Using Normalised Radial Based Functions (NRBF's) to Prodict Energy Consumption in the National Grid

Connor Wheeler

I. Introduction

Training a Neural Network to prodict the energy Consumptions of the national was not the easiest of tasks for the network to proform. There a number of intrsting occurances in the data, the output of the network and the results of the sigma optimisation and node optimisation.

II. Network

A. NRBF

Normalised Radial Based Functions (NRBF) work by using the activation of all nodes in the hidden layer to work out the output of the network. This is done by using the gaussian activation function of the nodes in the hidden layer, to work out how active the node is when a value is passed to it. if a node is very then its activation value will be one or very close to one, where as an inactive node will be much closer to zero. The activation of all of the nodes is later used to work out what the output of the net work will be.

When the activation has been calculated this can then be used to to get the output of the network as the more active nodes will contribute more to the final value that is output based on these inputs. To do this the the sum of the nodes weights multipied by the activation of the node is calculated. This part of the equaction can be seen in figure 1:

$$\sum_{n=1}^{N} W_n \phi(\|x - x_n\|)$$

figure 1 : sum of all node activations multipied by weights of all nodes

After this the total sum of all node activations is calculated and summed. The equation for this can be seen figure 2.

$$\sum_{n=1}^{N} \phi(\|x - x_n\|)$$

figure 2 : sum of all node activations

When these have been calculated the 2 values are devided. to get the final output from the hidden layer. The whole equation can been seen in figure 3.

$$f(x) = \frac{\sum_{n=1}^{N} W_n \phi(\|x - x_n\|)}{\sum_{n=1}^{N} \phi(\|x - x_n\|)}$$

figure 3 : sum of all node activations multipied by weights of all nodes devided by sum of all node activations

Node Activation Equation

The node activation equation is used to calculated the activation of the node. if the value before the exponential is calculated is 0 then the activation of the node will be 1.

$$y = exp(-\frac{1}{2\sigma^2}\sum_{k=1}^{K}(x_k - w_{jk})^2)$$

Root Mean Squar Equation

The Root Mean Square equation is used to calculated how incorrect the network was with its outout. This was used to comapar diffrent sigmas to see which has profomed the best on the testing data set.

$$RMS = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(y_i^p - y_{id}^p)^2}$$

Weight Update Equation

The weight update equation is used to adjust the weights in the hidden layer. This will allow the network to become more accurate over time as the weights get adjusted more and more, as the network becomes more accurate these adjustments become smaller. To do this the old weight is added to using the learning rate (α) multipied by the target value - the networks output, multipied by the activation of the node (ϕ).

$$W \leftarrow W + \alpha * (target - Networkoutput) * \phi$$

1) Task 1:

For task one an NRBF was made to work on a small and evanly distributed data set, to get the understanding of the network and the maths correct. To make sure the network was working correctly the sigma was set to 0.01 to see the step of the network function. this can be seen in figure 7. This was useful as it allowed to check if the network was working and was covering all of the traing data points with the network function.

Network function with a sigma of 0.01

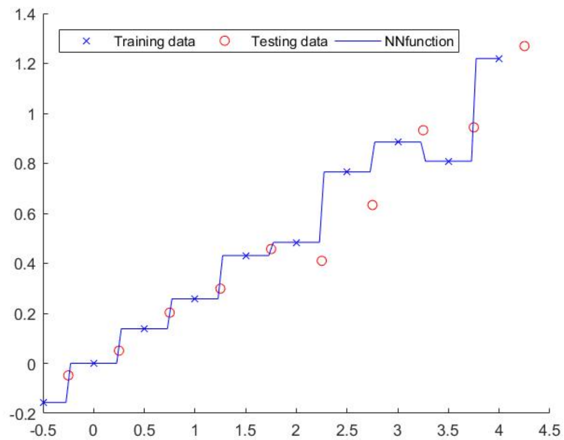

figure 7 : Network function when sigma is 0.01

Once the network was working properly, the sigma optimiation could begin to see which sigma would be best to use for the network. To do this the network was run over the data set 100 times and then the final test and train error where taken and stored. This allowed for the best value on the testing data to be found. The sigma was tested between 0.1 - 1 and the table of results can be seen in figure 10. The best sigma value that could be found from this testing was 0.9 as it had the lowest test error of all of the value tried with a error value of 0.0915. The error graph and network function graph for this sigma can be seen in figure 8 and 9.

Error plot for a sigma of 0.9

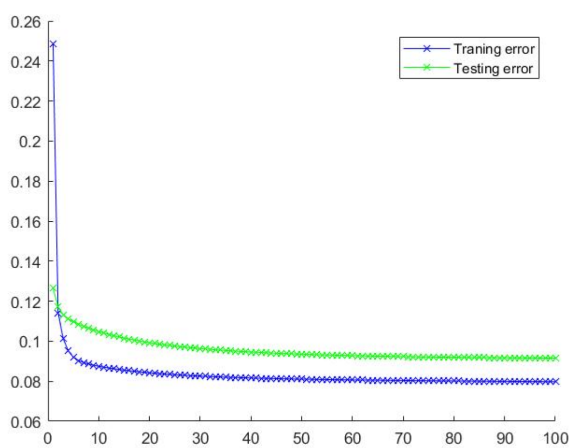

figure 8 : error plot for the network with sigma 0.9

Network function plot for a sigma of 0.9

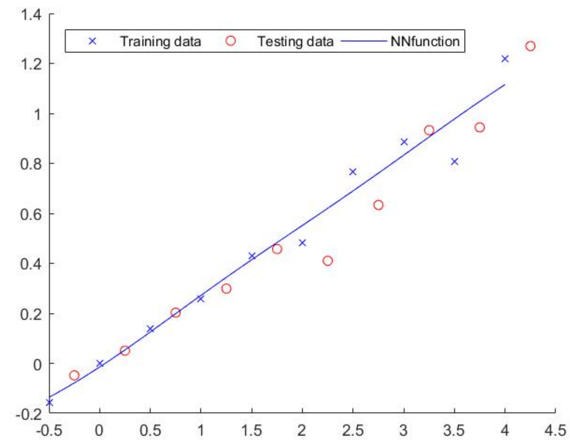

figure 9 : Network function when sigma is 0.9

Sigma Value	Train Error	Test Error
0.1	0.2733e-16	0.1005
0.2	2.0126e-10	0.1016
0.3	1.3562e-4	0.1094
0.4	0.0131	0.1161
0.5	0.0439	0.1073
0.6	0.0632	0.0983
0.7	0.0730	0.0933
0.8	0.0781	0.0919
0.9	0.0797	0.0915
1.0	0.0800	0.0918

figure 10 : Sigma optimisation table

2) *Tast 2:*

Number of Nodes	Train Error	Test Error

Sigma Value	Train Error	Test Error
0.1		
0.2		
0.3		
0.4		
0.5		
0.6		
0.7		
0.8		
0.9		
1.0		

B. MLP

1) *Task 2:*

III. Data

A. Data processing methods

B. Problems with the data

IV. results

V. conclution