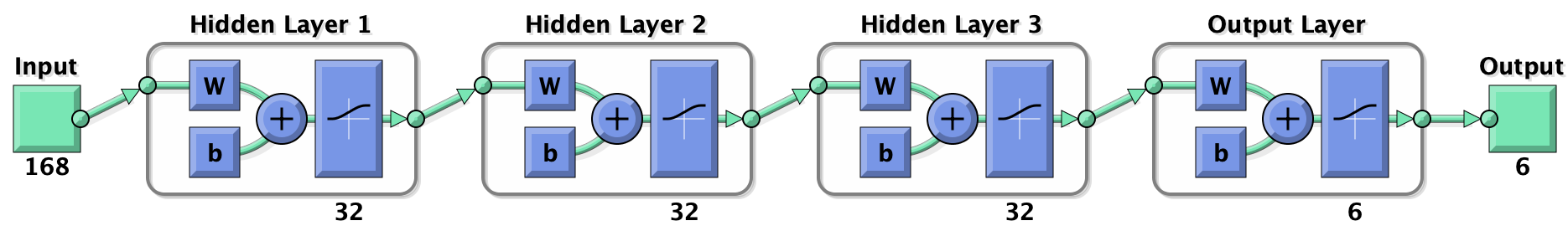
**Introduction**

Multilayer Perceptrons (MLPs), one of the classical types of artificial neural network, are widely used in classification prediction problems where inputs are assigned into classes or labels. The simple network 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 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 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 number of hidden units, learning rate and momentum could have significant influence of the model performance. Therefore, I used 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 confusion matrix was applied to evaluate the prediction performance.

Figure 1: Components of custom MLP



**Experimental results and analysis**

According to Figure 2, I printed out each MSE value of the model in 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 (Figure 2)

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

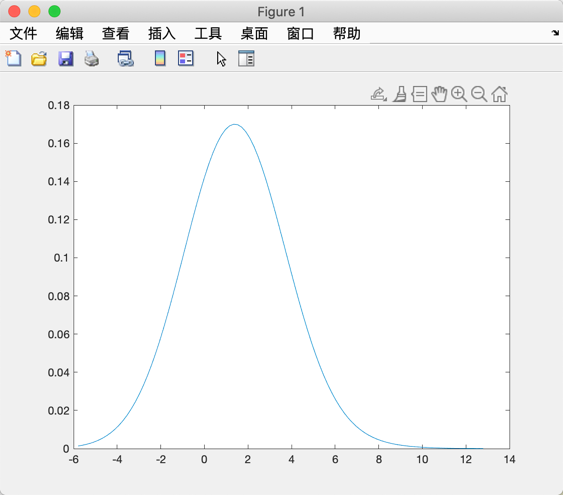
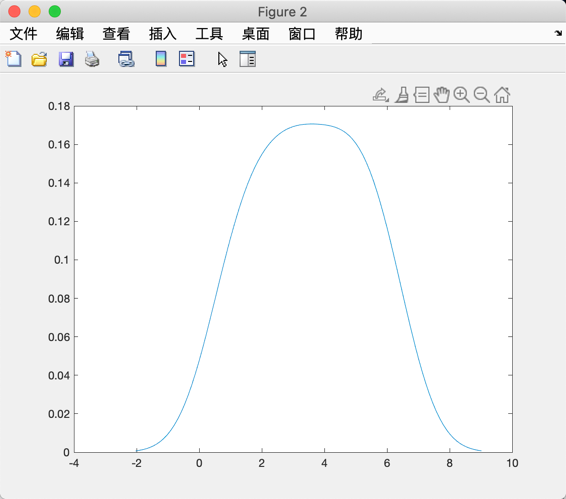
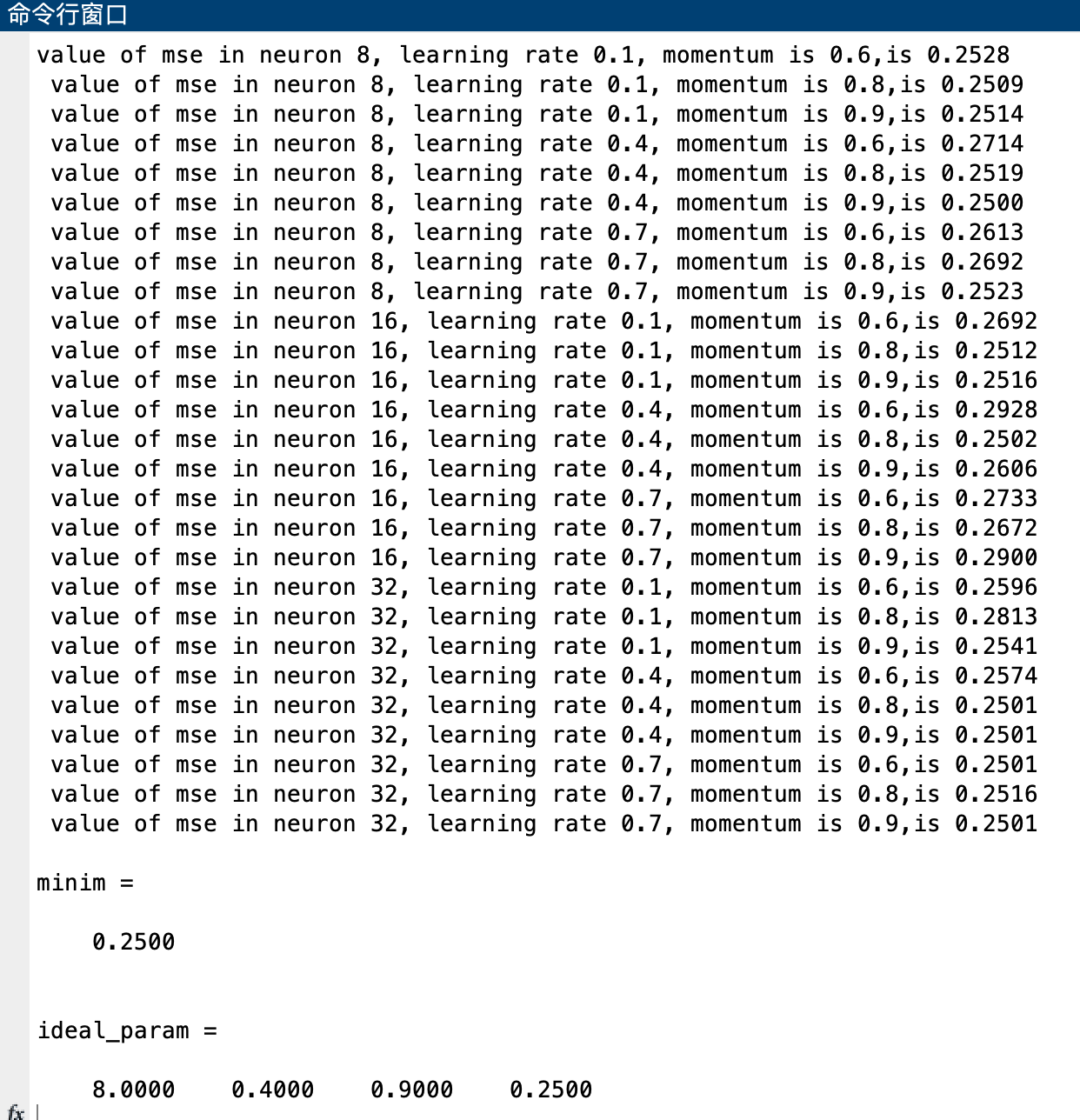


Figure 3: Density plots of predicted labels (A) and testing labels (B)

A

B

**Conclusion**

**Reference**

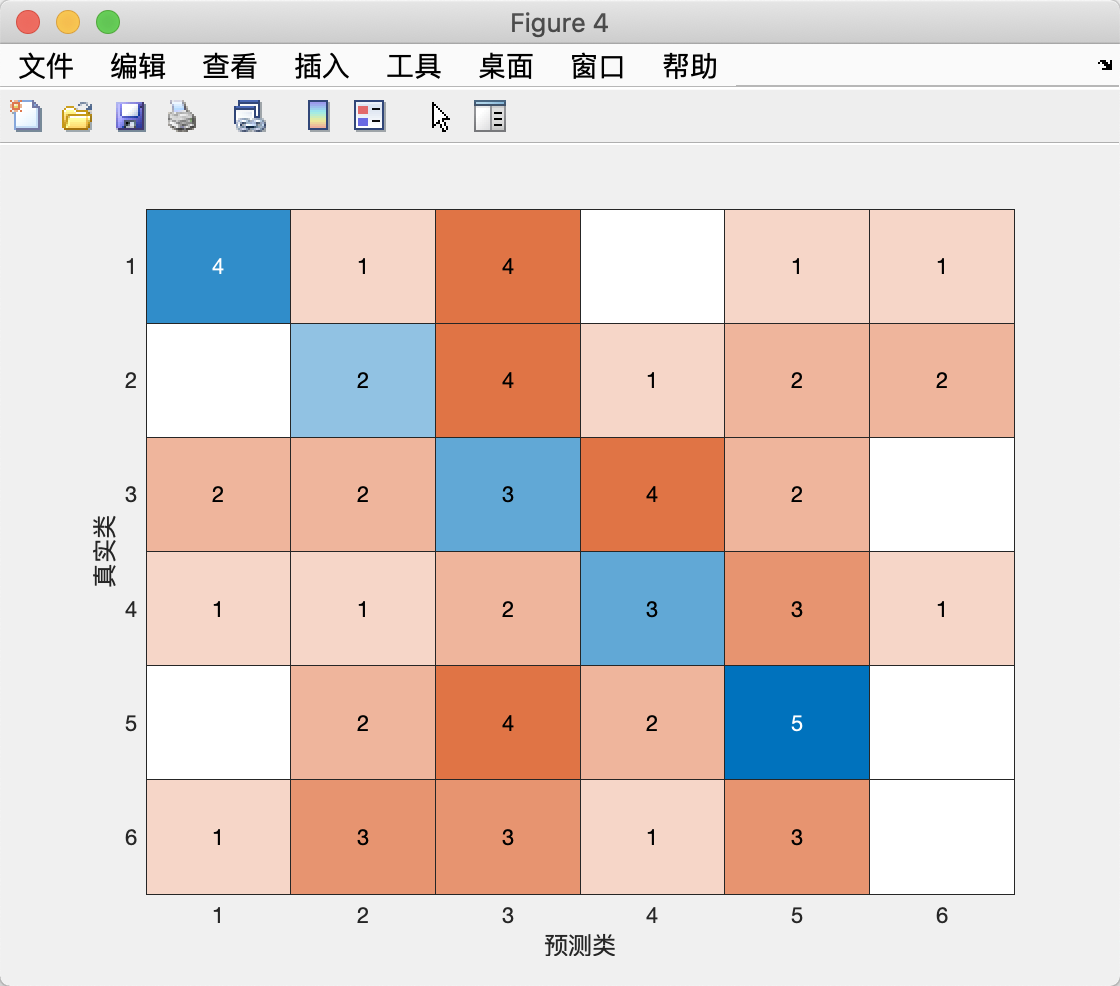


Figure 4: Confusion matrix of predicted labels and testing labels