Final Exam - Fall 2020

This is the first Question I of the final exam and has weight 20 pts.
Due to 4:00 PM, January 30^{th}, Saturday, 2021

You will be given two data sets that includes the x and y values. Your task is to fit a function to the given data in the file *"IE440Final20Training.txt"*.

1. **Least Square Method:**

   (a) Use a Linear Regression Model: $y = w_1 x + w_0$.

   (b) Use a Polynomial Regression Model: $y = w_3 x^3 + w_2 x^2 + w_1 x + w_0$.

   For these two methods, your task is to determine least square estimators of (a) and (b) by solving the following unconstrained optimization problems

   $$\min_{w_0, w_1} \sum_{p=1}^{P_{tra}} (y_p - w_0 - w_1 x_p)^2 \quad \text{and} \quad \min_{w_0, w_1, w_2, w_3} \sum_{p=1}^{P_{tra}} \left(y_p - w_0 - w_1 x_p - w_2 x_p^2 - w_3 x_p^3\right)^2.$$

   Here (x_p, y_p) is the p^{th} data pair, and P_{tra} and P_{test} are the size of the training and test sets.

   You can either solve them analytically or using one of the algorithms you have already programmed. The use of one of the optimization or statistical inference package's optimization regression modules is forbidden.

2. **Nonlinear Regression:**

   (a) Design a multi-layer perceptron with one output unit, two input terminals and one hidden layer using the training data given in file *"IE440Final20Training.txt"* and determine the number of hidden units using the test data given in file *"IE440Final20Test.txt"*. The perceptron is illustrated in Figure 1.

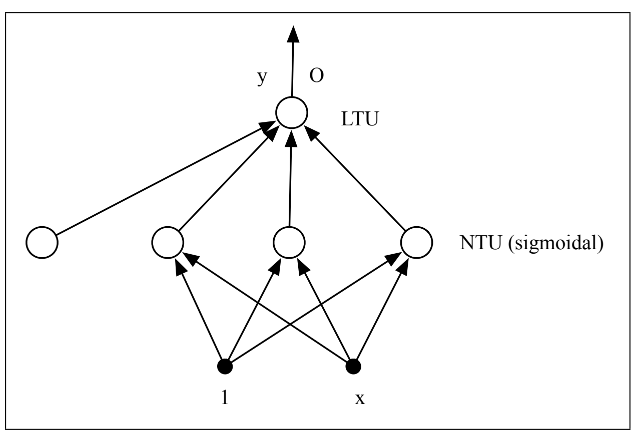

Figure 1: Sample perceptron for 2.a.

(b) Design a multi-layer perceptron with one output unit, four input terminals (one for x, one for x^2, one for x^3 and one for $x_0 = 1$), and one hidden layer using the training data given in file *"IE440Final20Training.txt"* and determine the number of hidden units using the test data given in file *"IE440Final20Test.txt"*. The perceptron is illustrated for three hidden units in Figure 2.

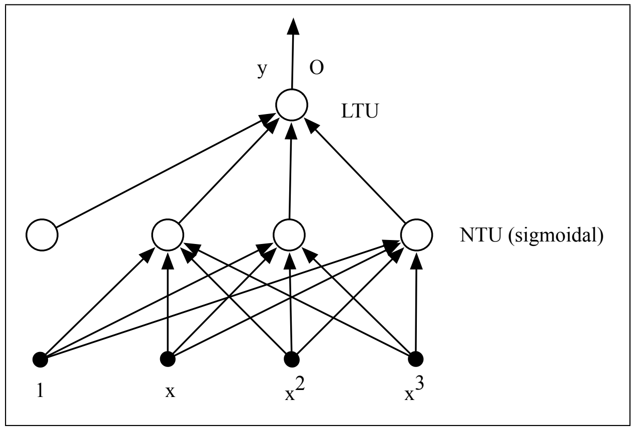

Figure 2: Sample perceptron for 2.b.

For both cases set $\alpha^{(0)} = 0.5, \eta = 0.9, \epsilon = 0.001, J^{(0)} = 3$ (i.e. initial number of hidden units), and assume that the output neuron is a linear threshold unit (i.e. $g(o) = o$) and each hidden neuron is a nonlinear threshold unit with sigmoidal activation factor $g(h) = \frac{1}{1+e^{-h}}$.

Report for all cases:

Method	Training SSE	Test MSE	s^2 for Test MSE
1.(a)	$\sum_{p=1}^{N_{tra}} e_p^2$	$\frac{1}{P_{test}} \sum_{p=1}^{P_{test}} e_p^2$	$\frac{1}{P_{test}-1} \sum_{p=1}^{P_{test}} (TestMSE - e_p^2)^2$
1.(b)	...	...	...
2.(a)	...	...	...
2.(b)	...	...	...

where $e_p^2 = (y_p - O_p)^2$ is the error term of the related data (O_p is the estimated value).

- Plot the training data and the regression functions.
- Plot the test data and the regression functions.
- Comment on the results for training and test errors.

You are repsonsible to write your own backpropagation programs. The use of one of the neural network software is forbidden.

For all parts of this question include the screen shots of your outputs and your source codes in your reports. Submit the soft copy of your **individual** report and the source code which are named as **FINAL20-SchoolID** to moodle before 4:00 PM, January 30^{th}, 2021.