Project 4: Neural Networks Project

The objective of this project is to determine the configuration of a feedforward neural network and its hyper-parameters using a data set with five input variables and a single continuous output variable. You can use any statistical package, such as MatLab, R, SAS, or Python with all the available statistical functions, or functions from different packages.

Data set

Use the data set with your login name posted in moodle. The data set was generated using your student id number, and it is the same as the one you used in the regression project, only it has been expanded to 2300 points. In each csv file, the first five columns give the data for the five input variables X_1, X_2, X_3, X_4 and X_5. The output variable Y is in the last column. (Note: If your data set does not appear to be good, let us know and we will provide you with another set.)

Task 1. Automatic grid search

1.1. Use the first 2000 points to train your neural network using an automatic grid search tool. Use the remaining 300 points as test data, if needed.
1.2. Discuss your results.

Task 2. Compare the trained neural network with multivariable regression

2.1. Re-run your multiple regression model that you obtained in the regression project using the first 2000 data points, and calculate the SSE. Do not do any residual analysis, as you did in the regression task.
2.2. Compare the SSE with the neural network determined above.
2.3. Comment on which model is the best.

What to submit

1. For each task submit the following: The code you used for the task. It does not have to run on eos, and it may be a number of different pieces of code from different packages.

   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.

2. Your results (graphs, tables, etc) and your conclusions.

How to submit your work

Submit your report and code in a **single.zip file** (not .tar or .7z, or any other format). Your report should be a **single pdf file** that contains all your graphs, tables and conclusions.

Grading

The TA will first verify that your code works and produces the results you submit. The break down of the grades will be as follows:

Task 1: 60 points
Task 2: 40 points

You will receive a bad grade if you submit results without substantive conclusions, or conclusions that are not backed by sufficient results.