Project: Neural Nets

Objectives
To develop a regression model using an A-NN based on the data that you used for the regression project. For this analysis you can use any package that solves A-NNs.

Data Set
The original data for the regression project consisted of only 500 6-tuplets. This data set has now been extended to 5000 6-tuplets: $(Y, X_1, X_2, X_3, X_4$ and $X_5)$ where Y is the dependent variable (label), and X_1, X_2, X_3, X_4 and X_5 are the independent variables.

Main Task
1. *Train a neural net with 6 input units and a single output unit.*
   Assume a number of hidden layers and units per layer and then train your neural net model using the cross validation method. That is,
   - Split the data into 10 subsets of 500 observations each. Then combine subsets 1 to 9 to create the training data set, and use subset 10 for testing. Determine the estimated $\hat{y}$ values. Calculate the error, i.e., $y - \hat{y}$, and then compute the sum of squared errors (SSE), the root mean squared error, and R^2.
   - Repeat the same process by combining subsets 1 to 8 and 9 to create the training data set 2, and use subset 9 for testing.
   - Continue in this manner until you have done all 10 combinations of the folds.
2. *Search for a good model*
   Vary the number of hidden layers and units per layer and repeat the above step in order to obtain the best model as judged by the metrics *SSE*, the root mean squared error, and R^2.

Extra Credit
Re-run your multiple regression model that you determined in the regression project and obtain all the residuals. Calculate the *SSE*, the root mean squared error, and R^2. Compare the neural net model with the multiple regression model from the point of accuracy as measured by *SSE*, the root mean squared error, and R^2.

What to submit
Submit a single report for the main task and for the extra task along with the code that you wrote to obtain the results. Indicate clearly each part of your report. Provide enough results to support your conclusions. Conclusions without insufficient results will make you lose grades. Also, it is important that you develop your own code. Sharing code is not allowed and constitutes cheating, in which case both students (the one that aids and the one that receives) will get a zero for the project and will be reported to the student conduct office.