

BUS/CSC 386 Homework #2

Regression Neural Network

Using the housing.csv dataset from california-house-prices (provided), do the following:

1. Import the data using the read.csv function
2. Inspect the data using the glimpse function and document the number of variables and samples
3. Move the median-house-value response variable to the first column
4. Split the data into training and test, features and outcomes, respectively using 80% of the dataset for training and 20% for test
5. Create new vectors for training and test outcomes, and delete the response variables from the predictor datasets
6. Show the dimension of the training and test predictor sets
7. Show the structure of the training and test predictor sets
8. Create one-hot dummy variables for ocean_proximity using recipe
9. Center and scale both training and test predictors using recipes
10. Build a keras sequential neural densely connected network with:
 - a. 16 input neurons; relu activation; initializer_he_normal
 - b. 16 hidden layer neurons; same as first layer
 - c. 1 output layer neuron, sigmoid activation
11. Compile the learning section as follows:
 - a. Optimizer = adam
 - b. Loss = mse
 - c. Metric = c(mae)
12. Fit the network to the training data, run for 10 epochs with a batch size of 128 and a validation split of 20%.
13. Paste the R code into your homework submission
14. Paste the validation accuracy graph into your homework submission
15. Run 5 different permutations of this initial network and keep the best model using the following as a guide:
 - a. Change the number of layers
 - b. Change the number of neurons per layer
 - c. Change the activation function of each layer
 - d. Change the optimizer and/or the learning rate
 - e. Change the batch size
 - f. Change the number of epochs
16. Document every change event with the corresponding validation accuracy and enter into the homework (brief comments and text only)
17. Paste the R code and validation accuracy graph of your best model into your homework
18. Run the best model on the test data ONE TIME ONLY using the evaluation function
19. Enter the best-model test accuracy into your homework
20. Submit via Moodle