DL HW1

人工智慧碩二 311505011 黃丰楷

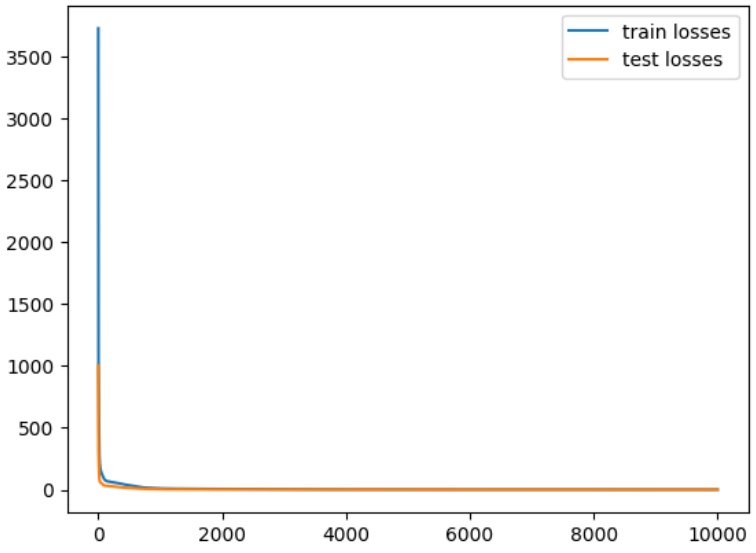
1. Regression

a. Network Architecture

layer name	output size	operations
fc1	8	fc [16,8], sigmoid
fc2	4	fc [8,4], sigmoid
fc3	1	fc [4,1]

b. Learning Curve

- Hyperparameters
  - epoch: 10000
  - batch size: 128
  - learning rate: 0.01
- Training/testing loss curve

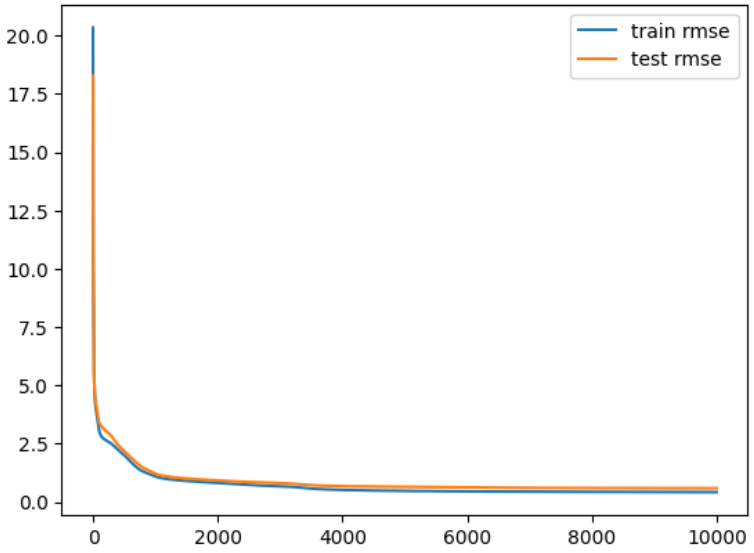


c. Training/Testing RMS error

- Training/testing RMSE

	RMSE
Training	0.4168
Testing	0.5859

- Training/testing RMSE curve



d. Feature Importance

- Method

  All 8 features were entered into the trained model to predict the heating load, and the features that were not selected in the process were brought in with a zero. In this way, the effect of each feature on predicting performance is tested, and the feature importance is analysed.

- Experiment

Selected feature	RMSE
Relative Compactness	10.3329
Surface Area	14.1452
Wall Area	11.3486

Selected feature	RMSE
Roof Area	8.2843
Overall Height	6.6521
Glazing Area	12.2029
Orientation	12.3862
Glazing Area Distribution	11.5541

- Feature importance result

  Overall Height > Roof Area > Relative Compactness > Wall Area > Glazing Area Distribution > Glazing Area > Orientation > Surface Area

1. Classification

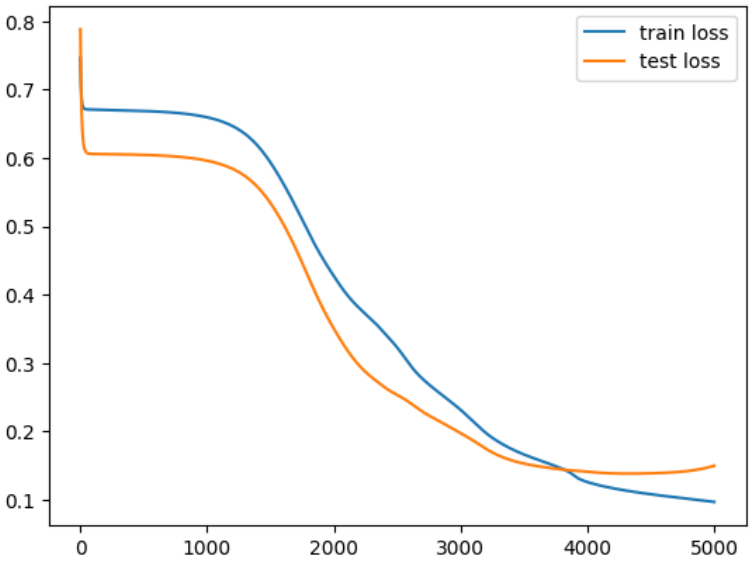
a. Network Architecture

layer name	output size	operations
fc1	4	fc [34,4], sigmoid
fc2	4	fc [4,4], sigmoid
fc3	2	fc [4,2], sigmoid
fc4	1	fc [2,1], sigmoid

b. Learning Curve

- Hyperparameters
  - epoch: 5000
  - batch size: 64
  - learning rate: 0.03

- Training/testing loss curve

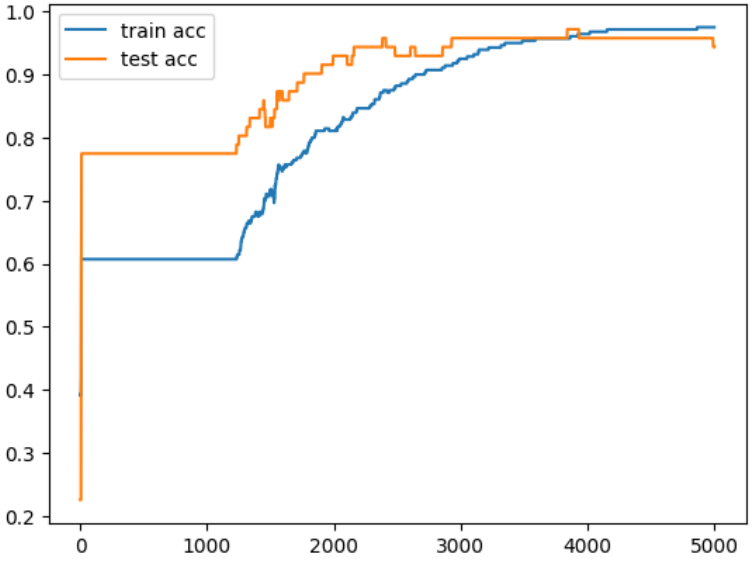


c. Training/Testing accuracy

- Training/testing accuracy

	Accuracy
Training	0.9786
Testing	0.9014

- Training/testing accuracy curve



d. Experiments — number of nodes

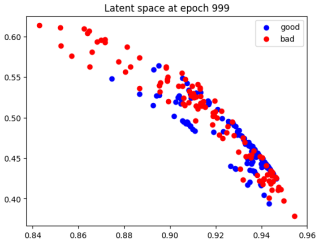
- Method

  I compared two different architecture, which have 2 and 3 nodes in the layer before output layer respectively.

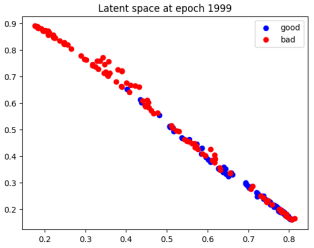
- Two nodes
  - Architecture

layer name	output size	operations
fc1	4	fc [34,4], sigmoid
fc2	4	fc [4,4], sigmoid
fc3	2	fc [4,2], sigmoid
fc4	1	fc [2,1], sigmoid

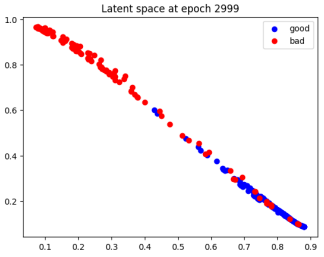
  - Distribution of latent features



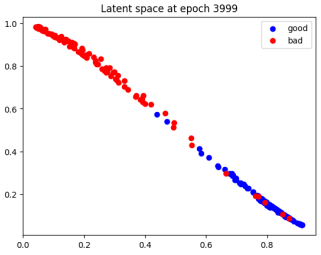
a. epoch 999



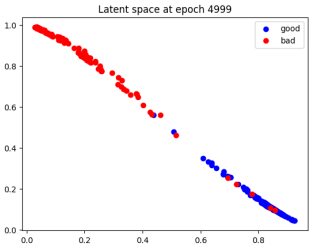
b. epoch 1999



c. epoch 2999



d. epoch 3999

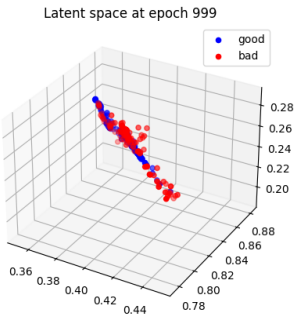


e. epoch 4999

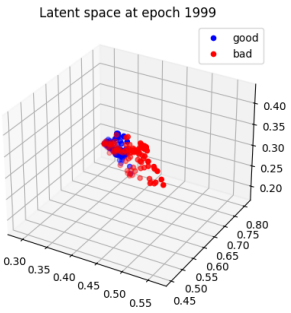
- Three nodes
  - Architecture

layer name	output size	operations
fc1	4	fc [34,4], sigmoid
fc2	4	fc [4,4], sigmoid
fc3	***3***	fc [4,***3***], sigmoid
fc4	1	fc [2,1], sigmoid

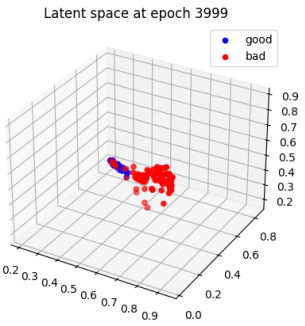
  - Distribution of latent features



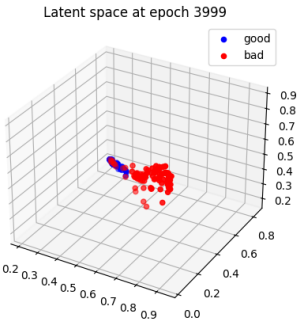
a. epoch 999



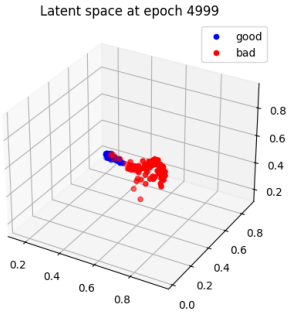
b. epoch 1999



c. epoch 2999



d. epoch 3999



e. epoch 4999