HW8

B10209040 陳彥倫

8.1

"We implement the ANNs using Keras (Chollet, 2018). The ANN architecture consists of two hidden layers with 10 neurons each and using the rectified linear activation function (ReLU). Its final layer contains two neurons and a softmax activation. The ANN uses categorical cross entropy for loss function and is trained for 50 epochs with the SGD optimizer with a fixed learning rate of 0.001. We manually tune the number of neurons of the hidden layers and other hyperparameters such as the activation function and the number of epochs. The ANN with the current choices of hyperparameters produces high accuracy as shown in Table 2. Given that the goal of the study is to understand the relevance of the input features for the ANN to make its decision, we do not perform a thorough search of hyperparameters to obtain a perfect ANN model. We use the accuracy and F1 score as the performance metrics. The ANN is trained 60 times with random initialization and on different subset of data. Meanwhile, for each of the 60 experiments, we apply a logistic regression on the same data as a baseline."
Source: Wu et al. (2022)

The role of logistic regression:

A logistic regression model was used as a baseline to compare the ANN's performance, helping to validate the effectiveness of the ANN in handling the given task.

8.2

1.

We use logistic regression model to classify wheather a case is a Sudden Stratospheric Warming (SSW) or not. First create training set from 70% of dataset, and testing set from the rest. The input and output sizes of the model are 144 and 1 respectively.

2.

Training Loss: The measure of error or the discrepancy between the model's predictions and the actual labels during the training phase.

Training Accuracy: The percentage of correct predictions made by the model on the training dataset. It's a direct indicator of how well the model is learning from the data it is trained on. The smaller the better.

Testing Loss: Similar to training loss, this is the measure of error between the model's pr edictions and the actual labels on a new, unseen dataset (the test dataset).

Testing Accuracy: The percentage of correct predictions made by the model on the training dataset. It's a direct indicator of how well the model is learning from the data it is trained on. The smaller the better.

In this case, the training and testing accuracy kept increasing during training. However the train accuracy went up to around 0.9857 and stopped. We can also observe that errors became less as time went on.