Data Analysis Assignment 1
Neural Network for regression

Developed a neural network function for regression based on principles of the algorithm. It has 3 different arguments, which are number of hidden neurons(M), number of epochs(epoch) and learning rate. And it demonstrated how the training and test MSE are affected by those parameters below.

[Table of test and training MSEs]

***Change of learning rate**

Parameters (M,epoch,learning rate)	**Training MSE**	**Test MSE**
(6,100,0.1)	7.56	10.87
(6,100,0.01)	10.59	13.71
(6,100,0.001)	344.9422	359.275

As learning rate goes up, training and test error also increase.

***Change of epoch**

Parameters (M,epoch,learning rate)	**Training MSE**	**Test MSE**
(3,100,0.1)	8.813904	11.30169
(3,1000,0.1)	8.767284	11.46102
(3,10000,0.1)	8.609831	11.87809

Adjustment of value of epoch affects to both MSE but the transitional patterns are not same. Smallest epoch, 100, has a good performance on test MSE. But, biggest epoch, 10000, shows a best training MSE result.

***Change of M(number of hidden neurons)**

Parameters (M,epoch,learning rate)	**Training MSE**	**Test MSE**
(10,100,0.1)	7.639044	11.1461
(8,100,0.1)	7.297144	11.24854
(6,100,0.1)	7.56	10.87
(5,100,0.1)	7.062872	10.86122
(4,100,0.1)	7.835164	11.48288

As epoch increase, training and test MSE move downward. After hit the bottom, it rises up.

[Comparison between linear regression and neural network for regression]

In terms of MSE performance, neural network for regression algorithm demonstrates better performance.

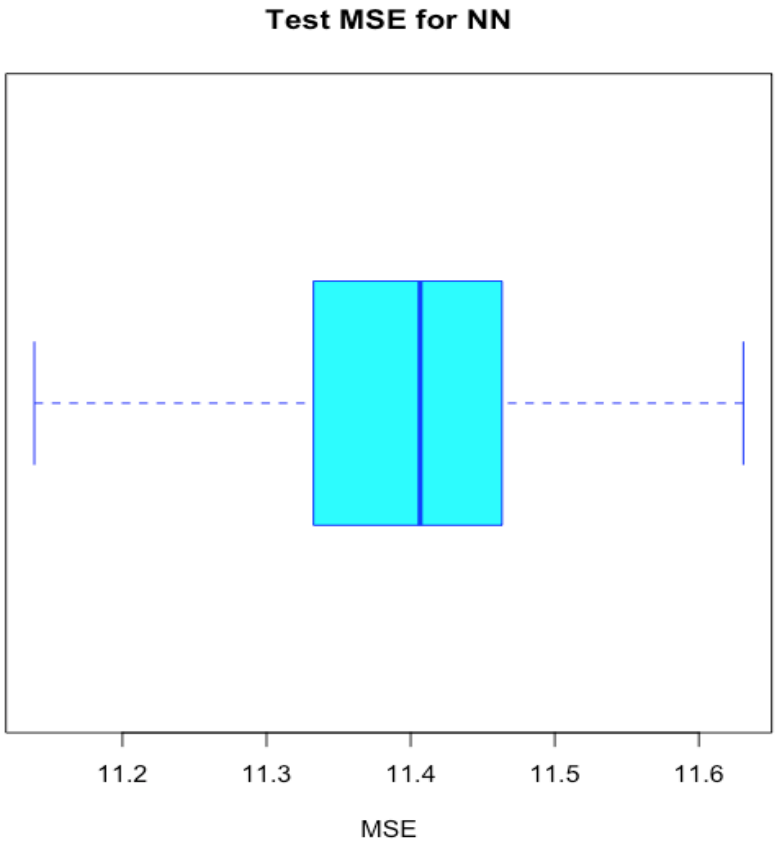
lm.train.MSE: 9.90965 lm.test.MSE: 12.45957

[K-fold cross validation error]

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	15	40.443053	53.17860	38.70156	38.33730
[2,]	16	23.302404	34.51766	27.87830	24.60892
[3,]	17	15.843694	26.61764	23.95193	19.17355
[4,]	18	10.475877	19.91578	21.57788	16.31838
[5,]	19	8.688048	17.76790	22.11022	15.68047
[6,]	20	8.342704	17.94054	23.23625	15.62642
[7,]	21	15.847606	24.22252	36.30807	27.10599
[8,]	22	22.713206	30.44224	46.38057	35.80005
[9,]	23	34.964465	41.59059	61.61611	50.28843
[10,]	24	45.591783	51.77653	74.42898	62.03050

In fold 1(K=1), best M is 20 and MSE is 8.34. The other best M values for each fold can be found on the image above. Average of value of best M in each fold is 19.2 though, the number of neuron can only take a natural number. Therefore, the best M based on this result is 19. Average of cross validation error is (8.34+17.77+21.58+15.63)/4 = 15.83 and the test error of M=20 is 11.27735. It is not the smallest value I could find. In my opinion, it would be possible to take smaller MSE through K-fold cross-validation procedure for every parameter (M, epoch, learning rate).

[Box plot of the 100 cases NN compute]



Interestingly, after computation of 100 times of Neural Network, each MSE converges on a point and the group has very small variance. Following the box plot, the mean of test.MSE is 11.40029 and variance of test.MSE is only 0.007941978, which is very small.

Additional Question.

In this assignment, I applied formula with threshold at first, but performance measure showed the model without threshold (alpha0 and beta0) has better result of MSE. It would have wrong assumption of formulas to weight both threshold values. I have attached the script as other.R file.