



4005 755 70 - Neural Networks and Machine Learning

Homework 1

Date due: April 26

This is a take home assignment. You may do it in groups of two or three students.

You may consult your books, your notes, papers distributed in class, the Web, the library and any other publicly available written sources. When you use a source other than the book, papers distributed in class, or your class notes, please give a reference to the source. For a Web site, give the URL.

Electronic submission by email to lr@cs.rit.edu as a MS-Word or PDF file is encouraged. Submit your report by 11.59 pm, April 26th.

Function predictor.

As you know neural networks structures have been proven to be universal approximators. This feature is used in design signal predictors based on neural networks. Of course, some neural networks approximate certain functions better than others. In this project you will have to study approximation and prediction abilities of neural networks by conducting experiments with them. Experiments will be done by computer simulation.

What needs to be done:

You have to build up a neural network model, which given values of the function $f(t)$, $f(t-1)$, $f(t-2)$, ..., $f(t-n)$ should be able to predict the value $f(t+1)$. This kind of predictor network could be applied for detecting novelty in data streams, as an erroneous prediction would mean that either the network was wrong in its prediction of change in the real signal has occurred.

More formally novelty detection would depend if the $\text{DIFF}(t) = \text{abs}(f(t+1)_{\text{predicted}} - f(t+1)_{\text{real}})$ is greater than the threshold value R . The increase of the DIFF over the given threshold value will mean that the novelty has been detected.

To facilitate function habituation, the network is constantly trained with the newest set of inputs.

The quality of your design can be evaluated with two indicators: a false positive ratio, when you system recognizes as a novel situation, when actually there is no novelty and a false negative ratio, when you do not recognize the novel event. Both things are bad and a good detector should minimize both ratios.

In this project you have to:

- 1) assume the novelty detection in the sine wave signal of the form $Y(t) = a \sin(bt + c)$.

- 2) modify the sine wave by adding some noise $Z(t) = Y(t) + N(t)$. You may simulate noise as a gaussian random value with zero mean value and the deviation D , which you might like to take as 1-2% of the $Y(t)$
- 3) you may use $z(t)$ as an input for your neural network
- 4) you may choose any initial values of a , b , c but you have to change each of them at least three times during your testing period. Any change will constitute the novelty you have to detect.
- 5) choose the type and the structure of the neural network. In a case you do this assignment in a group of two you can choose any two neural network types, if there are three of you in a group you have to try at least three different types. As a bonus question you may try some unsupervised neural network.
- 6) Choose the value of the threshold. In this project you have to investigate how false positive and false negative ratios depend on the threshold values. To investigate it, you need to try at least a few threshold values.
- 7) Run experiments on novelty detection. Introduce some novelty (a few changes in a signal value) during the experiment. Calculate false positive and negative ratios.
- 8) Repeat experiments with other threshold values. Try to investigate how the optimal threshold value depends on the noise characteristics.
- 9) Process your experiment results: draw a few graphs representing your results
- 10) Write down conclusions:
 - a) how the false positives and negatives depend on the threshold values
 - b) how they depend on the number of inputs and the signal history
 - c) which neural network structure worked better
 - d) anything else you wish to add
- 11) Write down your report and a Word or PDF file.
Include into your report:
 - a) description of your experiments, what you did and how
 - b) brief description (user's guide) of your code you used
 - c) your code (could be in any language or package)
 - d) your experimental results processed and presented as graphs, diagrams, etc
 - e) your conclusions

In design your experiments you may use any design package such as Matlab, neuroSolutions, Joone or any other one or write your own code in any language.