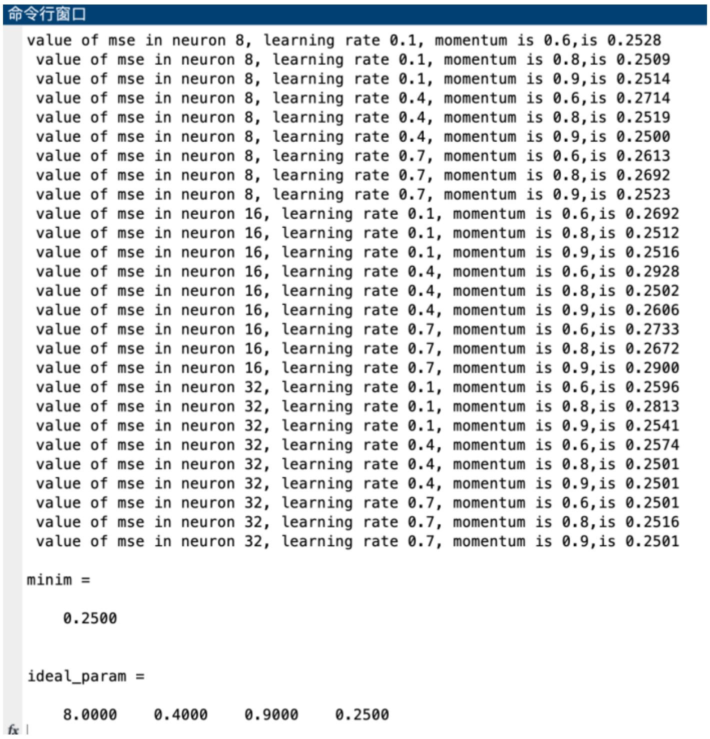


Figure 1: Components of custom MLP

Experimental results and analysis

According to Figure 2, I printed out each MSE value of the model in a different combination of hyperparameters. As can be seen from the figure, when the hidden layers contain 8 neurons, learning rate set to 0.4 and momentum set to 0.9, the model could achieve the best performance, with MSE value in 0.2500.



```
命令行窗口
value of mse in neuron 8, learning rate 0.1, momentum is 0.6,is 0.2528
 value of mse in neuron 8, learning rate 0.1, momentum is 0.8,is 0.2509
 value of mse in neuron 8, learning rate 0.1, momentum is 0.9,is 0.2514
 value of mse in neuron 8, learning rate 0.4, momentum is 0.6,is 0.2714
 value of mse in neuron 8, learning rate 0.4, momentum is 0.8,is 0.2519
 value of mse in neuron 8, learning rate 0.4, momentum is 0.9,is 0.2500
 value of mse in neuron 8, learning rate 0.7, momentum is 0.6,is 0.2613
 value of mse in neuron 8, learning rate 0.7, momentum is 0.8,is 0.2692
 value of mse in neuron 8, learning rate 0.7, momentum is 0.9,is 0.2523
 value of mse in neuron 16, learning rate 0.1, momentum is 0.6,is 0.2692
 value of mse in neuron 16, learning rate 0.1, momentum is 0.8,is 0.2512
 value of mse in neuron 16, learning rate 0.1, momentum is 0.9,is 0.2516
 value of mse in neuron 16, learning rate 0.4, momentum is 0.6,is 0.2928
 value of mse in neuron 16, learning rate 0.4, momentum is 0.8,is 0.2502
 value of mse in neuron 16, learning rate 0.4, momentum is 0.9,is 0.2606
 value of mse in neuron 16, learning rate 0.7, momentum is 0.6,is 0.2733
 value of mse in neuron 16, learning rate 0.7, momentum is 0.8,is 0.2672
 value of mse in neuron 16, learning rate 0.7, momentum is 0.9,is 0.2900
 value of mse in neuron 32, learning rate 0.1, momentum is 0.6,is 0.2596
 value of mse in neuron 32, learning rate 0.1, momentum is 0.8,is 0.2813
 value of mse in neuron 32, learning rate 0.1, momentum is 0.9,is 0.2541
 value of mse in neuron 32, learning rate 0.4, momentum is 0.6,is 0.2574
 value of mse in neuron 32, learning rate 0.4, momentum is 0.8,is 0.2501
 value of mse in neuron 32, learning rate 0.4, momentum is 0.9,is 0.2501
 value of mse in neuron 32, learning rate 0.7, momentum is 0.6,is 0.2501
 value of mse in neuron 32, learning rate 0.7, momentum is 0.8,is 0.2516
 value of mse in neuron 32, learning rate 0.7, momentum is 0.9,is 0.2501

minim =

    0.2500


ideal_param =

    8.0000    0.4000    0.9000    0.2500
```

Figure 2: Grid search for the best model performance in hyperparameter tuning

Based on the optimal hyperparameters, I trained the best performance MLP model and simulated the labels on testing data. Figure 3 compares and visualizes the density plot between predicted labels and testing labels, the overall shape is similar, but the predicted labels seem to have a smaller width. We also made a confusion matrix to test the efficacy of MLP model's prediction (Figure 4).

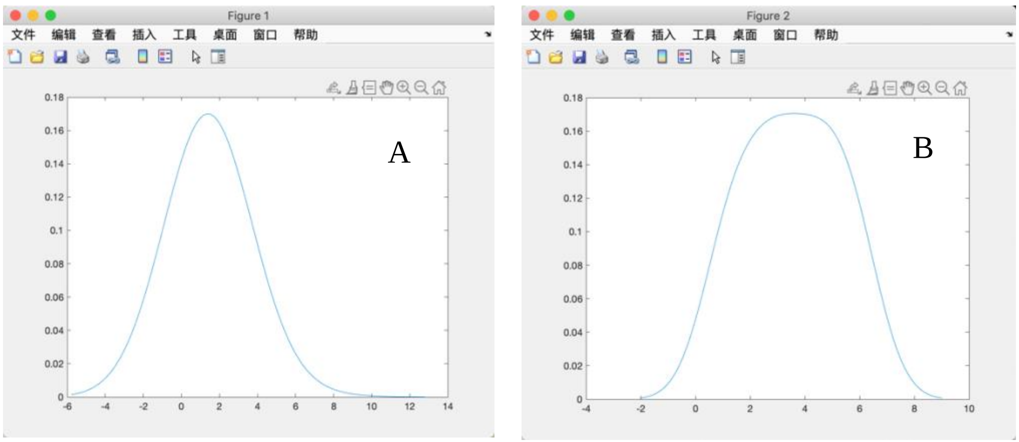


Figure 3: Density plots of predicted labels (A) and testing labels (B)

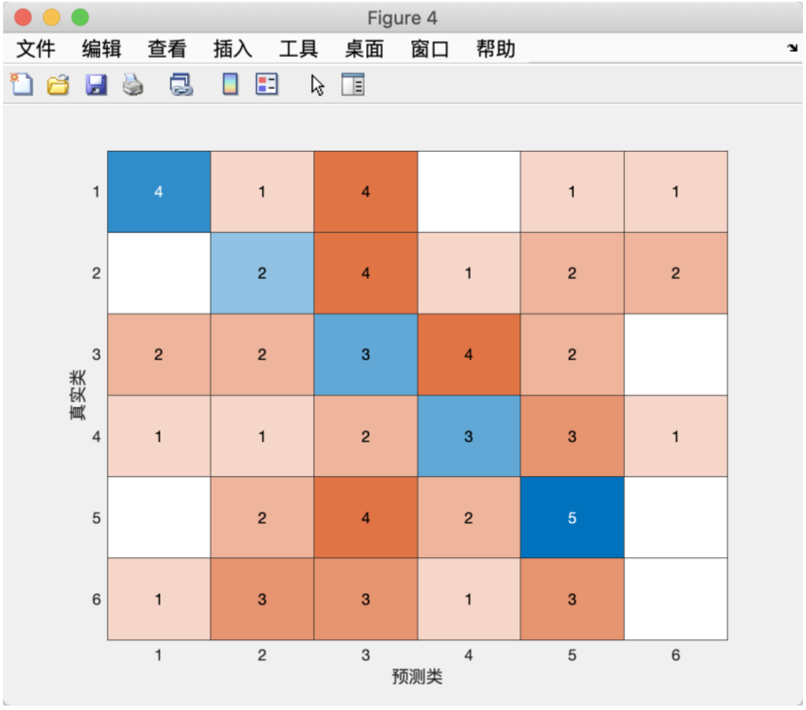


Figure 4: Confusion matrix of predicted labels and testing labels

Once we finished the finding of the best MLP model using the confusion matrix, we would like to demonstrate the effect of hyperparameters. By fixing the learning rate to 0.4, momentum to 0.9, the regarding MSE value was shown in Table 1. The overall MSE value decreases as the number of hidden units increases, but the higher numbers of neurons might lead to an overfitting problem, therefore it should limit into some ranges. Simultaneously, I fixed the neurons in each layer and adjusted the learning rate and momentum accordingly, it seemed that the patterns or trends were not clear to be seen.

Neurons	MSE
8	0.2810
16	0.2514
32	0.2504

Table 1: The effect of changing the number of neurons in each layer from 8, 16, 32 in MSE

Momentum / Learning rate	0.6	0.8
0.1	0.3662	0.2537
0.4	0.2512	0.2797

Table 2: The effect of changing the learning rate and momentum in MSE

Conclusion

In this experiment, we used the grid search to return the optimal hyperparameters, and the well-trained model could successfully predict the samples into different labels. However, the overall performance was not ideal enough, only 17 out of 70 were successfully predicted. Moreover, I adjusted the hyperparameters to see the corresponding performance change. In the future, I would tune the model and incorporate other algorithms to achieve high performance.

Reference

[1] "What is a multi-layered perceptron?" https://www.educative.io/edpresso/what-is-a-multi-layered-perceptron (accessed Nov. 26, 2020).