Neural Networks Projects

Undergraduate Radial Basis Neural Network

Tom Makkink

2018

**1 Introduction**

Universities often give entrance tests to incoming students, in order to assess their existing knowledge and to predict how well they may potentially do at the University. We would like to construct a neural network that explored the relationship between the Swedish points of students, the rating of the school they previously attended, and results they achieved on the entrance test in relation to their academic performance in semester 1 and semester 2 of their first year at the University.

**2 The Data**

The data set (ugraddata.txt) contains 500 samples of the form:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Swedish Points | School rating | Test | Semester 1 | Semester 2 |

Some notes on the data:

* The Swedish points are calculated according to a point system, with a maximum of 46 and a minimum of 20.
* The school rating is a value from 1 – 10, with 10 being the highest and 1 being the lowest.
* The test, semester 1 and semester 2 results are measured as a percentage.

The data was extracted and arranged into patterns and targets as follows:

%----------------------------------------------------------------------

% Arrange the data

%----------------------------------------------------------------------

Data = importdata('ugraddata5.txt');

P = Data(:, [1 2 3]);

T = Data(:, [4 5]);

% Arrange as rows

p = P';

t = T';

The input patterns consisted of the Swedish points, the school rating and the entrance test school, and the targets as the student’s academic performance in semester 1 and semester 2.

**3 Method**

The data was first normalized in the range [-1, 1], as this is the region the transfer functions have the greatest effect:

% Scale down the inputs by row

pMax = max(p,[],2);

pMin = min(p,[],2);

pf = 2./(pMax - pMin); % Factors to scale down rows

pc = -(pMax+pMin)./(pMax-pMin); % Additive terms

Dp = diag(pf); % pf down the diagonal of Dp

pn = Dp\*p+repmat(pc,1,size(p,2)); % Scale down

% Scale down the targets

tMax = max(t,[],2);

tMin = min(t,[],2);

tf = 2./(tMax-tMin);

tc = -(tMax + tMin)./(tMax-tMin);

Dt = diag(tf);

tn = Dt\*t + repmat(tc,1,size(t,2));

The data was further manually divided into training and testing sets for both a scaled and unscaled version:

%---------------------------------------------------------------------

% Divide training and test sets by index

%----------------------------------------------------------------------

% Train Index

I1 = randperm(floor(2\*q/3)); % Train 2/3rds of the data

q1 = length(I1);

% Test index

I2 = setdiff([1:q],I1);

q2 = length(I2);

% Training set:

p1 = p(:, I1);

t1 = t(:,I1);

ptrain = pn(:, I1); % Scaled versions

ttrain = tn(:,I1);

% Test set

p2 = p(:,I2);

t2 = t(:,I2);

ptest = pn(:, I2); % Scaled versions

ttest = tn(:, I2);

After a experimenting with a variety of different spread ranges, and line spaces, it was eventually decided to use a spread range of between .6 and 6, with 1000 points in order to find the best spread:

% sr = input('spread range = [min, max] = ')

sr = [.1,6];

%----------------------------------------------------------------------

% Find the best spread

%----------------------------------------------------------------------

% matrix for storing spread and r

R1 = [];

R2 = [];

for s=linspace(sr(1),sr(2),1000)

%train on training set

net = newrbe(ptrain,ttrain,s);

% simulate on test set

a = sim(net,ptest);

[r2,r] = correlation(a,ttest);

R2=[R2;[s r2(1,:) r2(2,:)]];

R1=[R1;[s r(1,:) r(2,:)]];

end

Using a 1000 points as opposed to say 100, produced slightly better results. Although the larger iteration count did impact performance and slow down the network, the better results were ultimately worth it.

The spreads and corresponding correlation coefficient and R-squared statistic for that spread were stored in the matrixes R1 and R2 and displayed as follows:

disp(' Feature 1 Feature 2 \n');

disp(' spread r2 r2');

fprintf('%8.4f\t%-8.4f\t%-8.4f\n',R2')

disp(' Feature 1 Feature 2 \n')

disp(' spread r r');

fprintf('%8.4f\t%-8.4f\t%-8.4f\n',R1')

The best spread across both features was calculated as follows:

%find the best spread wrt r2 stat

[mr2,i]= max(R2(:,2)+R2(:,3));

bs2 = R2(i,1);

%find the best spread wrt r: correlation coeff

[mrl,j] = max(R1(:,2) + R1(:,3));

bs = R1(j,1);

Ultimately, the spread with the best R-squared statistic was chosen, as the R-squared statistic is a more accurate measure of performance in the context, and used in the construction of a radial basis neural network as follows:

%----------------------------------------------------------------------

% Simulate with the best spread

%----------------------------------------------------------------------

net = newrbe (ptrain,ttrain,bs2);

The net was then simulated on the scaled training and testing data, and its performance accessed with using the R-squared statistic and correlation coefficient.

atrain = sim(net,ptrain);

atest = sim(net,ptest);

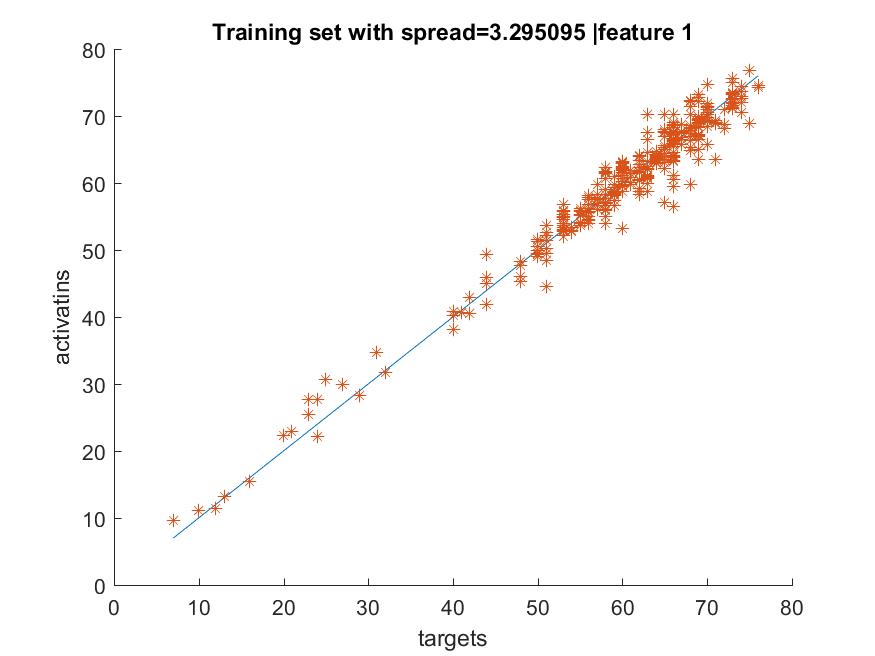
disp('Training set: ')

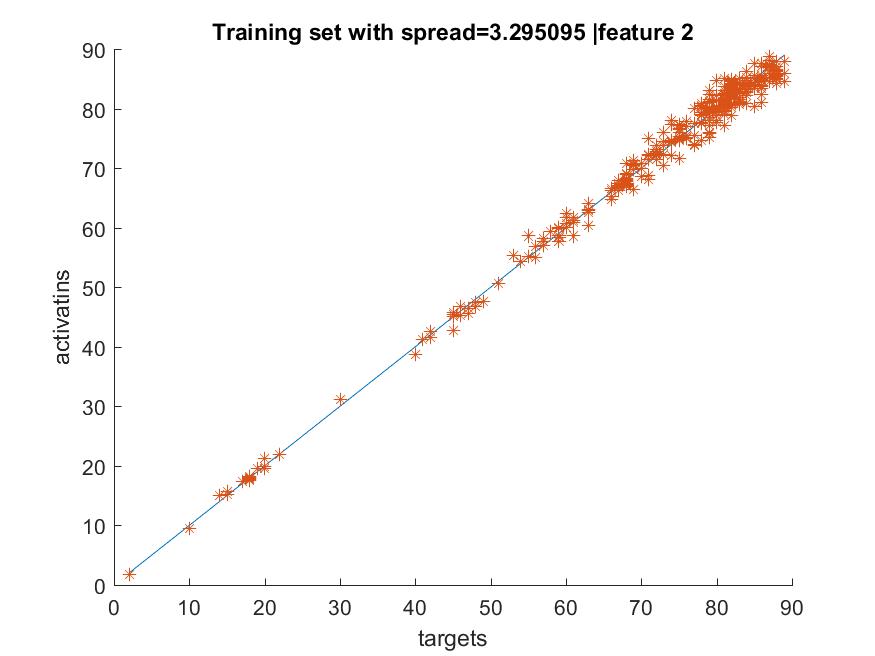
[r2train rtrain] = correlation(atrain,ttrain)

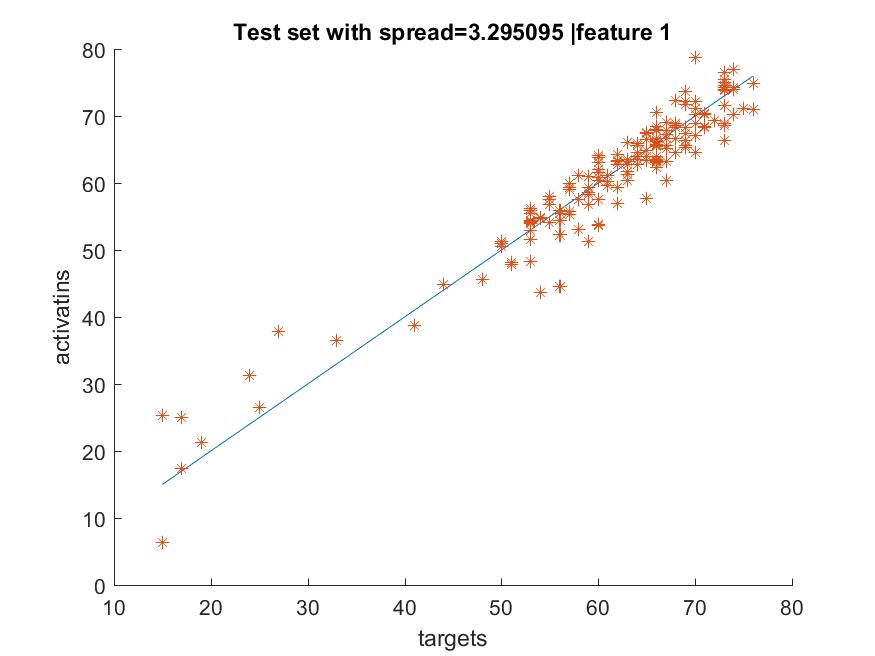
disp('Test set: ')

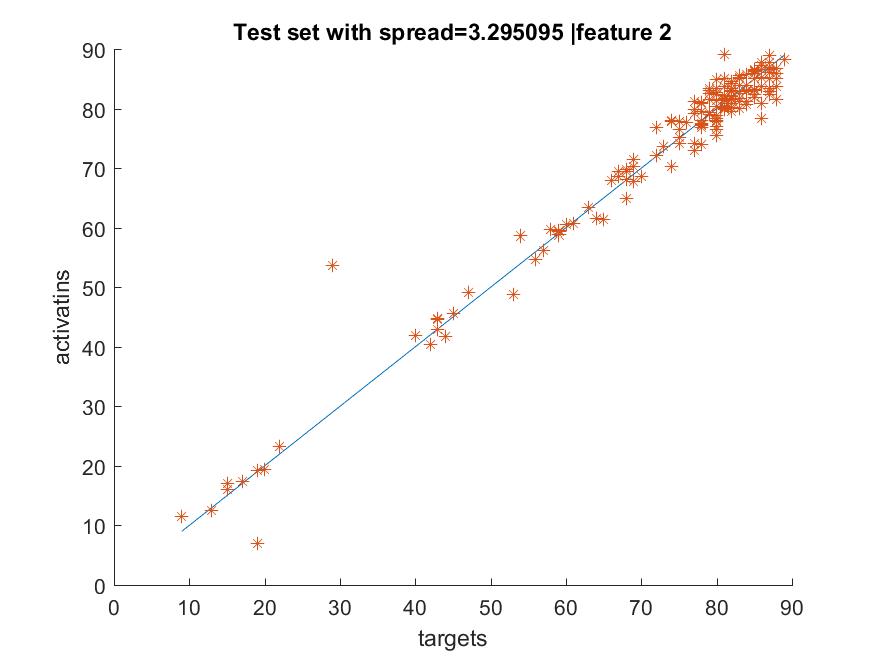
[r2test rtest] = correlation(atest,ttest)

The results were then scaled up, and plotted to produce the following results:









The final R-squared statistic and correlation coefficients were as follows:

Training set:

|  |  |  |
| --- | --- | --- |
|  | R-Squared Statistic | Correlation Coefficient |
| Feature 1 | 0.9609 | 1.0000 |
| Feature 2 | 0.9609 | 1.0000 |

Test set:

|  |  |  |
| --- | --- | --- |
|  | R-Squared Statistic | Correlation Coefficient |
| Feature 1 | 0.9223 | 0.9999 |
| Feature 2 | 0.9668 | 0.9999 |

**4 Conclusion**

The neural network performed well, and was able to provide a good prediction as to how well a student would perform in semester 1 and 2 at University, based on their Swedish points, school rating and entrance test score.

There were of course a few anomalies, as there are exceptions to any rule, and in light of this it should be noted that a neural network of this nature should be used as an indicator of how well a student will do, rather than an exact prediction.

**5 References**

*https://www.mathworks.com/help/newrbe.html*