R^2 (R Squared)

- *The coefficient of multiple determination* (also called *multiple correlation*)

R Squared is a statistical indicator usually applied to multiple regression analysis. It compares the accuracy of the model to the accuracy of a trivial benchmark model wherein the prediction is just the mean of all of the samples. That is, R^2 reflects the number of errors made when using the regression model to guess the value of the dependent (*y*), in ratio to the total errors made when using only the dependent's mean as the basis for estimating all cases.

Mathematically,

$$R^2 = 1 - \frac{SSE}{SST}$$

where SSE = error sum of squares = $\sum (y - \hat{y})^2$

SST = total sum of squares = $\sum (y - \bar{y})^2$

y is the actual value, $\hat{y}$ is the predicated value of *y*, and $\bar{y}$ is the mean of the *y* values.

Thus *R*-square is 1 minus regression error as a percent of total error and will be 0 when regression error is as large as it would be if you simply guessed the mean for all cases of *y*.

A perfect fit would result in an *R* squared value of 1, a very good fit near 1, and a very poor fit near 0. If your neural model predictions are worse than you could predict by just using the mean of your sample case outputs, the R squared value will be 0.

Do not confuse *R* squared, the coefficient of multiple determinations, with *r squared*, the coefficient of determination (the square of the correlation coefficient) for the simple linear regression case. The latter is usually the one that is found in spreadsheets. See any statistics book for more details.

Also note that sometimes the coefficient of multiple determination is called the multiple coefficient of determination, but in any case it refers to a multiple regression fit as opposed to a simple regression fit. Also, do not confuse it with ***r***, the *correlation coefficient*.

Note: R squared is not the ultimate measure of whether or not your network is producing good results, especially for classification networks. You might decide the network is OK by the number of correct classifications. For example, if you have a classification network with two outputs that generate output values of .6 and .4, the R squared value will not be very high.

Mean Squared Error (MSE)

A statistical measure of the differences between the values of the outputs in the training set and the output values the network is predicting. This is the mean over all patterns in the file of the square of the actual value minus the predicted value, i.e., the mean of (actual - predicted)^2. The errors are squared to penalize the larger errors and to cancel the effect of the positive and negative values of the differences.

RMS Error (root mean square error) - the square root of the MSE