

NeuroSolutions for Excel

Table of contents

Welcome to NeuroSolutions for Excel	5
Getting Started	5
Launching NeuroSolutions for Excel	5
Getting Started in Excel with NeuroSolutions	7
NeuroSolutions Modules in Excel	13
Express Builder Module	13
Get Started	14
Apply Solution	16
Manage Report Sheets	17
Data Tools Module	18
Preprocess Data	19
Difference	19
Randomize Rows	20
Sample	20
Shift	21
Moving Average	21
Encode Two Class Column	22
Translate Symbolic Columns	22
Insert Column Labels	24
Clean Data	24
Analyze	25
Correlation	26
Time Series Plot	26
XY Scatter Plot	26
Histogram	26
Summary Statistics	27
Trend Accuracy	28
Tag Data	28
Column(s) As Input	30
Column(s) As Desired	30
Column(s) As Symbol	30
Row(s) As Training	30
Row(s) As Cross Validation	31
Row(s) As Testing	31
Row(s) As Production	31
All Columns As Input	32
All Non-Numeric Columns As Symbol	32
All Rows As Training	32
Rows By Percentages	32
Clear Tags	33
Clear Column Tag	33
Clear Symbol Tag	33
Clear Row Tag	34
Clear All Tags	34
Select Cross-Section	34
Refresh Tag Formats	34
Neural Network Configuration	35

Classification MLP	35
Classification PNN	35
Classification TDNN	35
Regression MLP	36
Regression GRNN	36
Regression TDNN	36
New Custom Network	36
Open Existing Network	37
Train Network Module	38
Train	39
Train N Times	41
Vary a Parameter	43
Train Genetic	45
Advanced Train Network Module	49
Leave N Out	49
Test Network Module	52
Test	52
Test Reports	53
Sensitivity About the Mean	56
Apply Production Dataset	58
Automation	58
New Batch	58
Batch Manager	60
Run Batch	61
Create Network...	61
Train Network...	62
Test Network...	62
Create Data Files...	63
Miscellaneous	63
Data Sheets	63
Goto Active Data Sheet	63
Data Sheets	64
Reports	64
Goto Active Report	64
Reports	64
Data Files	65
All Files	66
Training Files	66
Cross Validation Files	66
Testing Files	67
Production Input File	67
View Data File	67
Delete Data Files	68
FAQ & Troubleshooting	68
General Technology	68
What is data preprocessing?	69
What is GPU (Graphical Processing Unit) processing?	69
What is Leave-N-Out?	69
What is the difference between NVIDIA CUDA and OpenCL?	69
What is the difference between Single vs Double Precision?	69

NeuroSolutions for Excel	69
Does my data need to be formatted any particular way?	70
What are the restrictions of the free trial?	71
What are the differences between NeuroSolutions & NeuroSolutions Pro?	72
Sales & License Information	73
Contacting NeuroDimension	73
Do you provide educational discounts?	74
Do you provide site licenses?	74
How to buy NeuroSolutions?	74
Related Products & Modules	75
What is NeuroSolutions Infinity?	76
Additional Information	76
System Requirements	76

Welcome to NeuroSolutions for Excel



NeuroSolutions is an easy-to-use neural network software package for Windows. It combines a modular, icon-based network design interface with an implementation of advanced artificial intelligence and learning algorithms using intuitive wizards or an easy-to-use Excel™ interface. Perform sales forecasting, sports predictions, medical classification, and much more with NeuroSolutions.

Getting Started

Follow the [Getting Started](#) guides to familiarize yourself with the processes of creating neural network models in NeuroSolutions.

How NeuroSolutions Works

Neural networks are long, complicated mathematical equations and NeuroSolutions is designed to make the technology easy and accessible to both novice and advanced neural network developers. There are three basic phases in neural network analysis: training the network on your data, testing the network for accuracy and making predictions/classifying from new data. Only the Express Builder in the NeuroSolutions Excel interface can accomplish all of this automatically in one simple step!

Easy to Use Excel Interface

With the NeuroSolutions for Excel interface, it has never been easier to get started quickly in solving your problem. It provides an easy-to-use and intuitive interface for users to easily setup a simulation that automatically builds, trains and tests multiple neural network topologies and generates an easy-to-read report of the results including the best-performing model.

Created with the Standard Edition of HelpNDoc: [Easy EBook and documentation generator](#)

Getting Started

Table of Contents

- [Launching NeuroSolutions for Excel](#)
- [Getting Started in Excel with NeuroSolutions](#)

Created with the Standard Edition of HelpNDoc: [Full-featured Help generator](#)

Launching NeuroSolutions for Excel

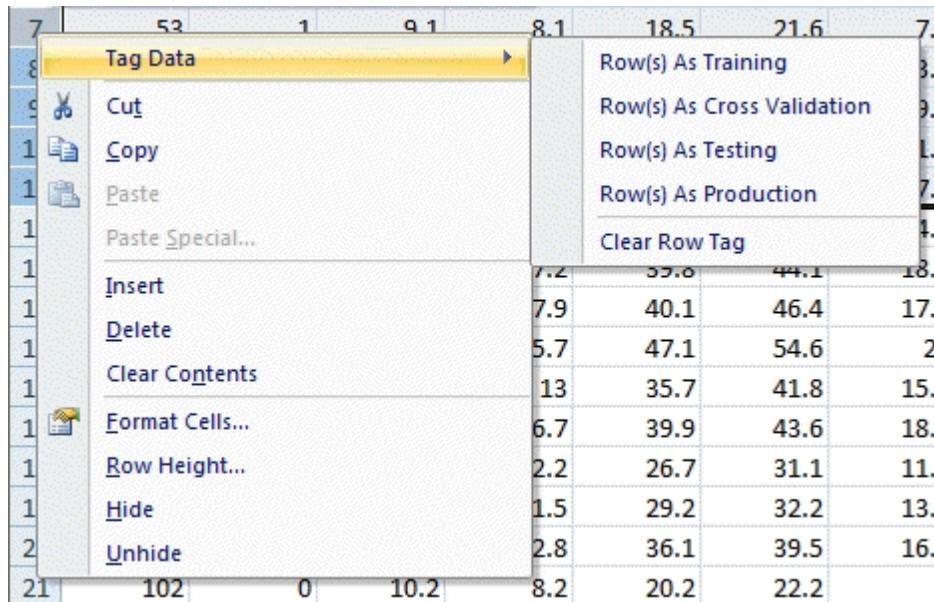
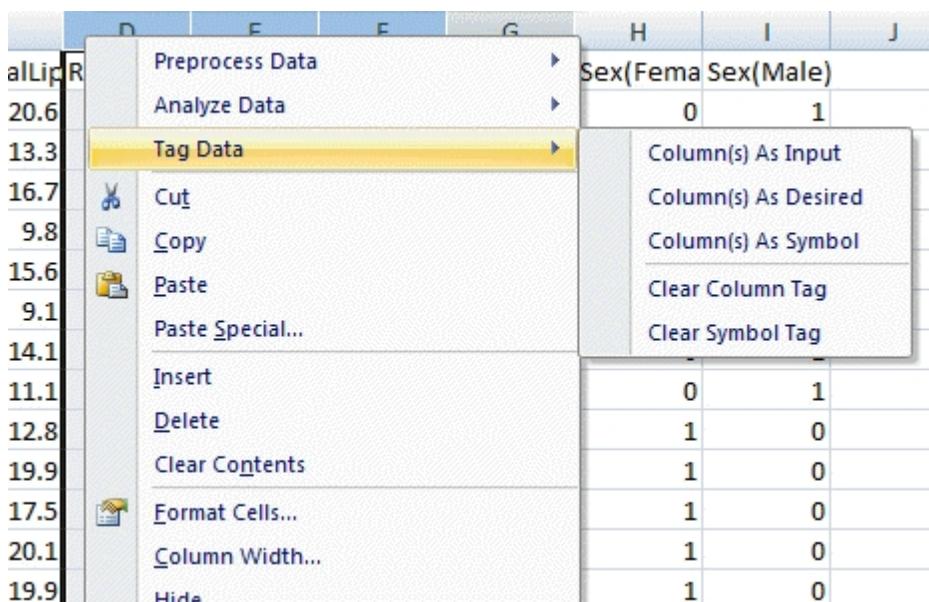
There are three primary ways to run NeuroSolutions for Excel:

1. Launch Microsoft Excel (after installation)
2. Launch NeuroSolutions and from the launcher select "Launch Excel with NeuroSolutions"
3. Click the "NS Excel" button located on the toolbar within NeuroSolutions.

*NeuroSolutions toolbar loaded in Excel 2007 and later*

All of these methods will launch Microsoft Excel and load the NeuroSolutions for Excel add-in. A “NeuroSolutions” menu will appear in the menu bar of Microsoft Excel. This is the menu you will use to perform most of the NeuroSolutions for Excel operations.

You can also perform column and row operations by right-clicking on one or more columns or rows. The appropriate NeuroSolutions for Excel menu items will appear on the resulting shortcut menu. Shortcut menus provide a faster way to perform an operation by displaying operations available in the current context (selection).

*Tagging rows as Training using row shortcut menu**Tagging columns as Desired using column shortcut menu*

Getting Started in Excel with NeuroSolutions

The Excel interface is accessible by clicking on the "Launch Excel with NeuroSolutions" button from the launcher OR by launching Excel directly. This quick tutorial will provide you with high level details of tagging your data, building a neural network, training it and finally testing the performance.

Importing Data in Excel

Excel provides the perfect interface for presenting data for neural networks. It can work with any column-formatted data such as comma-separated files (.csv), tab-delimited files (.txt) and of course Excel files (.xls and .xlsx).

If you do not have data available for this tutorial simply use the CrabData.csv located in the "C:\Users\[USER]\Documents\NeuroDimension\NeuroSolutions\SampleData" directory. Please note that you will need to File > Save As the data to a different writable directory such as My Documents so that the software has write access to store underlying files during processing.

Designating Inputs & Desired Outputs

	A	B	C	D	E	F	G	H
1	Specimen	Species	FrontalLip	RearWidth	Length	Width	Depth	Sex
2	142	0	20.6	14.4	42.8	46.5	19.6	Male
3	19	1	13.3	11.1	27.8	32.3	11.3	Male
4	169	0	16.7	14.3	32.3	37	14.7	Female
5	56	1	9.8	8.9	20.4	23.9	8.8	Female
6	164	0	15.6	14.1	31	34.5	13.8	Female
7	53	1	9.1	8.1	18.5	21.6	7.7	Female
8	113	0	14.1	10.5	29.1	31.6	13.1	Male
9	7	1	11.1	9.9	23.8	27.1	9.8	Male
10	73	1	12.8	12.2	27.9	31.9	11.5	Female
11	187	0	19.9	16.6	39.4	43.9	17.9	Female

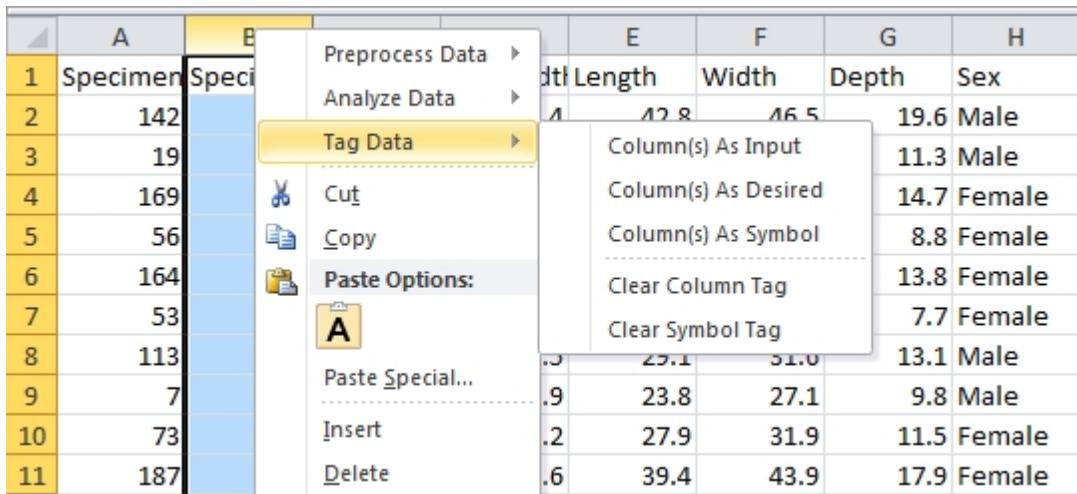


Figure 1. Sample of the CrabData.csv in Excel

Figure 2. Context Menu in Excel

Tagging data is accessible in two ways:

1. By right-clicking on the selected column(s) (see Figure 2).
2. Or, by choosing the Tag Data dropdown menu from the Excel toolbar (see Figure 3).

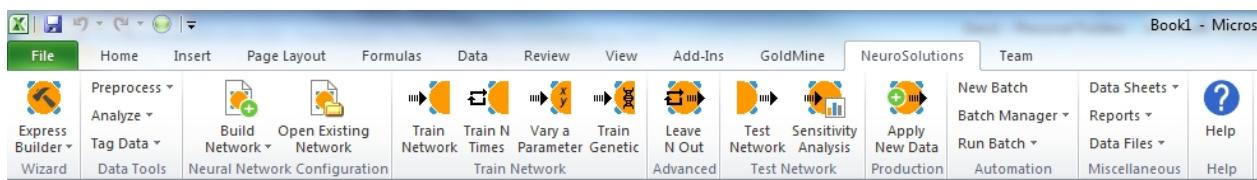


Figure 3. NeuroSolutions for Excel Ribbon Control for Excel 2007 and Later

For the CrabData.csv data file the Species, FrontalLip, RearWidth, Length, Width and Depth are all tagged as Inputs and the Sex is tagged as the DesiredOutput.

	Specime	Frontallip	RearWidth	Length	Width	Depth	(S)Sex
1	nNumber	Species	p	h			
2	142	0	20.6	14.4	42.8	46.5	19.6 Male
3	19	1	13.3	11.1	27.8	32.3	11.3 Male
4	169	0	16.7	14.3	32.3	37	14.7 Female
5	56	1	9.8	8.9	20.4	23.9	8.8 Female
6	164	0	15.6	14.1	31	34.5	13.8 Female
7	53	1	9.1	8.1	18.5	21.6	7.7 Female
8	113	0	14.1	10.5	29.1	31.6	13.1 Male
9	7	1	11.1	9.9	23.8	27.1	9.8 Male

Figure 4. Tagged inputs and desiredOutputs

You may (or not) have noticed that the Sex column has been prefixed with (S). The (S) stands for Symbolic which is used to designate inputs and/or desiredOutputs that are non-numerical. This is very important because neural networks only understand numerical values. So if you are using a column that contains textual information such as True/False, Yes/No, Red/Green/Blue then the column should be also tagged as "Column(s) As Symbolic", which is available in both the dropdown menu from the Excel toolbar or from the right-click context menu.

Specifying Training, Cross Validation and Testing Data Segments

It might be obvious the roles of the training and testing data sets, but we often get asked "What is Cross Validation?" Cross Validation is another set-aside data set that is used during the training process to help prevent the neural network model from overspecializing on the training data. So while the neural network is training, the software also uses the cross validation data to test the network simultaneously to stop the neural network training before it starts becoming too specialized on the training data.

The different data segments are designated by tagging the rows appropriately and it can be easily done from the **Tag Data menu** and choosing the option **Rows by Percentages...**

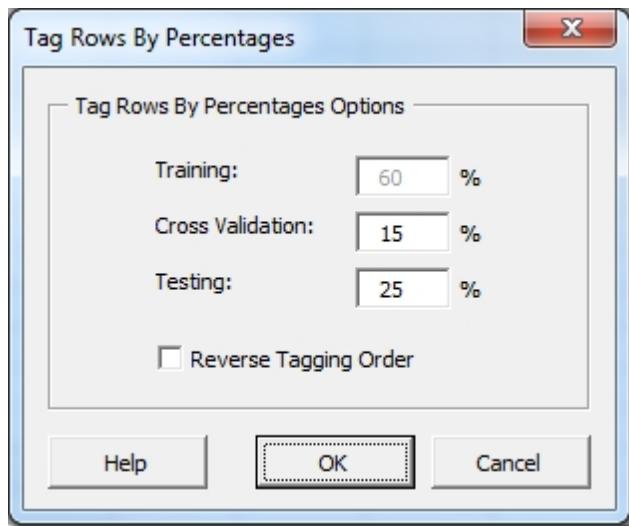


Figure 5. Rows by Percentages dialog

For this example simply choose the default, but you can obviously experiment with these values on your own or even tag the rows manually via right-clicking on a selected row number to bring up the context menu for Tag Data or from the Tag Data menu option.

Building a Neural Network Model

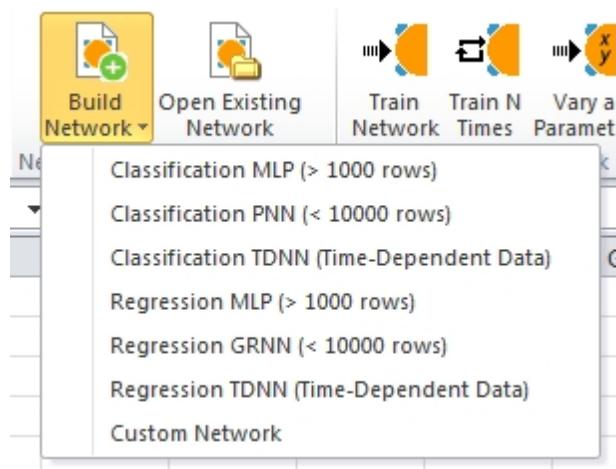


Figure 6. Build Network sub-menu

NeuroSolutions for Excel has access to all of the same supervised networks that NeuroSolutions does, but we offer an easy way to get started quickly with the most common neural network model configurations available in the sub-menu of **Build Network**. For this example, we will choose the default setup for "Classification PNN (< 1000 Rows)" since the CrabData.csv has less than 1,000 total samples. After the network is constructed, go to the File menu and choose Save As and save the breadboard to the same location of your data set (e.g. My Documents).

Training the Neural Network Model

To begin training our neural network mode simply press the **Train Network** button on the Excel toolbar. Before any data is loaded into the neural network in NeuroSolutions it must first verify that the data does not require preprocessing such as any columns tagged as Symbolic. So if you see a message like Figure 7 simply press OK to continue.

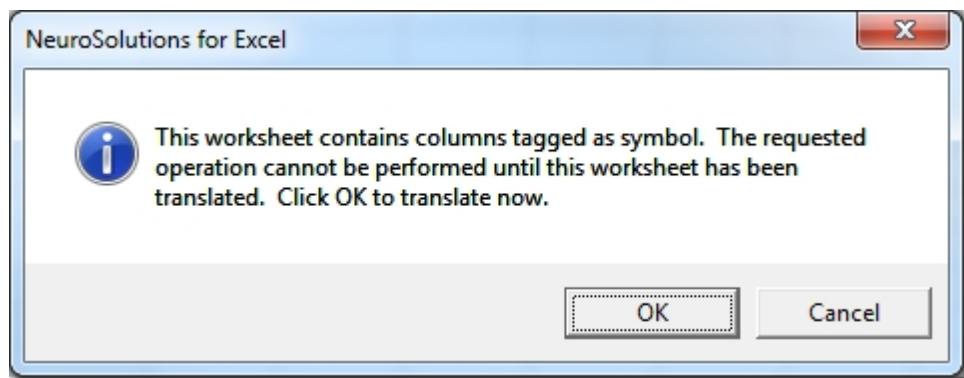


Figure 7. Symbolic translation

Following any required preprocessing the Train dialog will show up (see Figure 8).

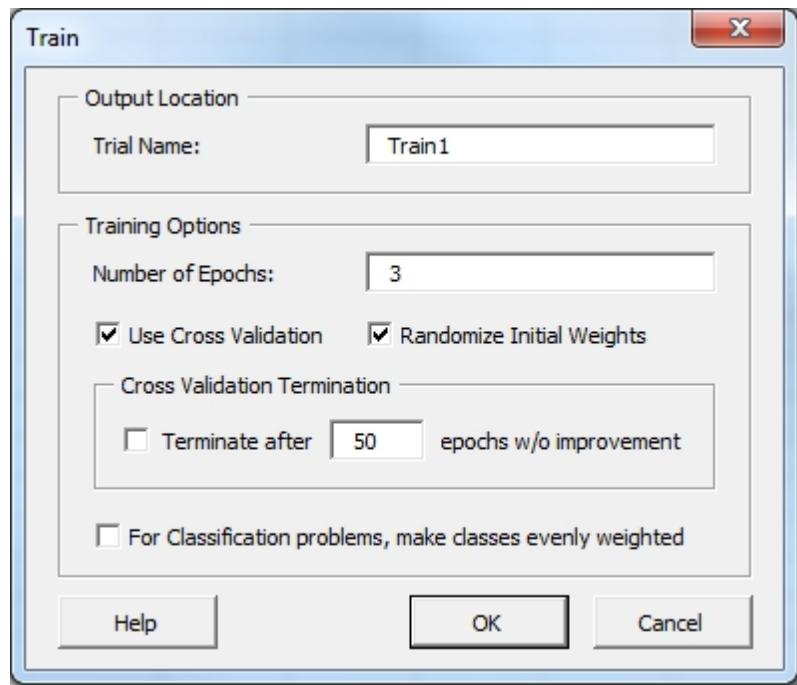
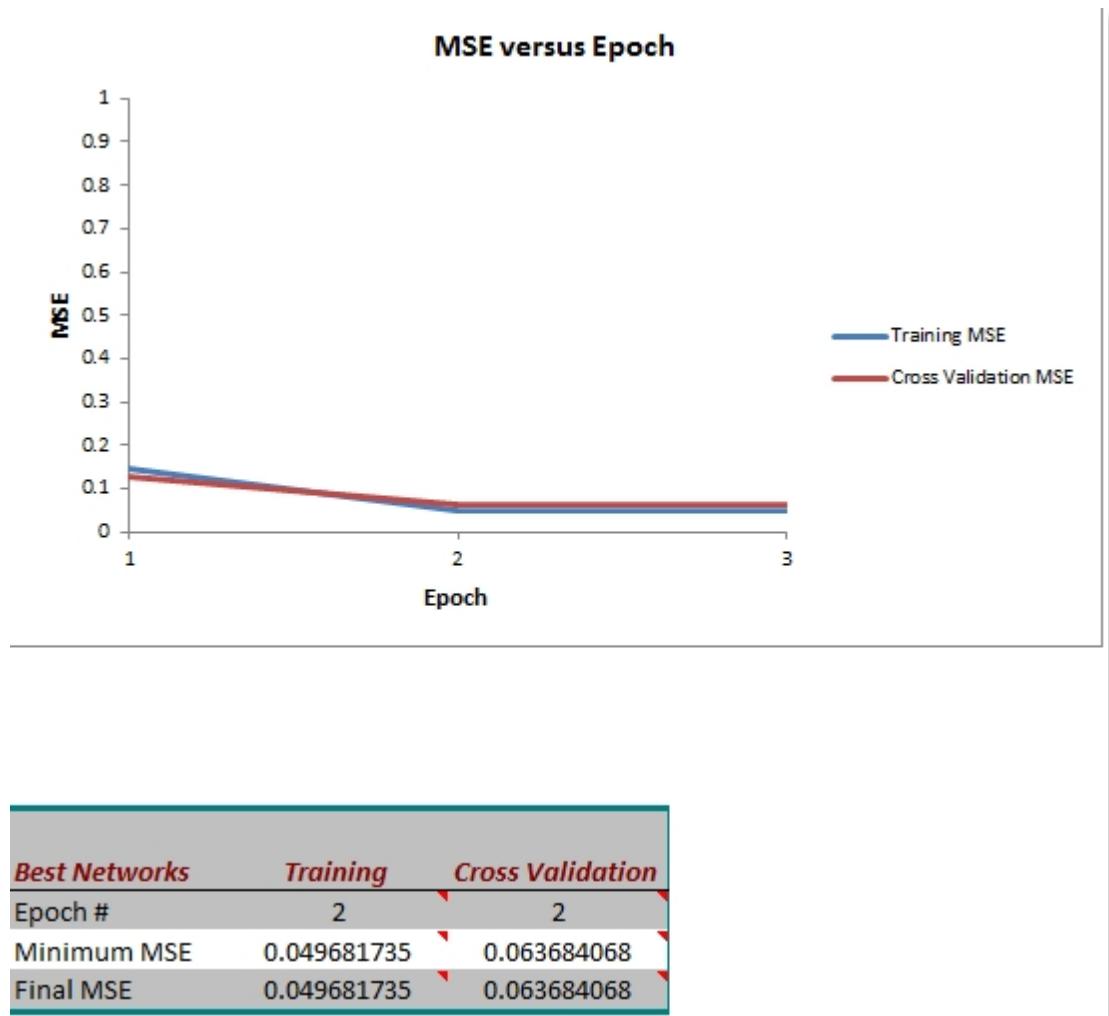


Figure 8. Train dialog

In this dialog you can choose from various items such as the number of epochs, termination and classification weighting. For our CrabData problem we will simply choose the defaults and press OK to continue.

Briefly Examining the Training Results

*Figure 9. Train1 Report*

OK. I trained the model on my data. What does this report even mean?

The chart at the top shows the progression of the Mean Squared Error for the 3 epochs of the training run. While the Training Report can provide valid information about the model's performance let's move on to the testing to see how the default network performed on out of sample data.

Testing the Neural Network on Out of Sample Data

Testing the network is easy. Simply press the Test Network button on the Excel toolbar to launch the Test dialog (see Figure 10).

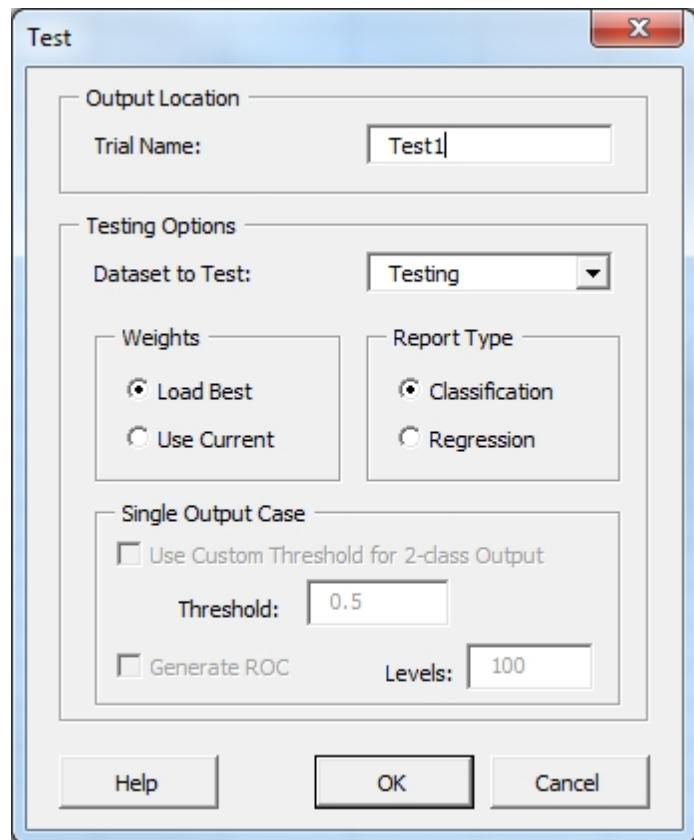


Figure 10. Test Network dialog

The Test Network dialog is designed to make the best choices based on your data such as loading the best weights and the report type. The defaults should be set for the CrabData so simply press OK to continue.

Analyzing the Test Data Results

Output / Desired	Sex(Female)	Sex(Male)
Sex(Female)	20	5
Sex(Male)	7	18

Performance	Sex(Female)	Sex(Male)
RMSE	0.416418073	0.416418073
NRMSE	0.416418073	0.416418073
MAE	0.372650138	0.372650138
Min Abs Error	0.030742446	0.030742446
Max Abs Error	0.758160705	0.758160705
r	0.588203766	0.588203766
Percent Correct	74.07407407	78.26086957

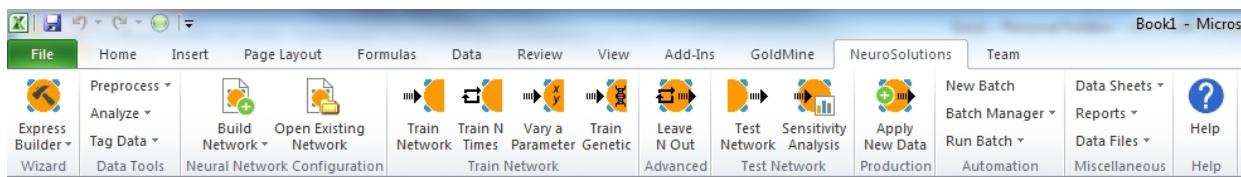
Figure 11. Test Report (for classification)

The end results is an easy to understand report displaying a confusion matrix of the results as well as some general statistics about the performance of the model on the out of sample data. The confusion matrix is read from the top left and goes diagonally to the right displaying the correct predictions. So for this particular training run we classified 20 Females and 18 Males correctly with 5 Males and 7 Females misclassified.

This is just one example of how easy it is to get quick results in NeuroSolutions through our easy to use Excel interface.

Created with the Standard Edition of HelpNDoc: [Single source CHM, PDF, DOC and HTML Help creation](#)

NeuroSolutions Modules in Excel



NeuroSolutions toolbar loaded in Excel 2007 and later

NeuroSolutions for Excel is broken down into 10 major modules. These are:

1. **Express Builder**

Used to create, train and test multiple neural networks simultaneously. This module results in a report that details all of the model(s) results with expanded details on the model with the best results.

2. **Data Tools**

Used to apply various preprocessing techniques to the raw data to prepare it for input into the neural network.

3. **Neural Network Configuration**

Used to create a NeuroSolutions breadboard (neural network) from scratch by utilizing the NeuralBuilder wizard or to open an existing NeuroSolutions breadboard.

4. **Train Network**

Used to train a network using one of several built-in training processes. This powerful module permits the user to easily find the optimum network for a particular problem.

5. **Advanced**

Used to train a network using one of several built-in training processes. This powerful module permits the user to easily find the optimum network for a particular problem.

6. **Test Network**

Used to test a network after the optimum network has been found using the Train Network module. In testing the network, various performance measures are computed. This module can also be used to perform sensitivity analysis on the trained network.

7. **Production**

Used to feed new input data (data tagged as Production Input) to the network and compute the network output. This is how the trained neural model is applied to new data where the output is unknown.

9. **Automation**

Used to create tab a delimited ASCII file for each cross-section tagged using the Tag Data module.

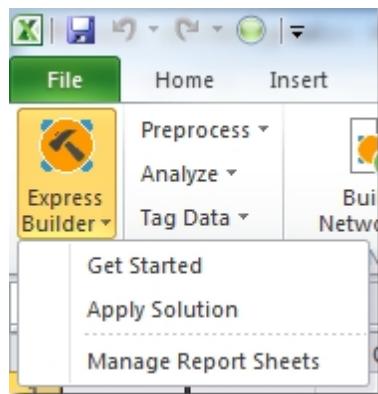
10. **Miscellaneous**

Used to create tab a delimited ASCII file for each cross-section tagged using the Tag Data module.

Each of these modules is a menu / menu item under first partition of the main NeuroSolutions menu.

Created with the Standard Edition of HelpNDoc: [Free Kindle producer](#)

Express Builder Module



Express Builder Menu Options

Menu Item	Description
Get Started	Automatically builds, trains and test multiple NeuroSolutions breadboards and then creates a report of the results.
Apply Solution	Applies the best model from the training run to production data.
Manage Report Sheets	By default all associated training and testing reports are hidden with exception to the final report. This dialog enables the user to display hidden sheets from the Express Builder's simulation.

Created with the Standard Edition of HelpNDoc: [Easy EPub and documentation editor](#)

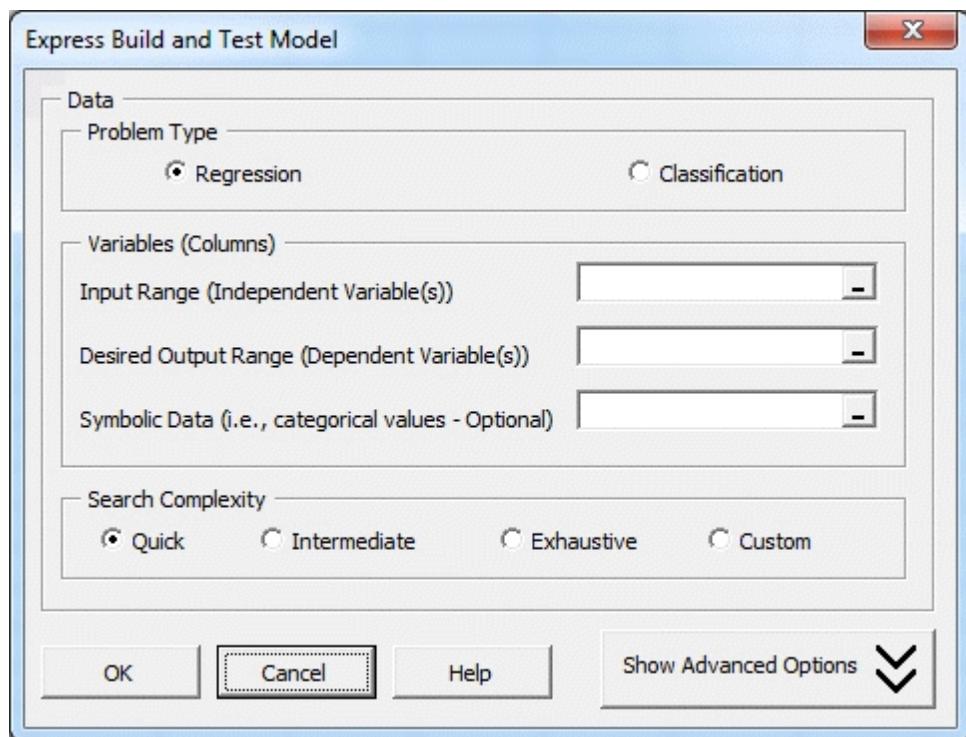
Get Started

Description

Selection of this menu item will open the Express Builder dialog. The Express Builder is a quick and easy way to get started in NeuroSolutions. The Express Builder offers flexibility in creating anywhere from a few different solutions to the entire arsenal of NeuroSolutions topologies. The process includes automatically testing the neural network's trained results and determines the best solution in the final report.

Data Sheet Name	Description
Express Summary	A new worksheet is created for the summarized report from the training and testing simulation.

Dialog Box Options



Problem Type

This option indicates the format of the Express Summary report. Classification problems will focus on confusion matrices and percent correct while Regression problems focus on results with the lowest Mean Squared Error (MSE).

Variables

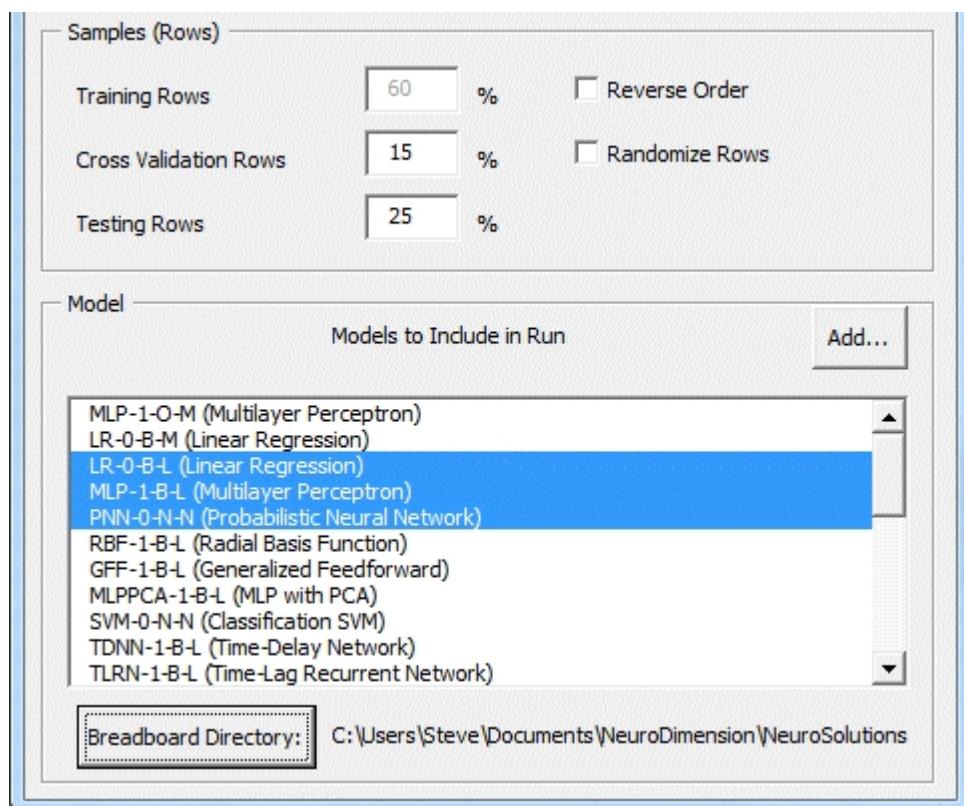
- Input Range (Independent Variable(s)):** Indicates the range for the input variable(s) by selecting the appropriate columns. If the column(s) have already been tagged accordingly, this field will require no further action.
- Desired Output Range (Dependent Variable(s)):** Indicates the range for the desired output variable(s) by selecting the appropriate column(s). If the columns have already been tagged accordingly, this field will require no further action.
- Symbolic Data (i.e. categorical values - Optional):** Indicates the range for the variable(s) that contain categorical values (e.g. Male/Female, Yes/No, etc.). If the column(s) have already been tagged accordingly, this field will require no further action.

Search Complexity

These options indicate the extensiveness that software will search for a solution.

- Quick:** This option will use a set of quick preconstructed networks including Linear Regression, Multilayer Perceptron and Probabilistic Neural Network.
- Intermediate:** This option will use an entire set of preconstructed networks based on your data set size (i.e. Small Data Set vs. Large Data Set). Some of the networks included in this simulation include: Radial Basis Function, Generalized Feedforward SVM (if Classification was selected for the Problem Type) and variances of Multilayer Perceptron networks.
- Exhaustive:** This option will use all of the preconstructed networks available including temporal related networks such as: Time-Delay Network, Time-Lag Recurrent Network and Recurrent Network.
- Custom:** This option will expand the Advanced Options and allow you to select the networks of your choice. Custom breadboards can be added to the list by selecting the Add button.

Advanced Options



Samples (Rows)

Indicates the breakdown of the Training, Cross Validation and Testing ranges for the simulation. If the ranges have already been selected prior to using the Express Builder, it will maintain those settings.

- Reverse Order:** The Reverse Order checkbox indicates to reverse the order of the sample rows by starting from the bottom for Training, Cross Validation and then Testing at the top.
- Randomize Rows:** The Randomize Rows checkbox indicates to randomize the data into a new worksheet in Excel.

Model

The model section indicates the preconstructed breadboards available to run the simulation. Additional breadboards can be added to the list by selecting the Add button and browsing to the appropriate directory.

Note: Models must be selected from a list, by holding down the *Ctrl* key while selecting the models to include. To select a block of models, select the first model in the block, then hold the *Shift* key and select the last model in the block.

Breadboard Directory

The Breadboard Directory button indicates the default location of where the breadboards will be saved. Clicking the button and browsing to the preferred directory can change this directory.

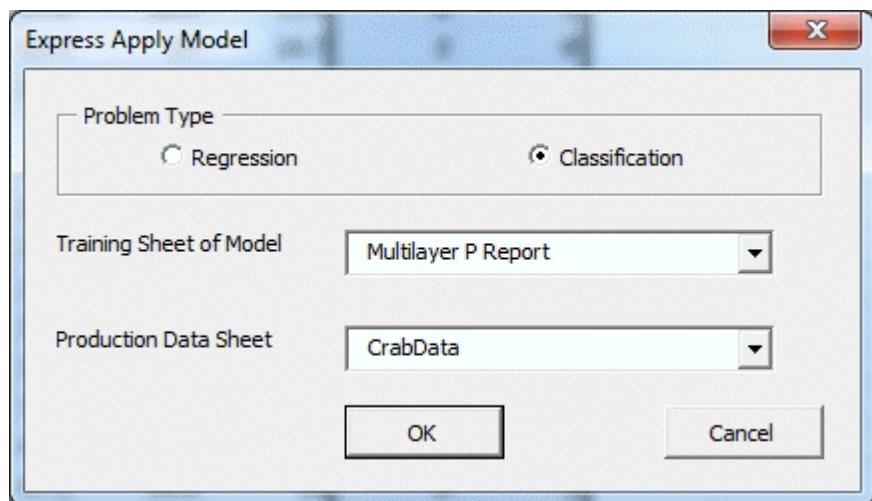
Created with the Standard Edition of HelpNDoc: [Full-featured Documentation generator](#)

Apply Solution

Description

Selection of this menu item will open the Express Apply Solution dialog. The Express Apply Solution dialog allows the user to apply the trained model from the Express Builder simulation to new (production) data in a separate worksheet.

Dialog Box Options



Problem Type

This option indicates the format of the Express Summary report. Classification problems will focus on confusion matrices and percent correct while Regression problems focus on the results with the lowest Mean Squared Error (MSE).

Training Sheet of Model

This option indicates which of the models the test results will be based on. The default selection is the model with the "best results" from the Express Summary.

Production Data Set

This option indicates the worksheet with the input data for the model to be applied to.

Created with the Standard Edition of HelpNDoc: [Free CHM Help documentation generator](#)

Manage Report Sheets

Description

Selection of this menu item will open the Manage Report Sheets dialog. The Manage Reports Sheets dialog enables users to display previously hidden worksheets by the Express Builder.

Dialog Box Options



Training Sheet of Model

This dropdown menu indicates the model that the reports should take action for. The Hide Reports button indicates that all of the selected model's reports should be hidden. The Show Reports button indicates that all of the selected model's reports should be shown. The Delete Reports button indicates that all of the selected model's reports should be permanently deleted.

Manage Report Sheets for All Models

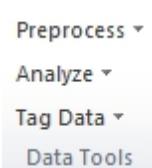
These three options indicate the chosen action affects All Models in the Excel workbook. The Hide All Reports button indicates that all of the model's reports should be hidden. The Show All Reports button indicates that all of the model's reports should be shown. The Delete All Reports button indicates that all of the model's reports should be permanently deleted.

Manage Report Sheets for Best-Performing Model

These two options indicate the chosen action affects All Models EXCEPT the Best Model in the Excel workbook. The Hide Reports for All Except Best button indicates that all of the model's reports should be hidden with exception to the Best. The Delete Reports for All Except Best button indicates that all of the model's reports should be permanently deleted with exception to the Best.

Created with the Standard Edition of HelpNDoc: [Easily create Help documents](#)

Data Tools Module



Data Tools Menu Options

Menu Item	Description

[Preprocess Data](#)

Various tools for preprocessing data including shifting data and randomizing.

[Analyze](#)

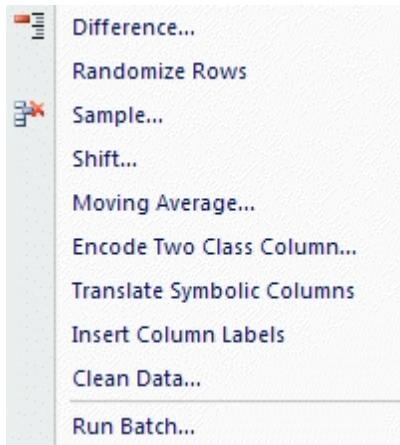
Various tools for visually representing your data including scatter plots and histograms.

[Tag Data](#)

Various methods of tagging both rows and columns as well as segmenting the data into various data sets.

Created with the Standard Edition of HelpNDoc: [Full-featured multi-format Help generator](#)

Preprocess Data



Preprocess Menu Options

Menu Item	Description
Difference	Computes the difference or percent difference along a column of data.
Randomize Rows	Randomly arranges the rows of data and writes the result to a new sheet.
Sample	Creates a new worksheet made up of every Nth row of data within the active worksheet.
Shift	Shifts the dataset up or down by a user specified number of rows.
Moving Average	Computes the moving average of a column using the chosen window length.
Translate Symbolic Columns	Translates columns that have been tagged as symbol.
Encode Two Class Column	Translates single columns that have been verified to contain two classes.
Insert Column Labels	Inserts a row of column labels into the first row of the active worksheet.
Clean Data	Cleans the data by replacing blank cells, error codes, and/or user-defined text with an interpolated value, the column average, a random value, or the closest value in a column.

Created with the Standard Edition of HelpNDoc: [Easy to use tool to create HTML Help files and Help web sites](#)

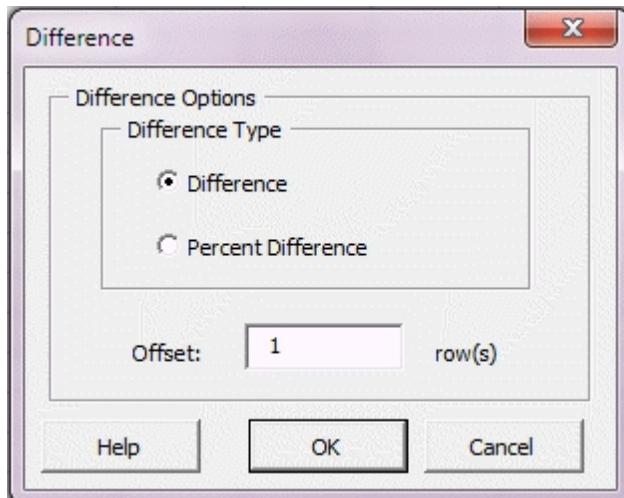
Difference

Description

This menu item is used to find the difference or percent difference of a column of data. Before selecting this operation, you must select a single column or multiple columns. If multiple columns are selected, they may be contiguous or non-contiguous. **Note:** To properly select a column, the **entire** column must be selected (see [Selecting Columns and Rows](#)). The results of the differencing operation will be written starting at the first blank

column in the current worksheet.

Dialog Box Options



Difference

Choose this option if you want to compute the difference between data values in a column.

Percent Difference

Choose this option if you want to compute the percent difference between data values in a column.

Offset

Set the value you want to compute for either difference or percent difference between data values in a column.

Created with the Standard Edition of HelpNDoc: [Easily create Help documents](#)

Randomize Rows

Description

Selection of this menu item will result in the creation of a new worksheet named “<Active Worksheet Name> Randomized” where <Active Worksheet Name> is the name of the active worksheet when this menu item was selected. All of the data on the active worksheet will be copied to this new worksheet. The rows of data on this new worksheet will then be randomized. If the original data had already been tagged, the column tags will be maintained, but the row tags will be cleared.

This function is useful if you want to randomly tag exemplars (rows) within your data. To accomplish this task you first run the randomize rows function. Then within the “Randomized Data” worksheet, you tag the rows in blocks (either manually using [Row\(s\) As Training](#), [Row\(s\) As Cross Validation](#), or [Row\(s\) As Testing](#), or automatically using [Rows By Percentages](#)). This process is equivalent to randomly tagging rows within your original data worksheet.

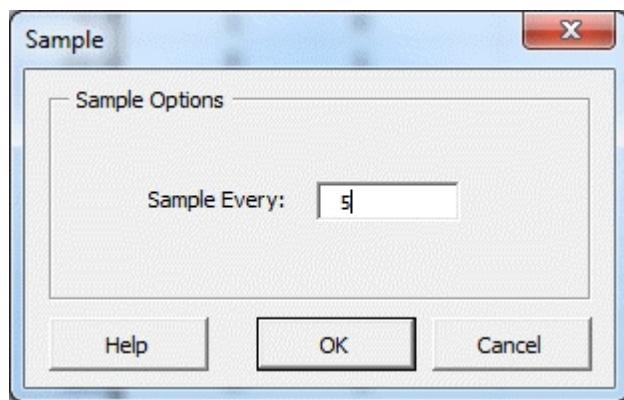
Created with the Standard Edition of HelpNDoc: [Generate EPub eBooks with ease](#)

Sample

Description

This menu item is used to periodically sample your data. Selection of this menu item will result in the creation of a new worksheet named “<Active Worksheet Name> Sampled” where <Active Worksheet Name> is the name of the active worksheet when this menu item was selected. The labels and the first row of data followed by every Nth row of data will be written to this new worksheet.

Dialog Box Options



Sample Every

Enter the sampling period into this edit box.

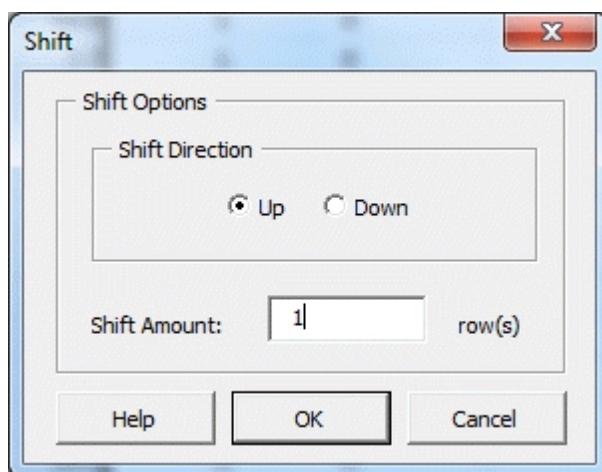
Created with the Standard Edition of HelpNDoc: [Easily create Web Help sites](#)

Shift

Description

When this menu item is selected, the input data is adjusted to either move the inputs back by a specified shift value to do predictions or move the inputs forward to lead your desired output.

Dialog Box Options



Shift Direction

Choose the direction that you want to shift your selected data.

Shift Amount

Enter the number of rows that the data will be shifted by in the direction selected.

Created with the Standard Edition of HelpNDoc: [Easily create Help documents](#)

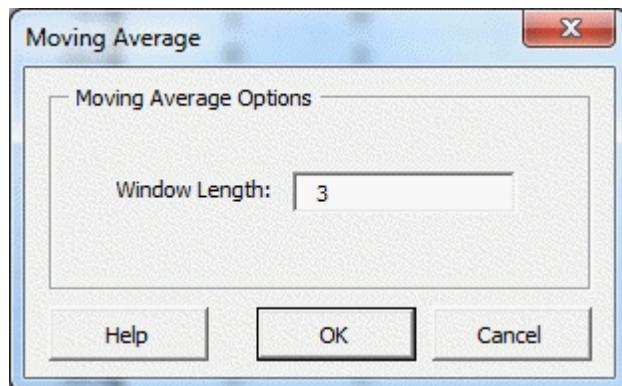
Moving Average

Description

This menu item is used to find the moving average of a column of data. Before selecting this operation, you must select a single column or multiple columns. If multiple columns are selected, they may be contiguous or non-contiguous. **Note:** To properly select a column, the **entire** column must be selected (see [Selecting Columns and Rows](#)).

The moving average is the average value of a variable computed using a sliding window. A moving average provides trend information that would not be evident in the average over all of a variables historical data. The moving average of the selected column will be written in the first blank column in the current worksheet.

Dialog Box Options



Window Length

Enter the number of data values to use for computing each average value.

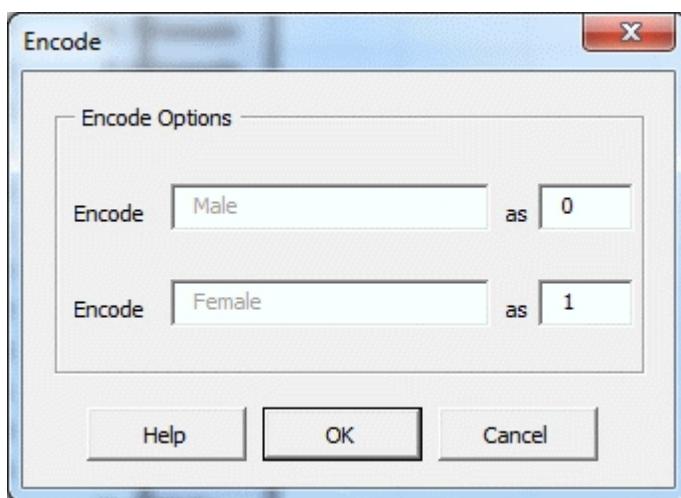
Created with the Standard Edition of HelpNDoc: [Full-featured multi-format Help generator](#)

Encode Two Class Column

Description

When this menu item is selected, the selected column will be checked to verify that there are two classes contained within the column and is then encoded in another column. The data to be encoded can be textual or numeric, but must be translated to only numeric, integer codes. The encoded column will be written in the first empty column in the dataset.

Dialog Box Options



Encode Options

The classes from the selected column will be visible in the dialog box. The desired code value is entered in the "as" textbox. Codes must be whole, numeric values.

Created with the Standard Edition of HelpNDoc: [Write eBooks for the Kindle](#)

Translate Symbolic Columns

Description

Selection of this menu item will result in the translation of all columns tagged as [Symbol](#). A new worksheet named “<Active Worksheet Name> Translated” (where <Active Worksheet Name> is the name of the active worksheet when this menu item was selected) will be created. The columns of data on the original worksheet will be copied column by column to the new worksheet. When this procedure encounters a column tagged as Symbol, this column will first be translated (using unary encoding) and the result of this translation will be copied to the new worksheet.

Note: When a column of data is translated using unary encoding, one column is created for each unique symbol. The label for each of the translated columns becomes “<Original Column Label> (<Symbol>)” where <Original Column Label> is the label of the original column and <Symbol> is the symbol that the 1's in the translated column represents. Within a translated column, a “1” signifies the existence of the corresponding symbol and a “0” signifies the existence of any other symbol.

Below are two examples of symbol translation:

	A	B	C
1	X	Y	{S}Z
2	-1	-1	-1
3	-1	1	1
4	1	-1	1
5	1	1	-1

X-Or data before symbol translation

	A	B	C	D
1	X	Y	Z{1}	Z{-1}
2	-1	-1	0	1
3	-1	1	1	0
4	1	-1	1	0
5	1	1	0	1

X-Or data after symbol translation

	F	G	H
	Width	Depth	{S}Sex
46.5	19.6	Male	
32.3	11.3	Male	
37	14.7	Female	
23.9	8.8	Female	

Gender Classification Data before translation

	F	G	H	I
	Width	Depth	Sex{Female}	Sex{Male}
46.5	19.6	0	1	
32.3	11.3	0	1	
37	14.7	1	0	
23.9	8.8	1	0	

Gender Classification Data after translation

Insert Column Labels

Description

Selection of this menu item will result in the insertion of a row of unique column labels. The column labels will be inserted into the first row of the active worksheet. The naming convention used for these labels is as follows:

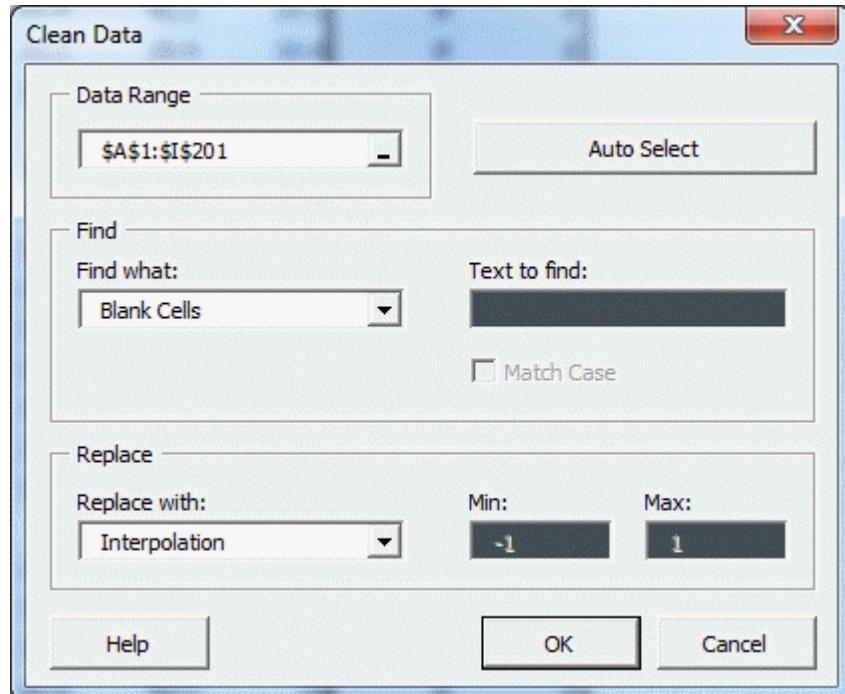
- 1.Columns tagged as Input are labeled sequentially as Input1, Input2, Input3, etc.
- 2.Columns tagged as Desired are labeled sequentially as Desired1, Desired 2, Desired 3, etc.
- 3.Untagged columns are labeled as Column# where # is the actual column number.

Clean Data

Description

This menu item is used to search for and replace invalid data. If your data contains missing values, Excel error codes, and/or text values, this operation can be used to replace these invalid values. The invalid values can be replaced using interpolation, the column average, a random value within the specified range, or the closest value in a column.

Dialog Box Options



Data Range

Enter the range of the data you want to clean. The range will be automatically entered by clicking inside the Data Range text box followed by using the mouse to select a range of cells on the data worksheet. The default range is the entire range of data found on the active sheet.

Auto Select

Click to restore the Data Range to the default range (the contiguous block of data starting in cell A1 and extending outward).

Find what

Choose the type of invalid data to find from one of the following choices:

- 1.Blank Cells – Cells that don't contain data.

- 2.Error Codes – Cells that contain Excel error codes such as "#DIV/0!".
- 3.User Specified Text – Cells that contain the textual data specified in the "Text to Find" text box.
- 4.All – Cells that contain all of the above 3 types of invalid data.

Text to find

Enter the invalid text to search for. Available only when the "Find what" choice is set to "User Specified Text" or "All".

Match Case

Check this box to make the text searching for the "User Specified Text" or "All" options case sensitive (i.e. the text in the "Text to Find" text box must match exactly).

Replace with

Choose the method to use for replacing the invalid data from one of the following choices:

- 1.Interpolation – Replaces invalid data with linearly interpolated values. To perform linear interpolation, the slope between the two closest valid values surrounding the invalid value is calculated. The interpolated value is then determined by adding the slope times the distance (from the first valid value) to the first valid value. For example, given a value of 2 in cell A2 and a value of 27 in cell A7, the slope would be 5 = [27 - 2] / [7 - 2], and the interpolated values for cells A3, A4, A5 and A6 would be 7, 12, 17, and 22, respectively. Note: Invalid values found at the endpoints (i.e. not between two valid values) are extrapolated in a manner similar to interpolation.
- 2.Column Average – Replaces invalid data with the average of all the numeric values in the column in which the invalid data resides.
- 3.Random Value – Replaces invalid data with a random value generated within a user-specified range.
- 4.Closest Value In Column – Replaces invalid data with the nearest valid value (in terms of proximity) in the column in which the invalid data resides.

Min

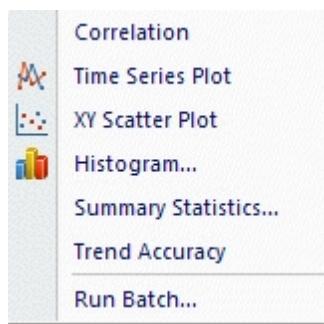
Enter the minimum bound to use for the random number generation when the "Replace with" method is set to "Random Value".

Max

Enter the maximum bound to use for the random number generation when the "Replace with" method is set to "Random Value".

Created with the Standard Edition of HelpNDoc: [Easy to use tool to create HTML Help files and Help web sites](#)

Analyze



Analyze Menu Options

Menu Item	Description
<u>Correlation</u>	Computes the correlation between the each of the columns of data on the active worksheet.
<u>Time Series Plot</u>	Creates a Times Series Plot of the selected columns.
<u>XY Scatter Plot</u>	Creates an XY Scatter Plot of the selected columns.
<u>Histogram</u>	Computes the histogram of a selected column of data.
<u>Summary Statistics</u>	Computes various statistics for a selected column of data.

[Trend Accuracy](#)

Compute the trend accuracy of the selected columns.

Created with the Standard Edition of HelpNDoc: [Full-featured EPub generator](#)

Correlation

Description

Measures the relationship between columns of data. The returned correlation between two columns will be near 1 if the data in the two columns tend to move together, near -1 if the data in the two columns tend to move in inverse directions, or near 0 if the data in the two columns are unrelated. The computed correlations are presented in a table format within a new worksheet named “<Active Worksheet Name> Correlation” where <Active Worksheet Name> is the name of the active worksheet when this menu item was selected.

Created with the Standard Edition of HelpNDoc: [Free EPub producer](#)

Time Series Plot

Description

Creates a time series plot of the selected column(s) of data. If multiple columns are selected, they can be non-contiguous. The plot will be created on the active worksheet.

Created with the Standard Edition of HelpNDoc: [Produce Kindle eBooks easily](#)

XY Scatter Plot

Description

Creates an XY scatter plot of the selected column(s) of data. The first column of data will be used for the x-axis. The remaining column(s) will be used as y-axis traces. At least two columns of data must be selected (1 for the x-axis and 1 for the y-axis). The selected columns can be non-contiguous. The plot will be created on the active worksheet.

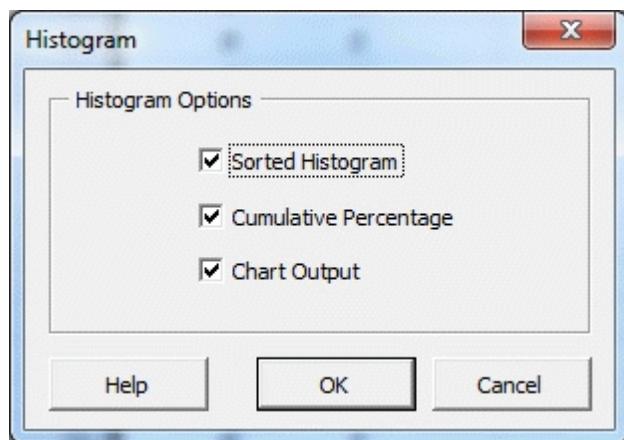
Created with the Standard Edition of HelpNDoc: [Write eBooks for the Kindle](#)

Histogram

Description

This menu item is used to find the histogram of a column of data. Before selecting this operation, you must select a single column of data. The histogram is calculated by counting the number of times (frequency) that a data value within the chosen column falls between a set of data bins equally spaced over the entire data range. This provides a useful indication of the relative distribution of the data values. The results are written to a new worksheet named “<Active Worksheet Name> Histogram” where <Active Worksheet Name> is the name of the active worksheet when this menu item was selected. By default, the results are presented in a table, but can also be plotted.

Dialog Box Options



Sorted Histogram

Check this box if you want to include a column in the output table which presents the result in descending order of frequency.

Cumulative Percentage

Check this box if you want to include a column in the output table containing the cumulative percentages.

Chart Output

Check this box if you want the function to produce a plot of the histogram. If the **Cumulative Percentage** box is checked, it will also be plotted.

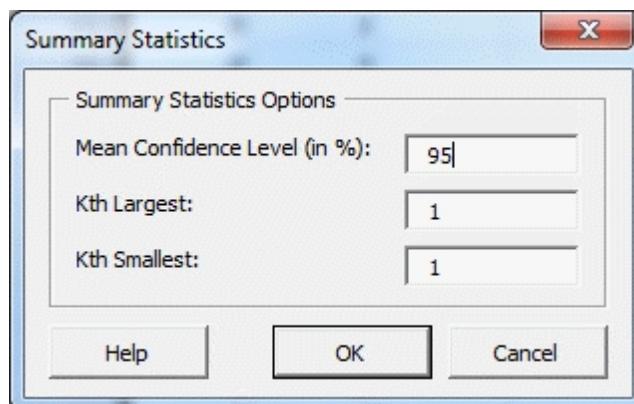
Created with the Standard Edition of HelpNDoc: [Free help authoring environment](#)

Summary Statistics

Description

Computes several key statistics for the selected column(s) of data. The column(s) must be selected before clicking this menu item. Multiple columns can be selected, but they must be contiguous. The default statistics computed are Mean, Standard Error (of the mean), Median, Mode, Standard Deviation, Variance, Kurtosis, Skewness, Range, Minimum, Maximum, Sum, and Count. The results are presented in a tabular format. The results are written to a new worksheet named "<Active Worksheet Name> Statistics" where <Active Worksheet Name> is the name of the active worksheet when this menu item was selected.

Dialog Box Options



Mean Confidence Level (in %)

Check this box if you want to include a row in the output table for the mean confidence level. Then enter the confidence level you want to use into the corresponding edit box.

Kth Largest

Check this box if you want to include a row in the output table which reports the kth largest value. In the corresponding edit box, enter the number to use for k.

Kth Smallest

Check this box if you want to include a row in the output table which reports the kth smallest value. In the corresponding edit box, enter the number to use for k.

Created with the Standard Edition of HelpNDoc: [Easily create PDF Help documents](#)

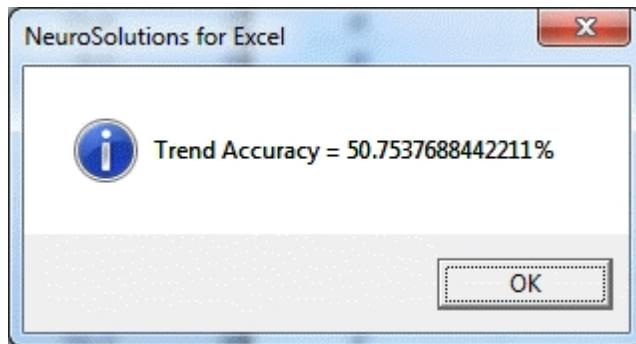
Trend Accuracy

Description

This menu item computes the trend accuracy of the selected columns of data. Exactly two columns must be selected before clicking this menu item. The first column of data selected will be interpreted as the desired network output and the second column of data selected will be interpreted as the actual network output. The two selected columns do not have to be contiguous.

Important: The interpretation of a column as the desired network output or the actual network output is based on the *order in which the columns were selected*, not the physical location of the columns.

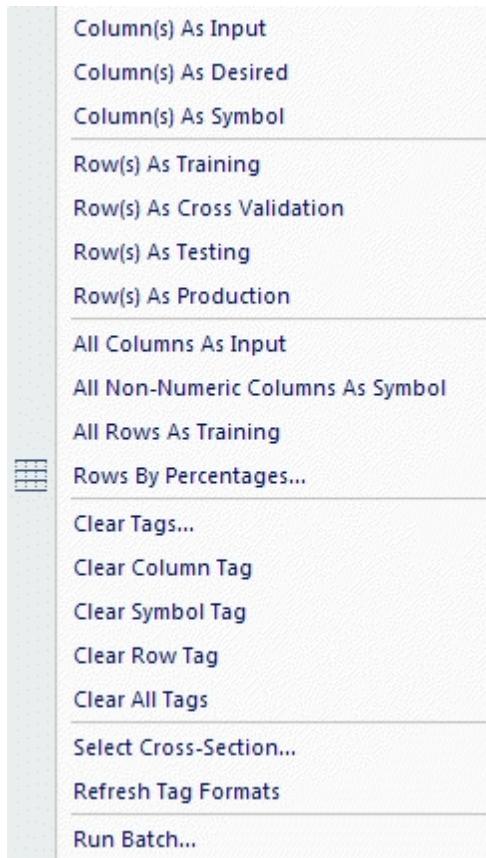
The trend accuracy measure is useful when working with time series data. It gives the percentage for which the actual output changed in the correct direction relative to the previous desired value. The following example demonstrates how Trend Accuracy is calculated.



Calculation of Trend Accuracy

Created with the Standard Edition of HelpNDoc: [Easily create EBooks](#)

Tag Data



Tag Data Menu Options

Menu Item	Description
Column(s) As Input	Tags the selected column(s) of data as Input.
Column(s) As Desired	Tags the selected column(s) of data as Desired.
Column(s) As Symbol	Tags the selected column(s) of data as Symbol.
Row(s) As Training	Tags the selected row(s) of data as Training.
Row(s) As Cross Validation	Tags the selected row(s) of data as Cross Validation.
Row(s) As Testing	Tags the selected row(s) of data as Testing.
Row(s) As Production	Tags the selected row(s) of data as Production.
All Columns As Input	Tags all columns as Input.
All Non-Numeric Columns As Symbol	Tags all non-numeric columns as symbol.
All Rows As Training	Tags all rows as Training.
Rows By Percentages	Tags the rows of data within the active worksheet as Training, Cross Validation, and Testing according to user-defined percentages.
Clear Tags	Allows you to clear any existing tag.
Clear Column Tag	Clears the tag(s) of the selected column(s).
Clear Symbol Tag	Clears the symbol tag for the selected column(s).
Clear Row Tag	Clear the tag(s) of the selected row(s).
Clear All Tags	Clears all of the tags on the active worksheet.
Select Cross-Section	Allows you to automatically select any existing cross-section.
Refresh Tag Formats	Refreshes the tag formatting.

Column(s) As Input

Description

Allows you to tag columns of data as Input. If a column is not tagged, the data in that column will not be used in the creation of the data files. To tag a column of data as Input, you must select the entire column (see [Selecting Columns and Rows](#)) then select the **Tag Data** menu item followed by the **Column(s) As Input** sub-menu item. Multiple (contiguous or non-contiguous) columns can be selected and tagged all at once. The following text describes the coloring/formatting scheme for the Input tag.

Input Tag Coloring/Formatting

If no row tags exist, the cells for a column tagged as Input will be colored **aqua** and the data will be colored **gray**. If any row tags exist, the cells (for the entire worksheet) will be colored **white** (the default color) and the color of the data (in the column tagged as Input) will depend upon the tag type for its corresponding row.

Created with the Standard Edition of HelpNDoc: [Full-featured Kindle eBooks generator](#)

Column(s) As Desired

Description

Allows you to tag columns of data as Desired. If a column is not tagged, the data in that column will not be used in the creation of the data files. To tag a column of data as Desired, you must select the entire column (see [Selecting Columns and Rows](#)) then select the **Tag Data** menu item followed by the **Column(s) As Desired** sub-menu item. Multiple (contiguous or non-contiguous) columns can be selected and tagged all at once. The following text describes the coloring/formatting scheme for the Desired tag.

Desired Tag Coloring/Formatting

If no row tags exist, the cells for a column tagged as Desired will be colored **aqua** and the data will be colored **gray** and formatted as bold italic. If any row tags exist, the cells (for the entire worksheet) will be colored **white** (the default color), the color of the data (in the column tagged as Desired) will depend upon the tag type for its corresponding row, and the data (in the column tagged as Desired) will be formatted as bold italic.

Created with the Standard Edition of HelpNDoc: [Easily create PDF Help documents](#)

Column(s) As Symbol

Description

Allows you to tag columns of data as Symbol. To tag a column of data as Symbol, you must select the entire column (see [Selecting Columns and Rows](#)) then select the **Tag Data** menu item followed by the **Column(s) As Symbol** sub-menu item. Multiple (contiguous or non-contiguous) columns can be selected and tagged all at once. When you tag a column as symbol, the "(S)" prefix will be appended to the beginning of that column's label. Any column with this prefix will be recognized as a symbolic column by NeuroSolutions for Excel. Columns tagged as Symbol are unary encoded when you click the [Translate Symbolic Columns](#) sub-menu item from the *Preprocess Data* menu item.

A Column should be tagged as symbol if either of the following two conditions exist:

- 1.The column contains non-numeric data.
- 2.The desired output column should be tagged as symbol if the problem being solved is a classification problem. For the *classification* option of the [Test](#) procedure to function properly, the desired output must be unary encoded.

Created with the Standard Edition of HelpNDoc: [Full-featured EBook editor](#)

Row(s) As Training

Description

Allows you to tag rows of data as Training. If a row is not tagged, the data in that row will not be used in the creation of the data files. To tag a row of data as Training, you must select the entire row (see [Selecting Columns and Rows](#)) then select the **Tag Data** menu item followed by the **Row(s) As Training** sub-menu item.

Multiple (contiguous or non-contiguous) rows can be selected and tagged all at once. The following text describes the coloring/formatting scheme for the Training tag.

Training Tag Coloring/Formatting

If no column tags exist, the cells for a row tagged as Training will be colored **black** and the data will be colored **gray**. If any column tags exist, the cells (for the entire worksheet) will be colored **white** (the default color), the data that exists within a tagged column will be colored **black**, and this data will be formatted based on the particular column tag.

Created with the Standard Edition of HelpNDoc: [Produce Kindle eBooks easily](#)

[Row\(s\) As Cross Validation](#)

Description

Allows you to tag rows of data as Cross Validation. If a row is not tagged, the data in that row will not be used in the creation of the data files. To tag a row of data as Cross Validation, you must select the entire row (see [Selecting Columns and Rows](#)) then select the **Tag Data** menu item followed by the **Row(s) As Cross Validation** sub-menu item. Multiple (contiguous or non-contiguous) rows can be selected and tagged all at once. The following text describes the coloring/formatting scheme for the Cross Validation tag.

Cross Validation Tag Coloring/Formatting

If no column tags exist, the cells for a row tagged as Cross Validation will be colored **red** and the data will be colored **gray**. If any column tags exist, the cells (for the entire worksheet) will be colored **white** (the default color), the data that exists within a tagged column will be colored **red**, and this data will be formatted based on the particular column tag.

Created with the Standard Edition of HelpNDoc: [Easy to use tool to create HTML Help files and Help web sites](#)

[Row\(s\) As Testing](#)

Description

Allows you to tag rows of data as Testing. If a row is not tagged, the data in that row will not be used in the creation of the data files. To tag a row of data as Testing, you must select the entire row (see [Selecting Columns and Rows](#)) then select the **Tag Data** menu item followed by the **Row(s) As Testing** sub-menu item. Multiple (contiguous or non-contiguous) rows can be selected and tagged all at once. The following text describes the coloring/formatting scheme for the Testing tag.

Testing Tag Coloring/Formatting

If no column tags exist, the cells for a row tagged as Testing will be colored **blue** and the data will be colored **gray**. If any column tags exist, the cells (for the entire worksheet) will be colored **white** (the default color), the data that exists within a tagged column will be colored **blue**, and this data will be formatted based on the particular column tag.

Created with the Standard Edition of HelpNDoc: [Easy to use tool to create HTML Help files and Help web sites](#)

[Row\(s\) As Production](#)

Description

Allows you to tag rows of data as Production. If a row is not tagged, the data in that row will not be used in the creation of the data files. To tag a row of data as Production, you must select the entire row (see [Selecting Columns and Rows](#)) then select the **Tag Data** menu item followed by the **Row(s) As Production** sub-menu item. Multiple (contiguous or non-contiguous) rows can be selected and tagged all at once. The following text describes the coloring/formatting scheme for the Production tag.

Testing Tag Coloring/Formatting

If no column tags exist, the cells for a row tagged as Testing will be colored **green** and the data will be colored **gray**. If any column tags exist, the cells (for the entire worksheet) will be colored **white** (the default color), the

data that exists within a tagged column will be colored **green**, and this data will be formatted based on the particular column tag.

Created with the Standard Edition of HelpNDoc: [Easy EPub and documentation editor](#)

All Columns As Input

Description

Provides a quick method for tagging multiple columns of data. Select this sub-menu item if you want to tag all of the columns in the active worksheet as Input.

Created with the Standard Edition of HelpNDoc: [Free EPub and documentation generator](#)

All Non-Numeric Columns As Symbol

Description

Provides a quick method for tagging multiple columns of data. Select this sub-menu item if you want to tag all of the non-numeric columns in the active worksheet as Symbol. If a column of data contains any non-numeric values (except for labels), it will be tagged as Symbol.

Created with the Standard Edition of HelpNDoc: [Free EPub producer](#)

All Rows As Training

Description

Provides a quick method for tagging multiple rows of data. Select this sub-menu item if you want to tag all of the rows in the active worksheet as Training.

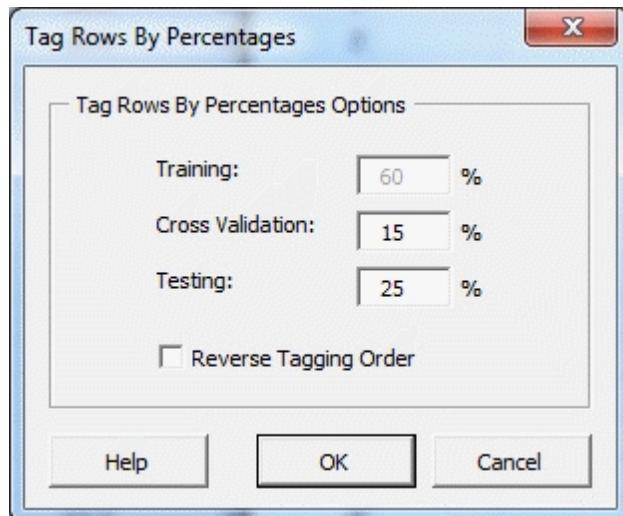
Created with the Standard Edition of HelpNDoc: [Create HTML Help, DOC, PDF and print manuals from 1 single source](#)

Rows By Percentages

Description

Provides a quick method for tagging multiple rows of data. Select this sub-menu item if you want to tag user-specified percentages of your data as Training, Cross Validation, and Testing. A dialog box will display allowing you to enter the percentages (of the total number of rows) to tag as Cross Validation and Testing. The remainder of the data will be tagged as Training.

Dialog Box Options



Training

Displays the percentage of data that will be used for Training. This cell cannot be changed manually. It will automatically be updated based on changes in the Cross Validation and Testing percentages.

Cross Validation

Enter the percentage of data to use for Cross Validation.

Testing

Enter the percentage of data to use for Testing.

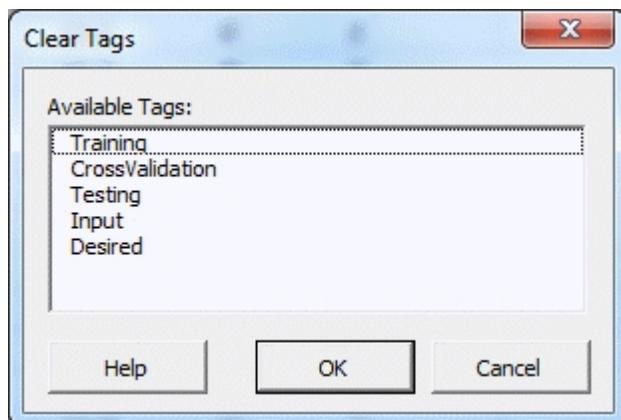
Reverse Tagging Order

By default, the data will be tagged in the following order: Training, Cross Validation, Testing. However, checking this option will cause the tagging order to be reversed such that the data is tagged in the following order: Testing, Cross Validation, Training.

Created with the Standard Edition of HelpNDoc: [Single source CHM, PDF, DOC and HTML Help creation](#)

[Clear Tags](#)**Description**

Select this sub-menu item if you want to clear one or more of the available tags on the active worksheet. This will display a dialog box showing all of the tags that exist on the active worksheet. Select one or more of these tags then click OK to clear the tags you selected.

Dialog Box Options**Available Tags**

Choose the tag or tags that you want to clear. When you click OK, the selected tags will be cleared.

Created with the Standard Edition of HelpNDoc: [Easy EBook and documentation generator](#)

[Clear Column Tag](#)**Description**

Select this sub-menu item if you want to clear the tag(s) for the selected column(s). Multiple (contiguous or non-contiguous) columns can be selected and have their tags cleared all at once. If multiple columns are selected, they may have different tag types (Input or Desired).

Created with the Standard Edition of HelpNDoc: [Write eBooks for the Kindle](#)

[Clear Symbol Tag](#)**Description**

Select this sub-menu item if you want to clear the Symbol tag for the selected column(s). Multiple (contiguous or non-contiguous) columns can be selected and have their Symbol tags cleared all at once. This operation has

no affect when performed on columns that are not tagged as symbol.

Created with the Standard Edition of HelpNDoc: [Easily create HTML Help documents](#)

[Clear Row Tag](#)

Description

Select this sub-menu item if you want to clear the tag(s) for the selected row(s). Multiple (contiguous or non-contiguous) rows can be selected and have their tags cleared all at once. If multiple rows are selected, they may have different tag types (Training, Cross Validation, Testing, or Production).

Created with the Standard Edition of HelpNDoc: [Full-featured EBook editor](#)

[Clear All Tags](#)

Description

Select the sub-menu item if you want to clear all of the tags for the active worksheet.

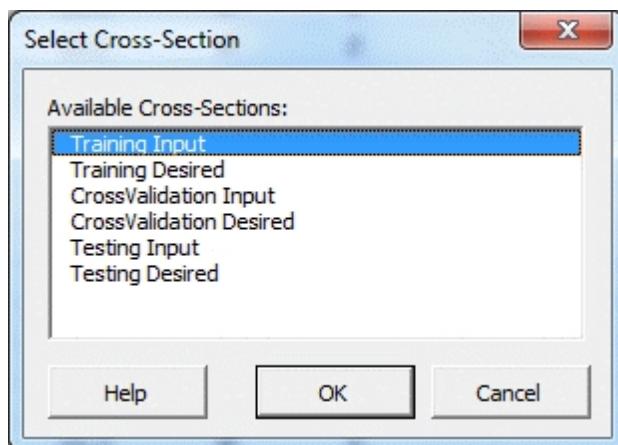
Created with the Standard Edition of HelpNDoc: [Easily create iPhone documentation](#)

[Select Cross-Section](#)

Description

Select this sub-menu item if you want to select one of the available cross-sections. This will display a dialog box showing all of the available cross-sections. Choose the cross-section that you want to select then click OK and the cross-section will be highlighted. This gives you a visual indication of the different sections of data. Furthermore, since the cross-section is automatically selected, you can easily cut or copy the data for a particular cross-section.

Dialog Box Options



Available Cross-Sections

Choose the cross-section that you want to select. When you click OK, the chosen cross-section will be selected.

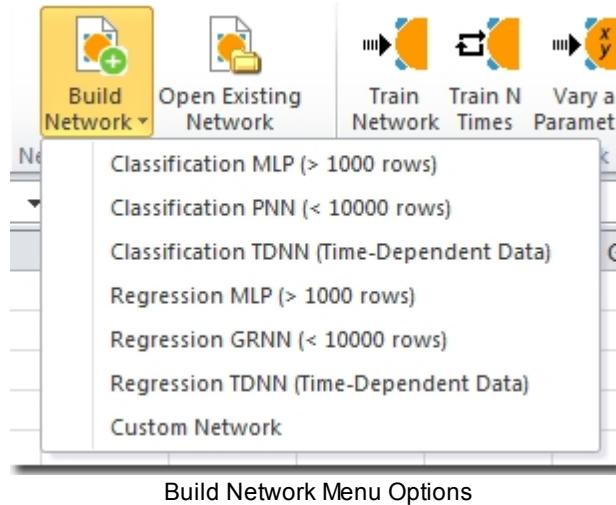
Created with the Standard Edition of HelpNDoc: [Produce Kindle eBooks easily](#)

[Refresh Tag Formats](#)

Description

Refreshes the tag coloring after you have made changes to the spreadsheet such as inserting another column of data. After selecting this menu item, the colors of the spreadsheet will again reflect what data is actually tagged as well as the tag type.

Neural Network Configuration



Build Network Menu Options

Menu Item	Description
New Classification Network	Creates a new NeuroSolutions breadboard with typical elements used for a classification problem.
New Function Approximation Network	Creates a new NeuroSolutions breadboard with typical elements used for a function approximation problem.
New Custom Network	Starts the NeuralBuilder which guides you step-by-step through the creation of a new NeuroSolutions breadboard.

Classification MLP

Classification MLP

Description

When this menu item is selected, a NeuroSolutions breadboard is created and Excel and NeuroSolutions are automatically tiled. The breadboard is a one-hidden-layer Multilayer Perceptron network (MLP) prepared with all the components for solving a classification problem. It is generally recommended for data sets larger than 1000 rows to avoid over-fitting, but smaller data sets can be used.

Classification PNN

Classification PNN

Description

When this menu item is selected, a NeuroSolutions breadboard is created and Excel and NeuroSolutions are automatically tiled. The breadboard is a Probabilistic Neural Network (PNN) prepared with all the components for solving a classification problem. It is generally recommended for data sets smaller than 10,000 rows to avoid a network with too many weights, but larger data sets can be used.

Classification TDNN

Classification TDNN

Description

When this menu item is selected, a NeuroSolutions breadboard is created and Excel and NeuroSolutions are automatically tiled. The breadboard is a focused Time-Delay Neural Network (TDNN) prepared with all the components for solving a classification problem with time-dependent data. Note that you want to the desired output be a future value of one of your data columns, you'll first need to create a shifted version of the column (see Shift) and then tag that column as Desired.

Created with the Standard Edition of HelpNDoc: [Free HTML Help documentation generator](#)

Regression MLP

Regression MLP

Description

When this menu item is selected, a NeuroSolutions breadboard is created and Excel and NeuroSolutions are automatically tiled. The breadboard is a one-hidden-layer Multilayer Perceptron network (MLP) prepared with all the components for solving a regression (i.e., function approximation) problem. It is generally recommended for data sets larger than 1000 rows to avoid over-fitting, but smaller data sets can be used.

Created with the Standard Edition of HelpNDoc: [Single source CHM, PDF, DOC and HTML Help creation](#)

Regression GRNN

Regression GRNN

Description

When this menu item is selected, a NeuroSolutions breadboard is created and Excel and NeuroSolutions are automatically tiled. The breadboard is a General Regression Neural Network (GRNN) prepared with all the components for solving a regression (i.e., function approximation) problem. It is generally recommended for data sets smaller than 10,000 rows to avoid a network with too many weights, but larger data sets can be used.

Created with the Standard Edition of HelpNDoc: [Free EPub producer](#)

Regression TDNN

Regression TDNN

Description

When this menu item is selected, a NeuroSolutions breadboard is created and Excel and NeuroSolutions are automatically tiled. The breadboard is a focused Time-Delay Neural Network (TDNN) prepared with all the components for solving a regression problem with time-dependent data (i.e., time-series prediction). Note that you want to the desired output be a future value of one of your data columns, you'll first need to create a shifted version of the column (see Shift) and then tag that column as Desired.

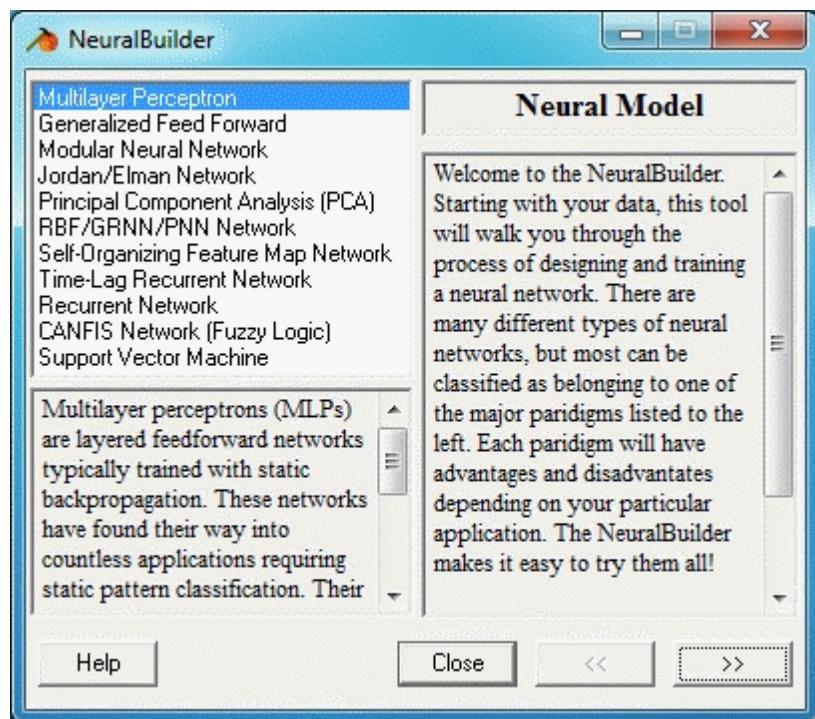
Created with the Standard Edition of HelpNDoc: [Easy to use tool to create HTML Help files and Help web sites](#)

New Custom Network

New Custom Network

Description

Selection of this menu item will open the NeuralBuilder.



The NeuralBuilder leads you through a series of panels that prompt you for information regarding the network you want to create. The *Build* button on the last panel will open NeuroSolutions (if it's not already open) and automatically create the network that you specified.

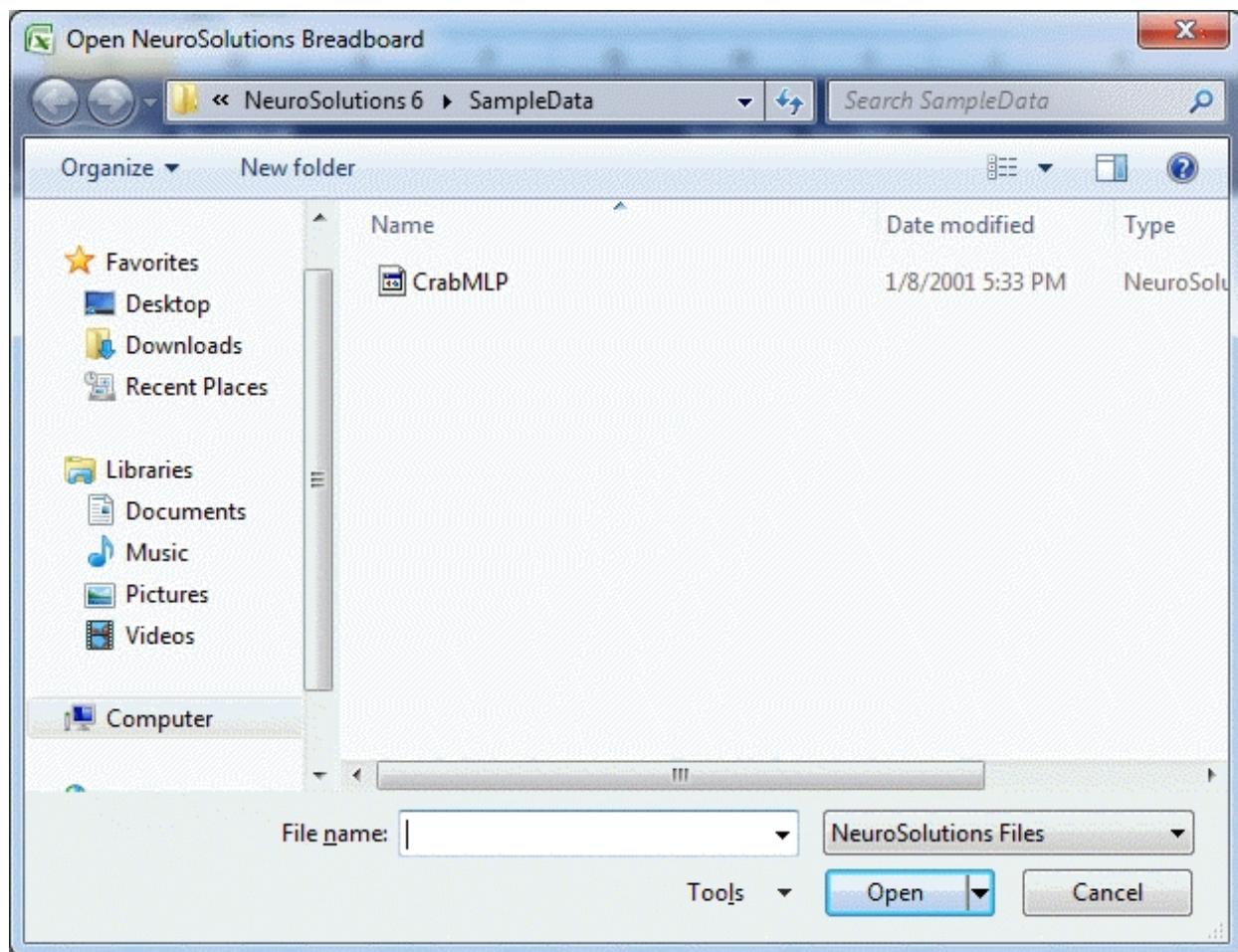
If the sheet that is active when you click the *New Custom Network* menu item contains tagged data, the default network parameters (number of PEs and *Step Sizes* for each layer) will be chosen intelligently based on the number of columns tagged as *Input*, the number of columns tagged as *Desired*, and the number of rows tagged as *Training*.

Created with the Standard Edition of HelpNDoc: [Easily create iPhone documentation](#)

Open Existing Network

Description

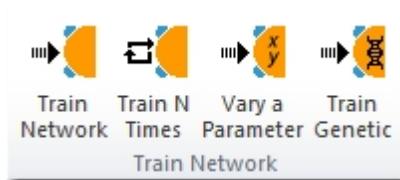
Selection of this menu item will display the Open NeuroSolutions Breadboard dialog box.



Use this dialog box to transverse the file system on your computer and find the breadboard file that you want to open. Double click this file or select it and click the *Open* button. This will run NeuroSolutions (if it is not already running) and the selected breadboard will be opened within NeuroSolutions. When you open a NeuroSolutions breadboard, a hidden sheet named *NSVariables* is created (if it does not already exist) within the active workbook. This sheet keeps track of the active *Breadboard Path* and the *Breadboard File Name*. If the *NSVariables* sheet already exists and another breadboard is already referenced within this sheet, you will be given the choice of whether or not to make the breadboard just opened the active breadboard. If you choose to make it the active breadboard, the *Breadboard Path* and the *Breadboard File Name* will be updated to reference the breadboard just opened.

Created with the Standard Edition of HelpNDoc: [Full-featured multi-format Help generator](#)

Train Network Module



Train Network Menu Options

Menu Item	Description
Train	Trains the active NeuroSolutions breadboard one time and creates a report of the results.
Train N Times	Trains the active NeuroSolutions breadboard N times with different random initial conditions and creates a report of the

results.

Vary a Parameter

Trains the active NeuroSolutions breadboard N times for each value of a network parameter as the parameter is varied from a user defined starting value by a user defined increment for a user defined number of variations.

Leave N Out

Trains the active NeuroSolutions breadboard multiple times leaving out different sections of data for each training run. This training procedure is very useful for testing the robustness of a model on small datasets.

Train Genetic

Trains the active NeuroSolutions breadboard while genetically optimizing the choice of inputs and parameter values to achieve the best model.

Note for users of 64-bit Excel: Due to a limitation of the programming environment under 64-bit Excel, the progress bars shown during training will only have the numerical values and not the graphical bars.

Created with the Standard Edition of HelpNDoc: [iPhone web sites made easy](#)

Train

Description

The active NeuroSolutions breadboard is trained one time and the best network weights are saved (Note: The best network weights are saved at the epoch when the cross validation error is minimum if a cross validation data set is used or the epoch at which the training error is minimum if only a training data set is used). A report of the training results is then generated. The generated report contains the following information:

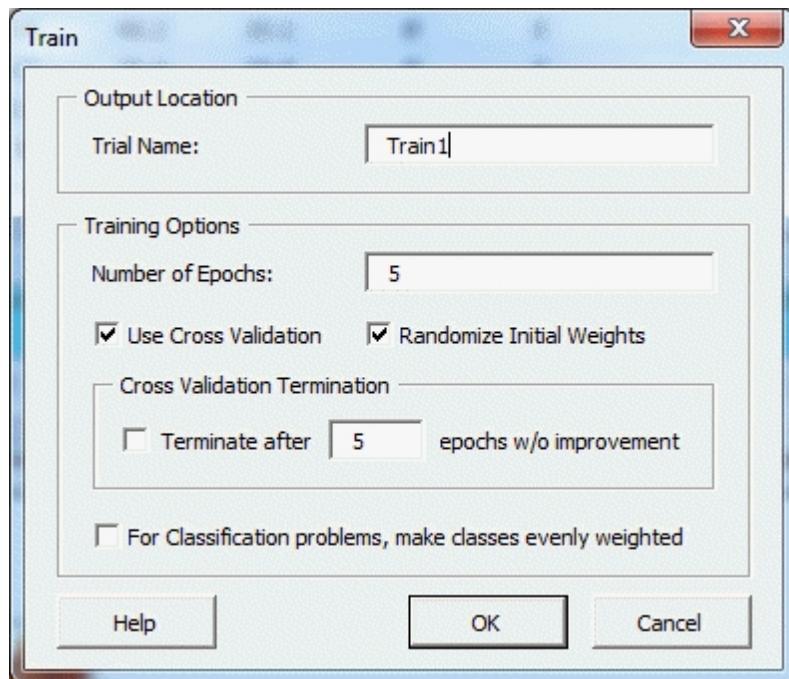
1. Plot of the training mean-squared error (MSE) versus Epochs. If the active worksheet contains Cross Validation Input and Cross Validation Desired tags and the user checks **Use Cross Validation** (see below), this plot will also show the cross validation MSE.
2. Table showing the minimum training MSE, the epoch at which this minimum training MSE occurred, and the final training MSE. If the active worksheet contains Cross Validation Input and Cross Validation Desired tags and the user checks **Use Cross Validation** (see below), this table will include the minimum cross validation MSE, the epoch at which this minimum cross validation error occurred, and the final cross validation MSE.

The following data sheets are also created:

Data Sheet Name	Description
TrialName MSE	Contains the training MSE for each epoch in the run (learning curve). Also, the minimum training MSE is given at the bottom of this column. If cross validation was used, the same results are also shown for the cross validation data set.

Note that the MSE in this report is actually one half of the normalized MSE, which should match what is reported by the Average Cost probe of NeuroSolutions.

Dialog Box Options



Trial Name

Enter the name for this particular training trial run. If the entered name has already been used, you will be given the option to overwrite it.

Number of Epochs

Enter the total number of epochs to train the network for. An epoch is defined as one complete presentation of all of the data. By default, this setting is initialized to the Epochs / Run setting of the active NeuroSolutions breadboard.

Use Cross Validation

Check this box if you want to use cross validation during training. Note: You will not be able to check this box if the active worksheet does not contain data tagged as Cross Validation.

Randomize Initial Weights

Check this box if you want to randomize the networks weights before training (this is the default). Unchecking this box allows you to start from where a previous training run left off.

Cross Validation Termination

Check this box to enable the termination of the training run if the cross validation error has not improved within the user-specified number of epochs. Note: You will not be able to check this box if the active worksheet does not contain data tagged as Cross Validation or if the **Use Cross Validation** option is unchecked.

Make Classes Evenly Weighted

Check this box to tell NeuroSolutions to evenly weight each class in a classification problem. This will enable the "Exemplar Weighting" feature of NeuroSolutions and will cause the gradients for each class to be weighted proportionately according to the number of samples of that class present in the training dataset. For classification problems in which one or more classes are disproportionately represented, using this feature will usually help the network to produce a more balanced model. Specifically, the underrepresented classes will be given more weight, which will usually improve their classification. For example, suppose you have a two-class problem where 99% of your samples are of the Class 1 type and 1% are of the Class 2 type. Without exemplar weighting, the network will likely take the path of least resistance and arrive at a model that classifies all of your data as Class 1. Overall, the network will be 99% accurate, but this is probably not what you want since it will be 0% correct for Class 2. With exemplar weighting, the gradients (used for updating the weights) for Class 2 samples will be weighted 99 times more than the gradients for Class 1 samples. This will usually result in a more correct classification percentage for Class 2 samples. For more information on "Exemplar Weighting", see the "Exemplar Weighting Inspector" help topic in the NeuroSolutions help. Note: This feature should only be enabled for classification problems.

Train N Times

Description

Trains the active NeuroSolutions breadboard N times (where N is chosen by the user) with different random initial conditions and the best network weights are saved (Note: The best network weights are saved at the run and epoch when the cross validation error is minimum if a cross validation data set is used or the run and epoch at which the training error is minimum if only a training data set is used). A report is then generated showing the results. The generated report contains the following information:

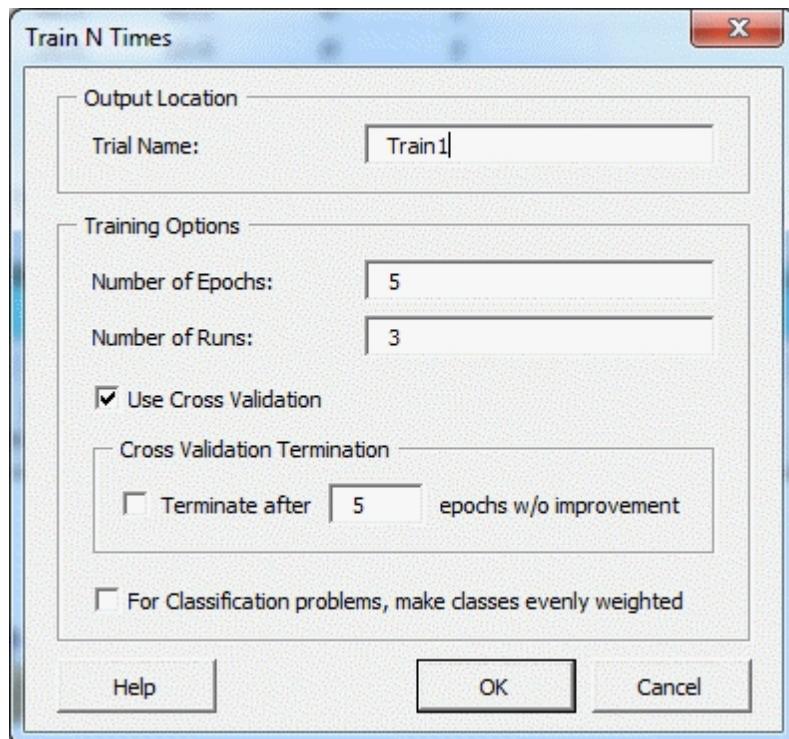
1. Plot of the average training mean-squared error (MSE) versus Epochs (the average is computed by averaging the MSE of the N runs at each Epoch). The high (average + 1 standard deviation) and low (average - 1 standard deviation) standard deviation boundaries are also shown on the plot using dashed lines of the same color as the average line (the standard deviation is computed across the N runs at each Epoch). If the active worksheet contains Cross Validation Input and Cross Validation Desired tags and the user checks **Use Cross Validation** (see below), this plot will also show the average cross validation MSE and the corresponding standard deviation boundaries.
2. Table showing the average of the minimum training MSEs and the average of the final training MSEs over the N runs along with their corresponding standard deviations. If the active worksheet contains Cross Validation Input and Cross Validation Desired tags and the user checks **Use Cross Validation** (see below), this table will include the average of the minimum cross validation MSEs and the average of the final cross validation MSEs over the N runs and their corresponding standard deviations.
3. Table reporting the best training MSE out of all N runs, the epoch and run # at which this best training MSE occurred, and the final training MSE for this best run. If the active worksheet contains Cross Validation Input and Cross Validation Desired tags and the user checks **Use Cross Validation** (see below), this table also presents the best cross validation MSE out of all N runs, the epoch and run # at which this best cross validation MSE occurred, and the final cross validation MSE for this best run.
4. Plot of the training MSE versus Epoch for each of the N runs.
5. Plot of the cross validation MSE versus Epoch for each of the N runs if the active worksheet contains Cross Validation Input and Cross Validation Desired tags and the user checks **Use Cross Validation** (see below).

The following data sheets are also created:

Data Sheet Name	Description
TrialName TrnMSE	Contains the training MSEs for each run. Also, the minimum training MSE for each run is given at the bottom of its corresponding column.
TrialName CVMSE	Contains the cross validation MSEs for each run. Also, the minimum cross validation MSE for each run is given at the bottom of its corresponding column (this sheet is only created if cross validation is used).
TrialName AvgMSE	Contains the average training MSE, average cross validation MSE (if cross validation is used), and corresponding standard deviation boundaries.

Note that the MSE measurements in this report are actually one half of the normalized MSE, which should match what is reported by the Average Cost probe of NeuroSolutions.

Dialog Box Options



Trial Name

Enter the name for this particular training trial run. If the entered name has already been used, you will be given the option to overwrite it.

Number of Epochs

Enter the total number of epochs to train the network for. An epoch is defined as one complete presentation of all of the data. By default, this setting is initialized to the Epochs / Run setting of the active NeuroSolutions breadboard.

Number of Runs

Enter the number of times you want to run this network. Since each run begins with different random initial weights, the more times you run the network the more statistically sound your results will be.

Use Cross Validation

Check this box if you want to use cross validation during training. Note: You will not be able to check this box if the active worksheet does not contain data tagged as Cross Validation.

Cross Validation Termination

Check this box to enable the termination of a training run if the cross validation error has not improved within the user-specified number of epochs. Note: You will not be able to check this box if the active worksheet does not contain data tagged as Cross Validation or if the **Use Cross Validation** option is unchecked.

Make Classes Evenly Weighted

Check this box to tell NeuroSolutions to evenly weight each class in a classification problem. This will enable the "Exemplar Weighting" feature of NeuroSolutions and will cause the gradients for each class to be weighted proportionately according to the number of samples of that class present in the training dataset. For classification problems in which one or more classes are disproportionately represented, using this feature will usually help the network to produce a more balanced model. Specifically, the underrepresented classes will be given more weight, which will usually improve their classification. For example, suppose you have a two-class problem where 99% of your samples are of the Class 1 type and 1% are of the Class 2 type. Without exemplar weighting, the network will likely take the path of least resistance and arrive at a model that classifies all of your data as Class 1. Overall, the network will be 99% accurate, but this is probably not what you want since it will be 0% correct for Class 2. With exemplar weighting, the gradients (used for updating the weights) for Class 2 samples will be weighted 99 times more than the gradients for Class 1 samples. This will usually result in a more correct classification percentage for Class 2 samples. For more information on "Exemplar Weighting", see the "Exemplar Weighting Inspector" help topic in the NeuroSolutions help. Note: This feature should only be enabled for classification problems.

Vary a Parameter

Description

For this training process, the user picks a network parameter to vary. This parameter can be selected from a default list or typed in. Next the starting value of the chosen network parameter is entered along with the desired number of variations and the size of the increment between each variation. When the user clicks *OK*, the active NeuroSolutions breadboard is trained N times (where N is also chosen by the user) for each value of the network parameter and the best network weights are saved (Note: The best network weights are saved at the parameter variation, run, and epoch when the cross validation error is minimum if a cross validation data set is used or the parameter variation, run, and epoch at which the training error is minimum if only a training data set is used). After training, a report is automatically generated summarizing the results. The generated report contains the following information:

1. Plot of the average of the minimum training mean-squared errors (MSEs) for each value of the varied parameter. The high (average + 1 standard deviation) and low (average - 1 standard deviation) standard deviation boundaries are also shown on the plot using dashed lines of the same color as the average line. If the active worksheet contains Cross Validation Input and Cross Validation Desired tags and the user checks **Use Cross Validation** (see below), this plot will also show the average of the minimum cross validation MSEs for each value of the varied parameter and the corresponding standard deviation boundaries.
2. Table reporting the best training MSE out of all N runs and all values of the varied parameter, the parameter value, epoch, and run # at which this best training MSE occurred, and the final training MSE for this best run. If the active worksheet contains Cross Validation Input and Cross Validation Desired tags and the user checks **Use Cross Validation** (see below), this table also presents the best cross validation MSE out of all N runs and all values of the varied parameter, the parameter value, epoch, and run # at which this best cross validation MSE occurred, and the final cross validation MSE for this best run.
3. Plot of the average training MSE versus Epochs (the average is computed by averaging the MSE of the N runs at each Epoch). This average is calculated and a curve is displayed for each value of the varied parameter.
4. If the active worksheet contains Cross Validation Input and Cross Validation Desired tags and the user checks **Use Cross Validation** (see below), a plot will be generated for the average cross validation MSE versus Epochs (the average is computed by averaging the MSE of the N runs at each Epoch). This average is calculated and a curve is displayed for each value of the varied parameter.

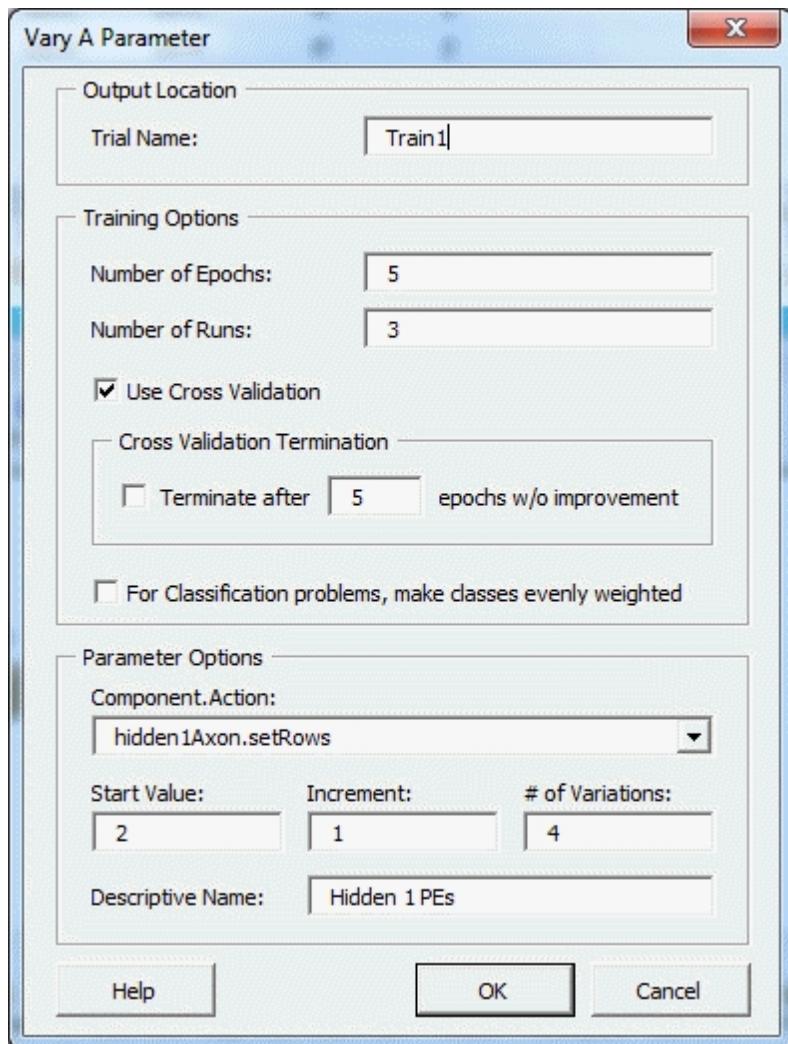
The following data sheets are also created:

Data Sheet Name	Description
TrialName TrnMSE(v#)	Contains the training MSEs for all N runs. Also, the minimum training MSE for each run is given at the bottom of its corresponding column. The # sign shown at the left signifies the variation number. Thus, a separate sheet is created for each parameter value.
TrialName CVMSE(v#)	Contains the cross validation MSEs for each run. Also, the minimum cross validation MSE for each run is given at the bottom of its corresponding column (this sheet is only created if cross validation is used). The # sign shown at the left signifies the variation number. Thus, a separate sheet is created for each parameter value.
TrialName AvgMinMSE	Contains the average of the minimum MSEs calculated over N runs and the corresponding standard deviations. These values are computed for each value of the varied parameter. This data is also computed for the cross validation set if cross validation is used.
TrialName AvgTrnMSE	Contains the average training MSE (averaged over the N runs) for each epoch. Each column contains this data for a different value of the varied parameter.
TrialName AvgCVMSE	If cross validation is used, the worksheet contains the average

cross validation MSE (averaged over the N runs) for each epoch. Each column contains this data for a different value of the varied parameter.

Note that the MSE measurements in this report are actually one half of the normalized MSE, which should match what is reported by the Average Cost probe of NeuroSolutions.

Dialog Box Options



Trial Name

Enter the name for this particular training trial run. If the entered name has already been used, you will be given the option to overwrite it.

Number of Epochs

Enter the total number of epochs to train the network for. An epoch is defined as one complete presentation of all of the data. By default, this setting is initialized to the Epochs / Run setting of the active NeuroSolutions breadboard.

Number of Runs

Enter the number of times you want to run this network. Since each run begins with different random initial weights, the more times you run the network the more statistically sound your results will be.

Use Cross Validation

Check this box if you want to use cross validation during training. Note: You will not be able to check this box if the active worksheet does not contain data tagged as Cross Validation.

Cross Validation Termination

Check this box to enable the termination of a training run if the cross validation error has not improved within the user-specified number of epochs. Note: You will not be able to check this box if the active worksheet does not

contain data tagged as Cross Validation or if the **Use Cross Validation** option is unchecked.

Make Classes Evenly Weighted

Check this box to tell NeuroSolutions to evenly weight each class in a classification problem. This will enable the "Exemplar Weighting" feature of NeuroSolutions and will cause the gradients for each class to be weighted proportionately according to the number of samples of that class present in the training dataset. For classification problems in which one or more classes are disproportionately represented, using this feature will usually help the network to produce a more balanced model. Specifically, the underrepresented classes will be given more weight, which will usually improve their classification. For example, suppose you have a two-class problem where 99% of your samples are of the Class 1 type and 1% are of the Class 2 type. Without exemplar weighting, the network will likely take the path of least resistance and arrive at a model that classifies all of your data as Class 1. Overall, the network will be 99% accurate, but this is probably not what you want since it will be 0% correct for Class 2. With exemplar weighting, the gradients (used for updating the weights) for Class 2 samples will be weighted 99 times more than the gradients for Class 1 samples. This will usually result in a more correct classification percentage for Class 2 samples. For more information on "Exemplar Weighting", see the "Exemplar Weighting Inspector" help topic in the NeuroSolutions help. Note: This feature should only be enabled for classification problems.

Component.Action

Use the combo box to select the parameter that you want to vary or manually enter a custom one into the box. The parameter is chosen using the form **Component.Action** where **Component** is the name of the NeuroSolutions component which contains the parameter to be varied (Note: this name is always shown on the Engine page of a components inspector) and **Action** is the action that you want to perform on the chosen component. For example, hiddenAxon1.setRows would be chosen if the parameter you want to vary is the number of rows (processing elements) in the first hidden layer. Note: If the word "selection" is used for the component name, the action is performed on the currently selected component(s) on the NeuroSolutions breadboard. This is useful if you want to vary the same parameter on multiple components at the same time. For example, to vary the *step size* for all of the Momentum components on a breadboard at once, you would select all of the Momentum components (by holding down the Shift key and clicking on each Momentum component one at a time) and use *selection.setStepSize* as the **Component.Action**.

Start Value

Enter the starting value for the parameter that you are varying.

Increment

Enter the value to increment the parameter by at each step.

of Variations

Enter the total number of parameter variations desired. Note: The starting value counts as a parameter variation. For example, if Start Value = 2, Increment = 1, and # of Variations = 4, the Vary A Parameter routine will use the following 4 parameter values: 2, 3, 4, and 5.

Descriptive Name

A default name is automatically assigned for the name of the parameter being varied (based on the selected **Component.Action**). This is the name that will be used in the report. The user can change this name if desired. If the user types in a custom **Component.Action**, the name for this parameter must be entered in this cell.

Created with the Standard Edition of HelpNDoc: [Easily create Help documents](#)

Train Genetic

Description

Trains the active NeuroSolutions breadboard while genetically optimizing the networks choice of inputs, step sizes, momentum values, and the number of processing elements in the hidden layer(s) (Note: Other parameters can also be optimized by setting them up manually within NeuroSolutions). The goal of the optimization is to find the parameter settings that result in the minimum error. If cross validation is used, the goal will be to minimize the cross validation error. Otherwise, the goal will be to minimize the training error. The best weights and parameter settings will automatically be saved during training and can be loaded into the NeuroSolutions breadboard by clicking the Load Best Weights menu item. However, this is not normally necessary since the best weights and parameters will automatically be loaded during testing (if the **Load Best** option is selected – the default) or during the application of the Production dataset.

To perform genetic training, first an initial population of networks is randomly created with each having a

different set of parameters. Each of these networks is then trained and evaluated (to determine its fitness) based on the minimum error it achieved. The characteristics of the good networks are then combined and mutated to create a new population of networks. Again, the networks in this population are evaluated and the characteristics of the best networks are passed along to the next generation of networks. This process is repeated until the Maximum Generations or Maximum Evolution Time is reached or the user stops the evolution.

After training, a report is automatically generated summarizing the results. The generated report contains the following information:

1. Plot of the best fitness achieved during each generation of the optimization. The best fitness is the overall minimum MSE (cross validation MSE if cross validation was used or the training MSE otherwise) among all of the networks within the corresponding generation.
2. Plot of the average fitness achieved during each generation of the optimization. The average fitness is the average of the minimum MSE (cross validation MSE if cross validation was used or the training MSE otherwise) taken across all of the networks within the corresponding generation.
3. Table summarizing the best fitness and the average fitness plots. For each of these plots, the minimum MSE (across all generations), the generation of this minimum, and the final MSE are displayed.

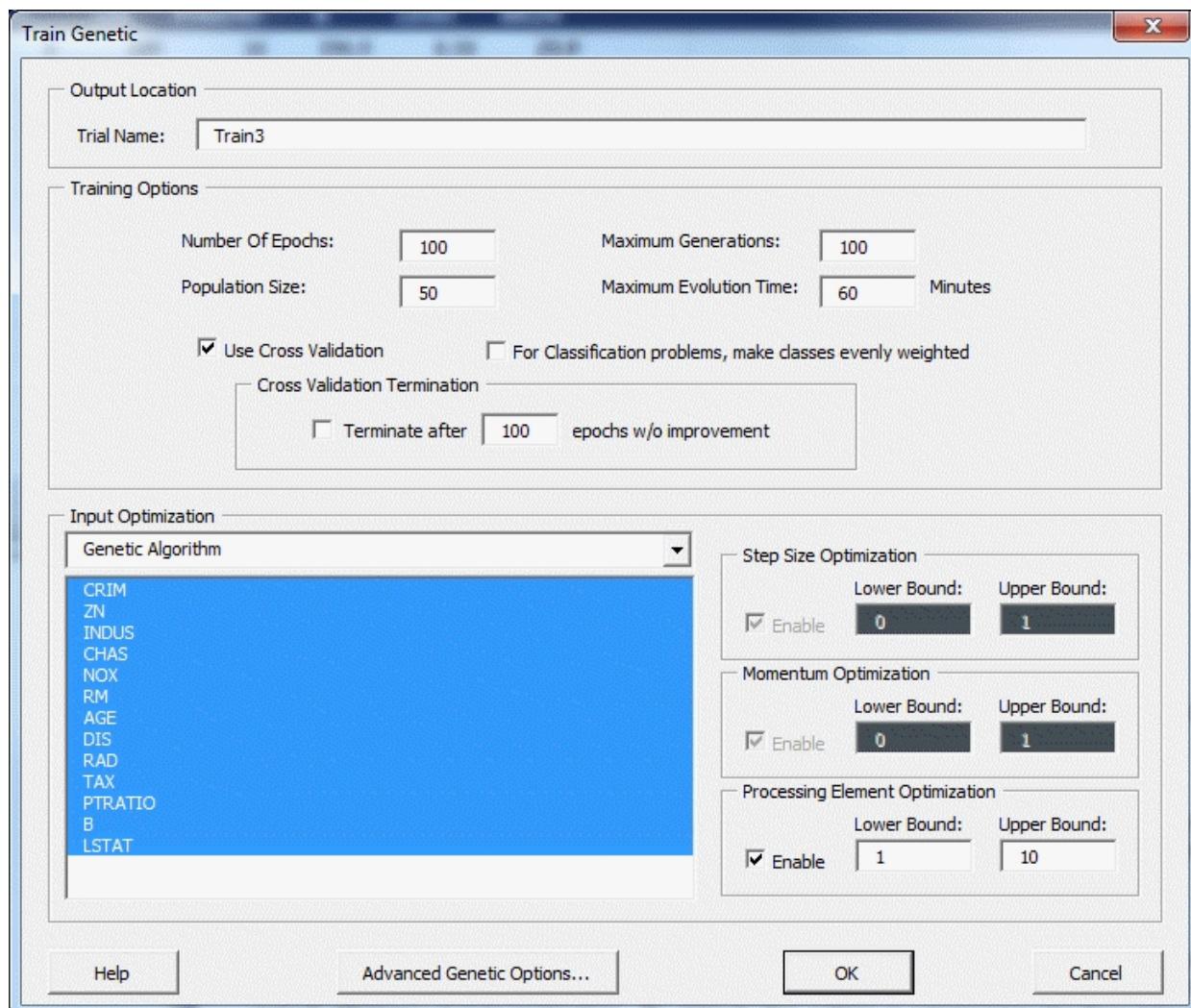
Note: If the user stops the genetic training before one generation has been completed, the report will not show any results.

The following data sheet is also created:

Data Sheet Name	Description
TrialName Fitness	Contains the best fitness and the average fitness for each generation. Also, the minimum of the best fitness and the minimum of the average fitness are given at the bottom of its corresponding column.

Note that the fitness measurements in this report are actually one half of the normalized MSE, which should match what is reported by the Average Cost probe of NeuroSolutions.

Dialog Box Options



Note: Models must be selected from a list, by holding down the Ctrl key while selecting the models to include. To select a block of models, select the first model in the block, then hold the Shift key and select the last model in the block.

Trial Name

Enter the name for this particular training trial run. If the entered name has already been used, you will be given the option to overwrite it.

Number of Epochs

Enter the total number of epochs to train the network for. An epoch is defined as one complete presentation of all of the data. By default, this setting is initialized to the Epochs / Run setting of the active NeuroSolutions breadboard.

Population Size

Enter the number of networks to create for each population. Increasing the population size increases the coverage of the parameter settings within a population, but if it is increased too much, it will reduce overall efficiency due to replication. By default, this setting is initialized to the Population Size setting of the active NeuroSolutions breadboard.

Maximum Generations

Enter the maximum number of generations to evolve for. A generation is defined as one complete evaluation of a population (the training and evaluation of all of the networks in the population). By default, this setting is initialized to the Maximum Generations setting of the active NeuroSolutions breadboard.

Maximum Evolution Time

Enter the maximum number of minutes to evolve for. By default, this setting is initialized to the Elapsed Time Termination setting of the active NeuroSolutions breadboard.

Use Cross Validation

Check this box if you want to use cross validation during training. Note: You will not be able to check this box if the active worksheet does not contain data tagged as Cross Validation.

Cross Validation Termination

Check this box to enable the termination of a training run if the cross validation error has not improved within the user-specified number of epochs. Note: You will not be able to check this box if the active worksheet does not contain data tagged as Cross Validation or if the ***Use Cross Validation*** option is unchecked.

Make Classes Evenly Weighted

Check this box to tell NeuroSolutions to evenly weight each class in a classification problem. This will enable the "Exemplar Weighting" feature of NeuroSolutions and will cause the gradients for each class to be weighted proportionately according to the number of samples of that class present in the training dataset. For classification problems in which one or more classes are disproportionately represented, using this feature will usually help the network to produce a more balanced model. Specifically, the underrepresented classes will be given more weight, which will usually improve their classification. For example, suppose you have a two-class problem where 99% of your samples are of the Class 1 type and 1% is of the Class 2 type. Without exemplar weighting, the network will likely take the path of least resistance and arrive at a model that classifies all of your data as Class 1. Overall, the network will be 99% accurate, but this is probably not what you want since it will be 0% correct for Class 2. With exemplar weighting, the gradients (used for updating the weights) for Class 2 samples will be weighted 99 times more than the gradients for Class 1 samples. This will usually result in a more correct classification percentage for Class 2 samples. For more information on "Exemplar Weighting", see the "Exemplar Weighting Inspector" help topic in the NeuroSolutions help. Note: This feature should only be enabled for classification problems.

Input Optimization

Check the inputs that you want to optimize (allow the inputs to be turned on or off). Uncheck the inputs that you want to always be on. The options for Input Optimization include:

1. **Genetic Algorithm:** This is the traditional method of progression for a genetic algorithm and has been proven to work well for a wide variety of problems. It tends to be a little slower than Steady State progression, but it tends to do a better job avoiding local minima.
2. **Greedy Search:** This is a type of input optimization that the evolution terminates immediately when adding a single input to the previous input set does not improve the fitness.
3. **Back Elimination:** This is a type of input optimization that the evolution terminates when removing a single input from the previous input collection leads to a worse fitness.
4. **Exhaustive:** This is a type of input optimization that tries all possible $2^N - 1$ combinations. Generally not recommended for more than 10 inputs.

Note for users of 64-bit Excel: Due to a limitation of the programming environment under 64-bit Excel, the check boxes shown in the above screen shot are not available. Instead, the inputs must be selected from a list, by holding down the *Ctrl* key while selecting the inputs to include. To select a block of inputs, select the first model in the block, then hold the *Shift* key and select the last model in the block.

Step Size Optimization

Check the Enable box if you want to genetically optimize the network's step sizes. This will enable the optimization of the step sizes for all Gradient components in NeuroSolutions. If you enable this optimization, you can use the Lower Bound and Upper Bound text boxes to specify the range that the optimized parameter(s) should stay within.

Momentum Optimization

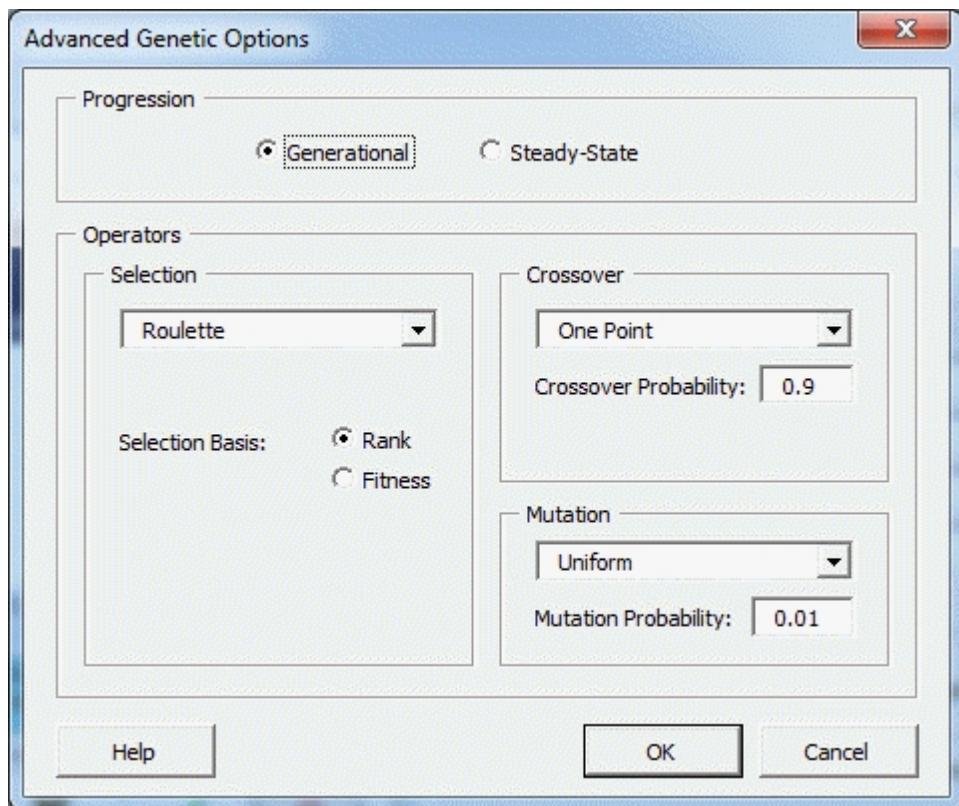
Check the Enable box if you want to genetically optimize the network's momentum values. This will enable the optimization of the momentum values for all Gradient components in NeuroSolutions that use momentum. If you enable this optimization, you can use the Lower Bound and Upper Bound text boxes to specify the range that the optimized parameter(s) should stay within.

Processing Element Optimization

Check the Enable box if you want to genetically optimize the number of processing elements in all of the network's hidden layers. This will enable the optimization of the **rows** parameter for each of the axons whose Component Name starts with "hidden" (see the Engine tab of the component's inspector). If you enable this optimization, you can use the Lower Bound and Upper Bound text boxes to specify the range that the optimized parameter(s) should stay within.

Advanced Genetic Options

Click this button to set up the advanced genetic options. The following dialog will be displayed:



For information on the advanced genetic settings, please download the [OptiGen Library Help](#) from our web site and read the chapter on Genetic Algorithm Theory.

Created with the Standard Edition of HelpNDoc: [Free Web Help generator](#)

Advanced Train Network Module



Train Network Menu Options

Menu Item	Description
Leave N Out	Trains the active NeuroSolutions breadboard multiple times leaving out different sections of data for each training run. This training procedure is very useful for testing the robustness of a model on small datasets.

Note for users of 64-bit Excel: Due to a limitation of the programming environment under 64-bit Excel, the progress bars shown during training will only have the numerical values and not the graphical bars.

Created with the Standard Edition of HelpNDoc: [Easy EPub and documentation editor](#)

Leave N Out

Description

The Leave-N-Out algorithm trains the network multiple times, each time omitting a different subset of the data and using that subset for testing. This enables you to use all of your data for training and all of your data for testing. The testing results are still out-of-sample since the rows being tested are not used to update the network weights during that particular training run.

To set up a Leave-N-Out training run, you must first tag a portion of your data set as Training and the remainder of the data set as Cross Validation. One thing to be aware of is that the cross validation set is not being used the same way as in the standard training routine. The cross validation set is actually the test set which is shifted between the training runs. The fewer rows you tag as cross validation, the more training runs will be required and the longer it will take to run. However, the testing results may be better with a smaller number of rows being tested since each run will have more training data.

Next, select Leave-N-Out from the Training menu and make any parameter changes if needed and click the OK button. You'll want to set the number of Epochs to be high enough for adequate training but low enough to prevent overtraining. You may want to try a standard Training run first to determine an appropriate setting (by looking at the epoch number of the minimum Cross Validation error).

The network will be trained for multiple runs, each time leaving out a different segment of the file for out-of-sample testing. The active NeuroSolutions breadboard is then trained one final time using all of the rows for training. The network output is collected and a report is then generated showing the training results and testing results. For the default settings, the testing report will contain the results from the entire data set. The contents of this report will vary based on whether the **Classification** or **Regression** report type was selected. The generated report contains the following information:

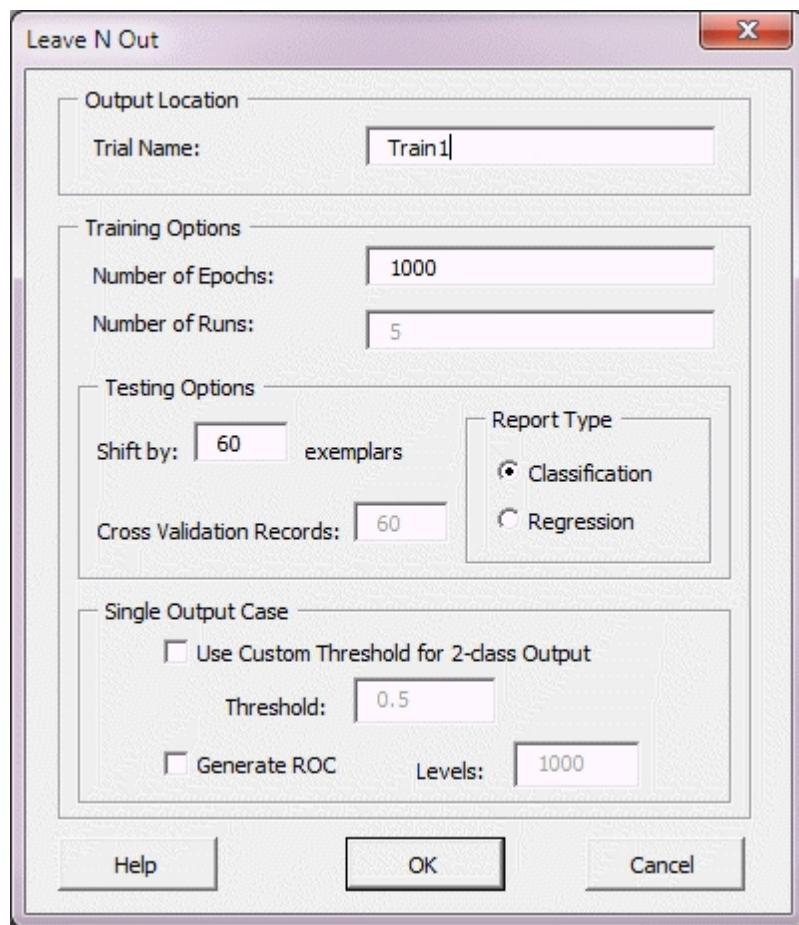
1. Plot of the training mean-squared error (MSE) versus Epochs.
2. Table showing the minimum training MSE, the epoch at which this minimum training MSE occurred, and the final training MSE.
3. **Regression selected:** Plot of the network output and the desired network output for each output. Each desired output will be plotted as a solid color and the corresponding network output will be a dashed line of the same color. For single-output problems a scatter plot will also be generated to show the network output and the desired output.
4. **Classification selected:** Confusion matrix showing the number of outputs classified as members of each class. For single-output problems there is also an option to generate an ROC table and plot.
5. Table reporting the mean-squared error (MSE), normalized mean-squared error (NMSE), mean absolute error (MAE), minimum absolute error, maximum absolute error, and correlation coefficient (r) for each output. Note that the MSE is measured using the original (i.e. non-normalized) desired outputs, which is different than the MSE reported in the Training portion of this report (which is one half of the normalized mean-squared error). If the **Classification** report type was selected, this table also includes the percent correct for each class.

The following data sheets are also created:

Data Sheet Name	Description
TrialName MSE	Contains the training MSE for each epoch in the run (learning curve). Also, the minimum training MSE is given at the bottom this column. If cross validation was used, the same results are also shown for the cross validation data set.
TrialName OutputData	Contains the output data for each run. (Each run contains the number of rows equal to your indicated shift value.)

Note that the MSE measurement for the training portion of this report is actually one half of the normalized MSE, which should match what is reported by the Average Cost probe of NeuroSolutions.

Dialog Box Options



Trial Name

Enter the name for this particular trial run. If the entered name has already been used, you will be given the option to overwrite it.

Number of Epochs

Enter the total number of epochs to train the network for. An epoch is defined as one complete presentation of all of the data. By default, this setting is initialized to the Epochs / Run setting of the active NeuroSolutions breadboard.

Cross Validation Records

The total number of rows tagged as cross validation. This number is set by the tagging only and cannot be changed in the form.

Shift by x exemplars

Enter the total number of rows to shift by for each run. Total number of runs will equal the Number of Cross Validation/shift value.

Report Type

1. **Classification:** Choose this option if you are trying to solve a classification problem. Note: If you choose this option, the desired output should be unary encoded. See the help for the [Column\(s\) As Symbol](#) sub-menu item of the Tag Data sub-menu.
2. **Regression:** Choose this option if the desired output(s) of the problem being solved are continuous values.

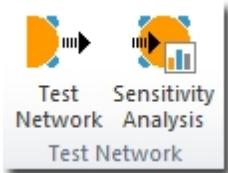
Single Output Case

1. **Use Custom Threshold for 2-class Output:** This field allows you to specify the threshold to use to differentiate between the two classes represented by the single output.
2. **Generate ROC:** The ROC Matrix consists of one row for each threshold generated. The thresholds are equally partitioned across the data range of the desired output.

3. **Levels:** This field specifies the number of thresholds generated within the data range.

Created with the Standard Edition of HelpNDoc: [Write eBooks for the Kindle](#)

Test Network Module



Test Network Menu Options

Menu Item	Description
<u>Test</u>	Tests the active NeuroSolutions breadboard on the chosen data set and creates a report of the results.
<u>Sensitivity About the Mean</u>	Performs sensitivity analysis on the chosen data set.

Created with the Standard Edition of HelpNDoc: [Produce Kindle eBooks easily](#)

Test

Description

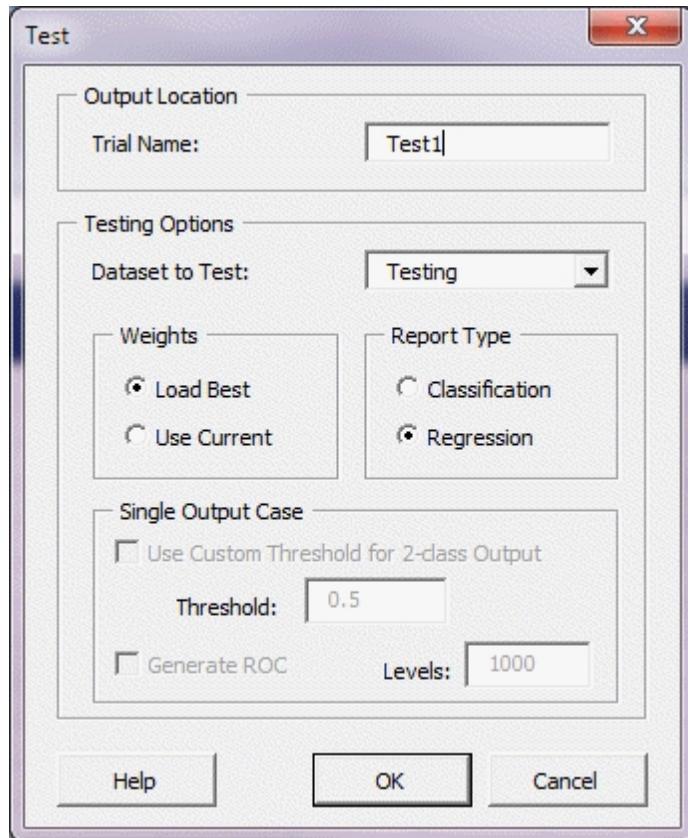
Allows the user to test a network on the chosen data set (Training, Cross Validation, or Testing). The user also has the option to use the current network weights or use the best network weights (Note: If a cross validation set is used during training, the best network weights are the ones that give the minimum cross validation error. Otherwise, the best network weights are the ones that give the minimum training error) saved during a training trial run (Note: if you choose to load the best network weights, the active breadboard must have the same topology as the one for which the best weights were created). During testing, the learning is turned off and the chosen data set is fed through the network. The network output is collected and a report is then generated showing the testing results. The content of this generated report varies based on whether the [Classification](#) or [Regression](#) report type was selected. If the desired output of the testing set is a single two-class column and the selected report type is Classification, the column is checked for the maximum of two classes and the report is generated without the necessity of translating the column. The generated report contains the following information:

- Regression selected:** Plot of the network output and the desired network output for each output. Each desired output will be plotted as a solid color and the corresponding network output will be a dashed line of the same color. For problems with a single output, a scatter plot will also be generated.
- Classification selected:** Confusion matrix showing the number of outputs classified as members of each class.
- Table reporting the mean-squared error (MSE), normalized mean-squared error (NMSE), mean absolute error (MAE), minimum absolute error, maximum absolute error, and correlation coefficient (r) for each output. Note that the MSE is measured using the original (i.e. non-normalized) desired outputs, which is different than the MSE reported in the Training reports (which is one half of the normalized mean-squared error). If the [Classification](#) report type was selected, this table also includes the percent correct for each class.

The following data sheet is also created:

Data Sheet Name	Description
TrialName IOData	The testing input(s) and desired output(s) used are copied to this sheet. After the testing data has been run through the network, the network output is also written to this worksheet.

Dialog Box Options



Trial Name

Enter the name for this particular testing trial run. If the entered name has already been used, you will be given the option to overwrite it.

Data Set to Test

Select the data set that you want to test. Only the data sets available for testing are shown.

Weights

1. **Load Best:** Choose this option to load the best network weights (saved during a training trial run) before the network is tested.
2. **Use Current:** Choose this option to test using the current network weights.

Report Type

1. **Classification:** Choose this option if you are trying to solve a classification problem. **Note:** If you choose this option, the desired output should be unary encoded. See the help for the [Column\(s\) As Symbol](#) sub-menu item of the Tag Data sub-menu.
2. **Regression:** Choose this option if the problem being solved is not a classification problem.

Single Output Case

1. **Use Custom Threshold for 2-class Output:** This field allows you to specify the threshold to use to differentiate between the two classes represented by the single output.
2. **Generate ROC:** The ROC Matrix consists of one row for each threshold generated. The thresholds are equally partitioned across the data range specified by the normalized data range of the input component. This field specifies the number of thresholds generated within the data range.

Created with the Standard Edition of HelpNDoc: [Generate EPub eBooks with ease](#)

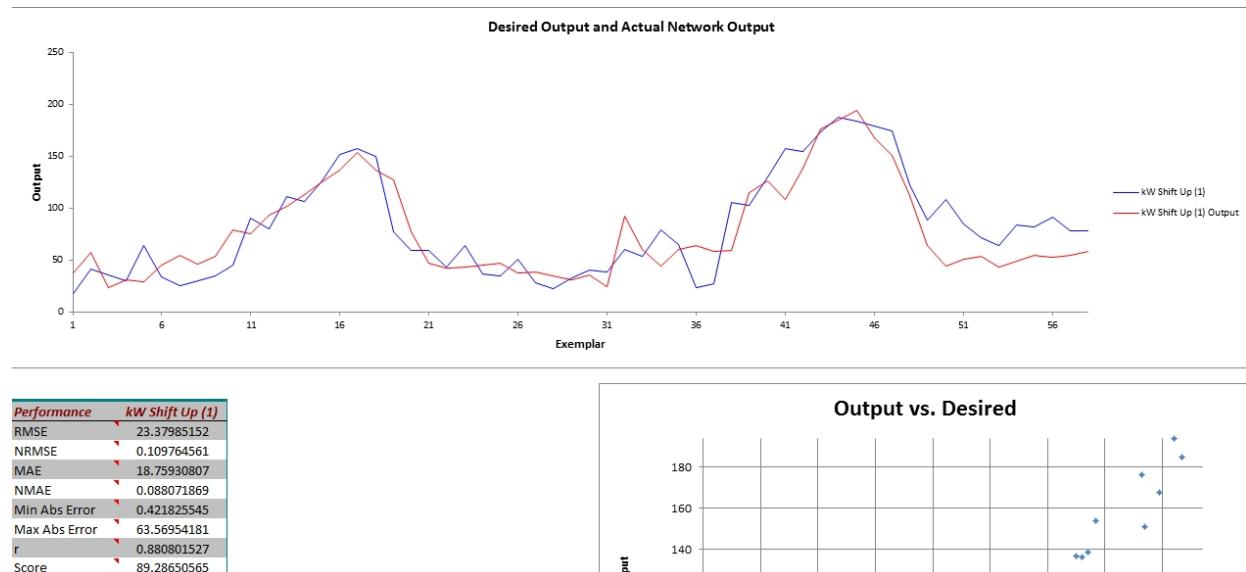
Test Reports

Description

The Test Reports are dynamic in that they are different depending on the options you chose in the [Test Network](#) dialog including Regression versus Classification and ROC.

Regression Report

This report shows a chart of the network output (dashed line) versus the desired output (solid line) for each desired output column. If there is only one output column, then the data will also be displayed as a scatter plot. The better the model, the more points will fall near the main diagonal line between the lower-left corner and the upper-right corner of the chart.



Performance Table

The Performance Table reports the root mean-squared error (RMSE), normalized root mean-squared error (NRMSE), mean absolute error (MAE), minimum absolute error, maximum absolute error, and correlation coefficient (r) for each output. The Score is based on a variety of statistics derived from the models performance such as: Area Under ROC (Classification Problems Only), Percent Correct (Classification Problems Only), Normalized Root Mean Squared Error, Normalized Mean Absolute Error, Average Percent Correct (Classification Problems Only).

Classification Reports

Threshold = 0.5			
Output / Desired		Benign(0)	Benign(1)
Benign(0)	149	8	
Benign(1)	9	176	

Performance	Benign(0)	Benign(1)
RMSE	0.226443139	0.209901447
NRMSE	0.226443139	0.209901447
MAE	0.063058404	0.048281565
Min Abs Error	4.03953E-55	1.31904E-09
Max Abs Error	1	0.999999999
r	0.903250302	0.903250302
Percent Correct	94.30379747	95.65217391

Confusion Matrix

A confusion matrix is a simple methodology for displaying the classification results of a network. The confusion

matrix is defined by labeling the predicted classification on the rows and the desired classifications on the columns. For each exemplar, a 1 is added to the cell entry defined by (desired classification, predicted classification). Since we want the predicted classification to be the same as the desired classification, the ideal situation is to have all the exemplars end up on the diagonal cells of the matrix (the diagonal that connects the upper-left corner to the lower right).

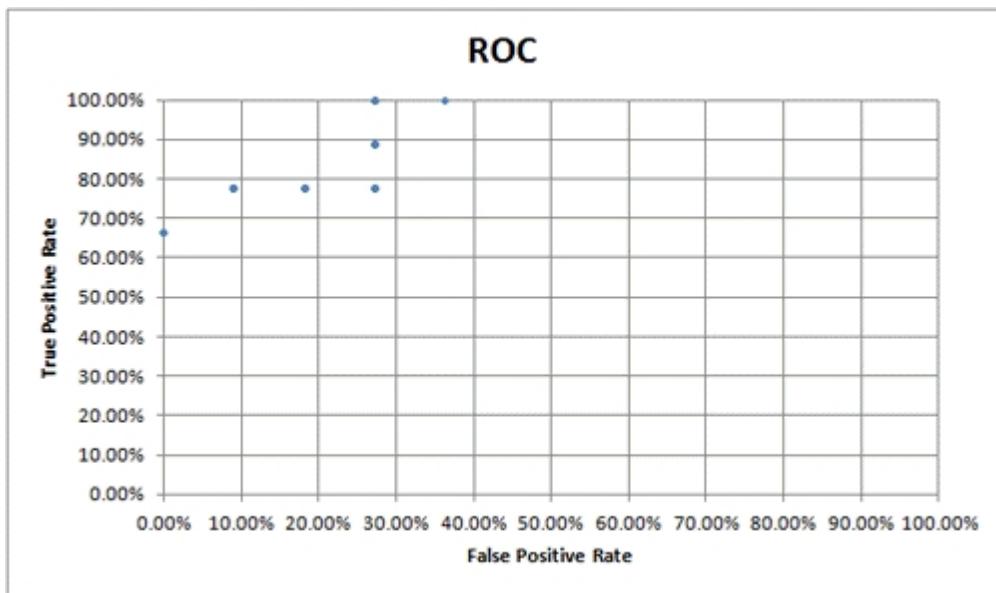
Performance Table

The Performance Table reports the root mean-squared error (RMSE), normalized root mean-squared error (NRMSE), mean absolute error (MAE), minimum absolute error, maximum absolute error, and correlation coefficient (r) for each output. This table also includes the percent correct for each class.

Classification Reports w/ ROC

ROC Curve

Receiver Operating Characteristic (ROC) matrices are used to show how changing the detection threshold affects detections versus false alarms. If the threshold is set too high then the system will miss too many detections. Conversely, if the threshold is set too low then there will be too many false alarms. Below is an example of an ROC matrix graphed as an ROC curve.



In NeuroSolutions for Excel, a ROC matrix is created through [Leave-N-Out Training](#) or the [Test Network](#) modules. The matrix contains the following: 1) True Positive Rate and 2) False Positive Rate.

ROC Matrix Definitions

1. **ROC Detection Threshold:** This field allows you to specify the threshold to use to differentiate between the two classes represented by the single output.
2. **Total Detections:** The number of samples predicted as 1.
3. **True Positive (TP):** If the outcome from a prediction is a 1 and the actual value is also a 1.
4. **False Positive (FP):** If the outcome from a prediction is a 1 and the actual values is a 0.
5. **True Negative (TN):** If the outcome from a prediction is a 0 and the actual value is also a 0.
6. **False Negative (FN):** If the outcome from a prediction is a 0 and the actual value is a 1.
7. **Detected as Positive ((TP+FP)/(TP+FP+TN+FN)):** The percentage of predicted as a 1 of all samples tested.
8. **False Positive Rate (FP/(FP+TN)):** The percentage of False Positives of those samples with an actual value of 0.
9. **True Positive Rate (TP/(TP+FN)):** The percentage of True Positives of those samples with an actual value of 1.
10. **False Discovery Rate (FP/(FP+TP)):** The percentage of False Positives of those samples predicted as a 1.
11. **Area Under the ROC Curve:** The sum of this column is displayed under the ROC chart, and it is a good

- measurement of the overall performance of the model. A measurement of 0.5 would be considered a poor model with no predictive power, while a measurement of 1.0 would be considered a perfect model.
12. **Positive Frequency:** The number of exemplars where the desired output was 1 and the network output was within the previous ROC Detection Threshold and the current ROC Detection Threshold. Note that this data is also shown in blue within the Histogram chart to the right of the Estimated Probability of Positive chart.
 13. **Negative Frequency:** The number of exemplars where the desired output was 0 and the network output was within the previous ROC Detection Threshold and the current ROC Detection Threshold. Note that this data is also shown in red within the Histogram chart to the right of the Estimated Probability of Positive chart.
 14. **Positive Cumulative %:** Of those exemplars where the desired output is 1, this is the percentage of where the network output was less than or equal to the ROC Detection Threshold.
 15. **Negative Cumulative %:** Of those exemplars where the desired output is 0, this is the percentage of where the network output was greater than or equal to the ROC Detection Threshold.
 16. **Estimated Probability of Positive:** The probability that a network output equaling the ROC Detection Threshold will have a desired output of 1. This data is also displayed in a chart to the right of the ROC chart.

		Actual Value				
Prediction Outcome	P'	P		N		Total P'
		True Negative	False Positive	False Negative	True Positive	
	N'					N'
		P		N		

Created with the Standard Edition of HelpNDoc: [Write eBooks for the Kindle](#)

Sensitivity About the Mean

Description

This testing process provides a measure of the relative importance among the inputs of the neural model and illustrates how the model output varies in response to variation of an input. By default the first input is varied between its mean +/- a user-defined number of standard deviations while all other inputs are fixed at their respective means. The network output is computed for a user-defined number of steps above and below the mean. This process is repeated for each input.

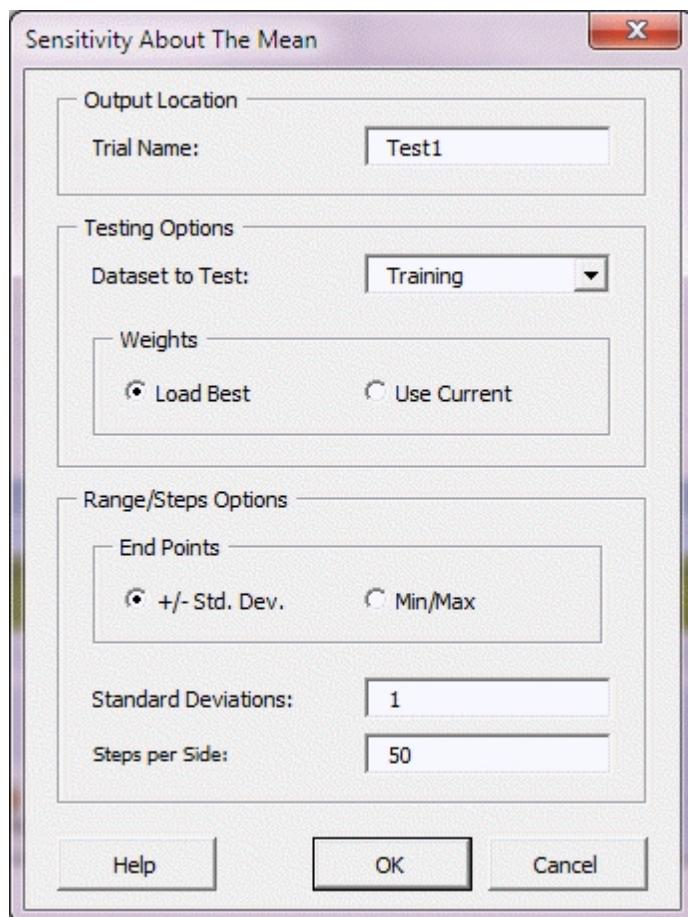
An alternate variation of this process is to vary the input of interest between its minimum value and its maximum value. This option is especially useful for binary inputs or inputs which have a non-Gaussian distribution.

A report is generated which summarizes the variation of each output with respect to the variation in each input. The generated report contains the following information:

1. A 3D Column plot of the table in 2.
2. Table reporting the standard deviation of each output for the input that was varied to create it. Each numerical value is simply the standard deviation of the network output for each of the X number of variations. For example, if the test is set to 50 steps per side, the network is tested 100 times for each input with the remaining inputs fixed at their mean values and while the tested input is varied between the min and the max (or +/- a number of standard deviations) for a total of 101 tests. The units of each sensitivity value (i.e., the standard deviation) would be whatever the output units are.
3. A plot is created for each input showing the network output(s) over the range of the varied input.

The following data sheet(s) is/are also created:

Data Sheet Name	Description
<i>TrialName</i> Output(#)	A new worksheet is created for each varied input. The varied input values are written to the worksheet along with the corresponding network output. The # sign shown at the left signifies the number of the varied input.

Dialog Box Options***Trial Name***

Enter the name for this particular testing trial run. If the entered name has already been used, you will be given the option to overwrite it.

Data Set to Test

Select the data set that you want to test. Only the data sets available for testing are shown.

Weights

1. **Load Best:** Choose this option to load the best network weights (saved during a training trial run) before the network is tested.
2. **Use Current:** Choose this option to test using the current network weights.

End Points

1. **+/- Std. Dev.:** Choose this option to vary each input around the column mean, plus or minus the number of standard deviations specified below.
2. **Min/Max:** Choose this option to vary each input between the minimum and maximum values of the column.

Standard Deviations

Enter the number of standard deviations to add and subtract from the mean of an input when calculating the range over which the input is varied.

Steps per Side

Enter the number of steps to use on each side of the mean. This is $\frac{1}{2}$ of the number of discrete values that will be used to calculate the output.

Created with the Standard Edition of HelpNDoc: [Easily create Help documents](#)

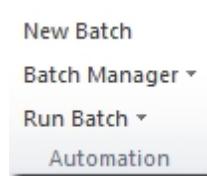
Apply Production Dataset

**Description**

Select this menu item to apply your neural model to the data tagged as [Production](#). This will feed the Production Input data into the neural model and produce the model output. The output will be written back to your data sheet at the location of the Production Desired cross-section. Note that by definition, Production data does not have a desired output since Production data is data that is used for the application of the neural model. Thus, the Production Desired cross-section will be empty until the Production dataset has been applied.

Created with the Standard Edition of HelpNDoc: [Full-featured Kindle eBooks generator](#)

Automation



Data Tools Menu Options

Menu Item	Description
New Batch	Option for creating a new batch process in Excel for NeuroSolutions.
Batch Manager	List of pre-built batches and user-defined batches.
Run Batch	Run pre-built batches and user-defined batchers.

Created with the Standard Edition of HelpNDoc: [Full-featured multi-format Help generator](#)

New Batch

Description

Select this menu item to create a new custom batch process. A dialog box will be displayed that will allow you to choose the type of custom batch that you want to create. There are 7 types of custom batches that are recognized by NeuroSolutions for Excel. These 7 types are listed below along with their required *Module Name Prefixes*:

Batch Type	Module Prefix
Preprocess Data	NSEPreprocess
Analyze Data	NSEAnalyze
Tag Data	NSETag
Create Network	NSECreateNet

Create Data Files	NSECreateFiles
Train Network	NSETrain
Test Network	NSETest

Select the type of batch you want to create, enter a name and a description for the batch, then click OK. A blank custom batch template module of the proper batch type will be copied to the active workbook. This batch template contains the following 3 procedures:

```

Function OnGetName() As String
OnGetName = "Your Batch Name"
End Function

Function OnGetDescription() As String
OnGetDescription = "Your Batch Description"
End Function

Sub OnRunBatchProcess()
'Enter the code for your custom <Batch Type> batch here
MsgBox "Custom batch process has not been implemented.", vbInformation + vbOKOnly
End Sub

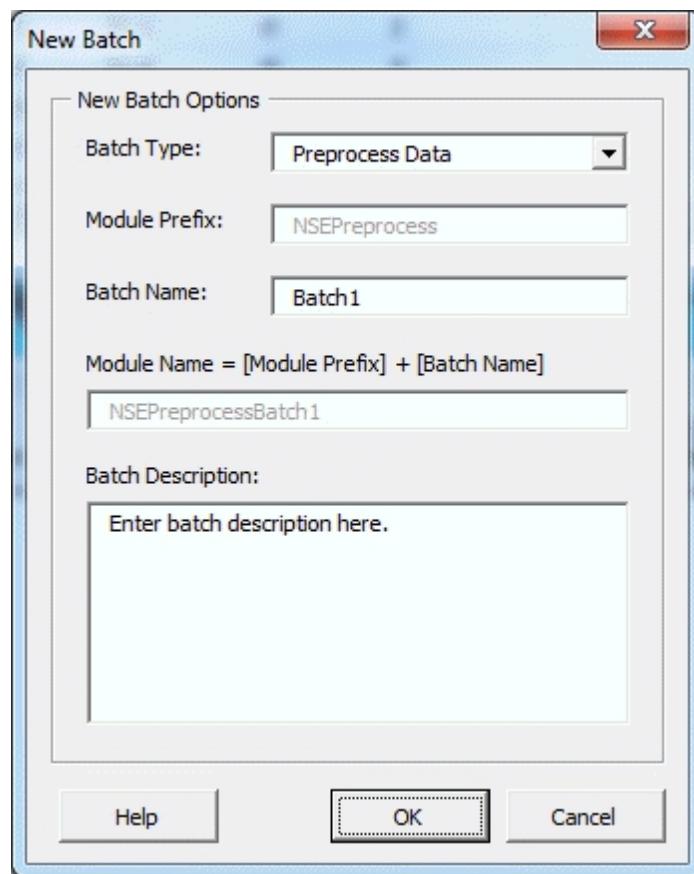
```

where <Batch Type> is one of the 7 batch types discussed above.

The OnGetName() and OnGetDescription() functions will contain the batch name and description that you entered into the New Batch dialog. This name and description will appear in the Run Batch dialog when you click the Run Batch menu item for the same batch type. To make changes to the batch name or description, simply edit these functions.

You will need to write the code to execute when this custom batch is run. This code should be inserted into the OnRunBatchProcess() subroutine. Note: The names of these three procedures cannot be modified.

Dialog Box Options



Batch Type

Choose the type of batch that you want to create.

Module Prefix

Shows the Module Prefix for the chosen batch type. This field is not editable.

Batch Name

Enter a name for the custom batch. This batch name is used to name the corresponding module (see Module Name below) and for display in the Run Batch dialog.

Module Name

This field shows the module name of the new batch. It is simply the concatenation of the Module Prefix with the Batch Name. This field is not editable.

Batch Description

Enter a description for the custom batch. The batch description is used for display in the Run Batch dialog.

Created with the Standard Edition of HelpNDoc: [Free EPub and documentation generator](#)

Batch Manager

Description

Selection of this menu item will display the batch manager dialog box for the chosen batch type. This dialog lists all of the batches found (for the chosen batch type) within the active workbook and within the code workbook (NSECodeWorkbook.xls). This dialog provides an easy interface for editing/deleting custom batches and for moving them between the active workbook and the code workbook. A batch process stored within the active workbook will only be available when that workbook is loaded and is the active workbook. If you want a custom batch process to always be available when NeuroSolutions for Excel is loaded, you should store it within the code workbook (simply use the Batch Manager to move it from the active workbook to the code workbook).

Note: The delete and move operations of the Batch Manager will be performed on all of the Excel objects whose name matches any of the following templates. This allows you to use Modules, Class Modules, Forms,

Worksheets, and Charts when creating a custom batch.

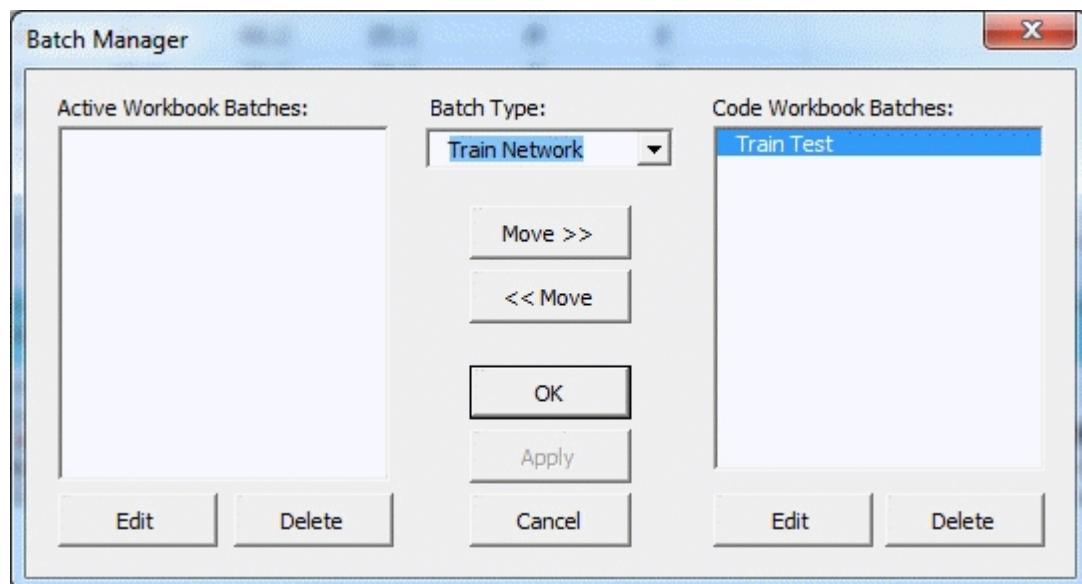
Modules:	[ModuleName]
Class Modules:	[ModuleName + "Class*"]
Forms:	[ModuleName + "Form*"] or [ModuleName + "Dialog*"]
Worksheets:	[ModuleName + "WSheet*"]
Charts:	[ModuleName + "Chart*"]

where

* = wildcard that can represent any number of characters.

+ = concatenation

Dialog Box Options



Active Workbook Batches

List of the batches available within the active workbook for the selected batch type.

Batch Type

Change the batch type to show the batches available for other batch types.

Code Workbook Batches

List of the batches available within the code workbook for the selected batch type.

Created with the Standard Edition of HelpNDoc: [Easily create EPub books](#)

Run Batch

Created with the Standard Edition of HelpNDoc: [Full-featured EPub generator](#)

Create Network...

Description

Selection of this menu item will display a dialog box listing all of the custom Create Network batches available within the Code Workbook (NSECodeWorkbook.xls) and within the active workbook. This dialog box displays the batch name and the batch description for each of the batches found. To run one of the Create Network batches displayed, click on its name then click the OK button.

A Create Network batch is identified by its Module name. All modules beginning with the prefix "NSECreateNet" are interpreted as Create Network batches and will be shown within the dialog. If a module's name starts with

this prefix, it **must** follow the custom batch protocol. For more information on this protocol and on creating custom batches see the help topic for the [New Batch](#) menu item.

Created with the Standard Edition of HelpNDoc: [Easily create HTML Help documents](#)

Train Network...

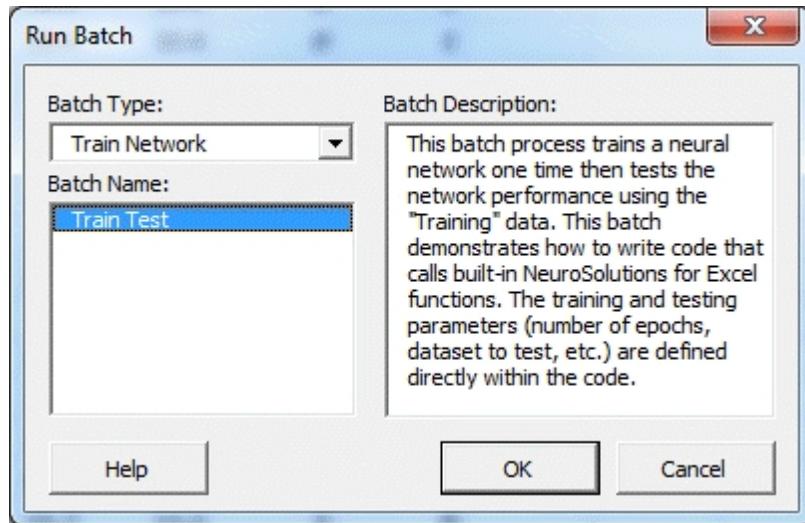
Description

Selection of this menu item will display a dialog box listing all of the custom Train Network batches available within the Code Workbook (NSECodeWorkbook.xls) and within the active workbook. This dialog box displays the batch name and the batch description for each of the batches found. To run one of the Train Network batches displayed, click on its name then click the OK button.

A Train Network batch is identified by its Module name. All modules beginning with the prefix "NSETrain" are interpreted as Train Network batches and will be shown within the dialog. If a module's name starts with this prefix, it **must** follow the custom batch protocol. For more information on this protocol and on creating custom batches see the help topic for the [New Batch](#) menu item.

The Code Workbook contains a custom Train Network batch named [Train Test](#) that is provided for demonstration purposes. Below is a picture of the Run Train Network Batch dialog and a short description of this batch.

Dialog Box Options



Batch Name

Lists all of the Train Network batches available. Select the one that you want to run then click **OK** to run it.

Batch Description

Displays the description for the currently selected Train Network batch.

Custom Train Network Batches

Batch Name	Description
Train Test	Trains the active NeuroSolutions breadboard one time then tests the performance using the Training data.

Created with the Standard Edition of HelpNDoc: [Full-featured EPub generator](#)

Test Network...

Description

Selection of this menu item will display a dialog box listing all of the custom Test Network batches available within the Code Workbook (NSECodeWorkbook.xls) and within the active workbook. This dialog box displays the batch name and the batch description for each of the batches found. To run one of the Test Network batches displayed, click on its name then click the OK button.

A Test Network batch is identified by its Module name. All modules beginning with the prefix "NSETest" are interpreted as Test Network batches and will be shown within the dialog. If a module's name starts with this prefix, it **must** follow the custom batch protocol. For more information on this protocol and on creating custom batches see the help topic for the [New Batch](#) menu item.

Created with the Standard Edition of HelpNDoc: [Free HTML Help documentation generator](#)

Create Data Files...

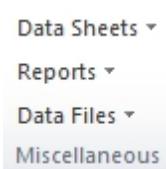
Description

Selection of this menu item will display a dialog box listing all of the custom Create Data Files batches available within the Code Workbook (NSECodeWorkbook.xls) and within the active workbook. This dialog box displays the batch name and the batch description for each of the batches found. To run one of the Create Data Files batches displayed, click on its name then click the OK button.

A Create Data Files batch is identified by its Module name. All modules beginning with the prefix "NSECreateFiles" are interpreted as Create Data Files batches and will be shown within the dialog. If a module's name starts with this prefix, it **must** follow the custom batch protocol. For more information on this protocol and on creating custom batches see the help topic for the [New Batch](#) menu item.

Created with the Standard Edition of HelpNDoc: [Free EPub producer](#)

Miscellaneous



Data Tools Menu Options

Menu Item	Description
Data Sheets	Access all visible and hidden data sheets in Excel.
Reports	Access to all reports in Excel generated by NeuroSolutions.
Data Files	Provides the ability to create data files manually for your tagged data.

Created with the Standard Edition of HelpNDoc: [Free help authoring tool](#)

Data Sheets

Created with the Standard Edition of HelpNDoc: [Single source CHM, PDF, DOC and HTML Help creation](#)

Goto Active Data Sheet

Description

Selection of this menu item will take you to the last data sheet for which a training or testing process was run. If no training or testing processes have been run, a dialog box will be displayed listing all of the tagged data sheets within the active workbook (see [Data Sheets](#)). If this dialog box is displayed, select the data sheet you

want to goto then click OK.

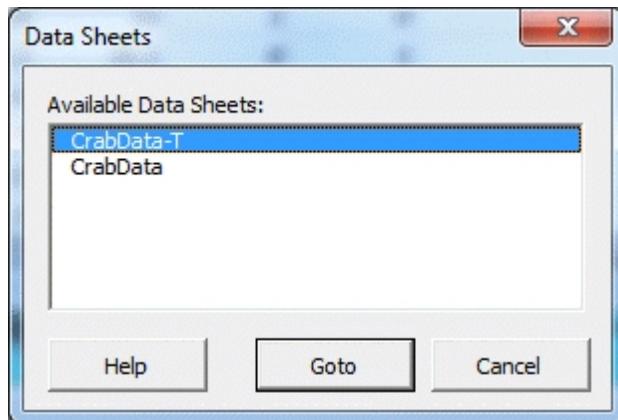
Created with the Standard Edition of HelpNDoc: Write EPub books for the iPad

Data Sheets

Description

Selection of this menu item will display a dialog box listing all of the tagged data sheets within the active workbook. Select the data sheet you want to goto then click OK.

Dialog Box Options



Available Data Sheets

Lists all of the tagged data sheets within the active workbook. Select the one you want to activate then click OK.

Created with the Standard Edition of HelpNDoc: [Full-featured Documentation generator](#)

Reports

Created with the Standard Edition of HelpNDoc: [Free Web Help generator](#)

Goto Active Report

Description

Selection of this menu item will take you to the last report generated by a training or testing process (the **active report**). The name of the active report is stored on the hidden project information sheet (NSVariables). If the stored active report name is blank or the NSVariables has not been created yet, the [Reports](#) dialog box is displayed listing all of the available reports within the active workbook. If this dialog box is displayed, select the report you want to view then click Goto. If there are no reports in the active workbook, a message is displayed conveying this information.

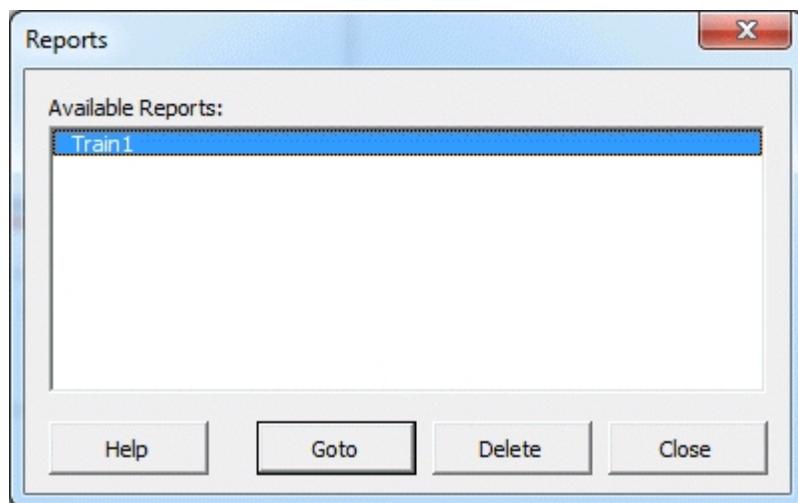
Created with the Standard Edition of HelpNDoc: [Full-featured Documentation generator](#)

Reports

Description

Selection of this menu item will display all of the reports available within the active workbook. Select the report you want to view then click Goto. Reports can also be deleted from this dialog. To delete a report, select the report that you want to delete then click Delete. This will permanently delete all of the sheets that were created for the chosen trial (i.e. the report and all of the associated data sheets).

Dialog Box Options

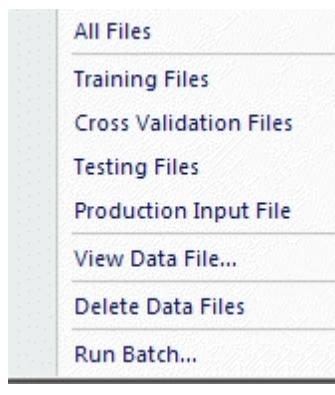


Available Reports

Lists all of the reports available within the active workbook. Select the one you want to goto or delete then click Goto or Delete, respectively.

Created with the Standard Edition of HelpNDoc: [Easy to use tool to create HTML Help files and Help web sites](#)

Data Files



Create Data Menu Options

Menu Item	Description
<u>All Files</u>	Creates data files for all tagged cross-sections within the active worksheet.
<u>Training Files</u>	Creates Training Input and Training Desired files from the correspondingly tagged cross-sections within the active worksheet.
<u>Cross Validation Files</u>	Creates Cross Validation Input and Cross Validation Desired files from the correspondingly tagged cross-sections within the active worksheet.
<u>Testing Files</u>	Creates Testing Input and Testing Desired files from the correspondingly tagged data cross-sections within the active worksheet.
<u>Production Input File</u>	Creates Production Input file from the correspondingly tagged data cross-section within the active worksheet.
<u>View Data File</u>	Allows you to view (in Notepad) a data file that was created for the active worksheet.
<u>Delete Data Files</u>	Deletes all of the files previously created for the active worksheet.
<u>Run Batch</u>	Displays a dialog box showing the available Create Data Files batches along with a description. From this dialog, you choose

which batch process to run.

Created with the Standard Edition of HelpNDoc: [Free iPhone documentation generator](#)

All Files

All Files

Description

When this menu item is selected, files for Training Input, Training Desired, Cross Validation Input, Cross Validation Desired, Testing Input, Testing Desired, and Production Input are created using the data on the active worksheet (provided that these cross-sections are tagged). The files are named and located according to the following naming convention:

Path = [Workbook Path]

Name = [Workbook Name (w/o extension)].[Worksheet Name].[Corresponding Suffix].asc

where [Corresponding Suffix] = “TrainingInput”, “TrainingDesired”, “CrossValidationInput”,
“CrossValidationDesired”, “TestingInput”, “TestingDesired”, or “ProductionInput”

Created with the Standard Edition of HelpNDoc: [Easy CHM and documentation editor](#)

Training Files

Training Files

Description

When this menu item is selected, files for Training Input and Training Desired are created using the data on the active worksheet (provided that these cross-sections are tagged). The files are named and located according to the following naming convention:

Path = [Workbook Path]

Name = [Workbook Name (w/o extension)].[Worksheet Name].[Corresponding Suffix].asc

where [Corresponding Suffix] = “TrainingInput” or “TrainingDesired”

Created with the Standard Edition of HelpNDoc: [Full-featured EBook editor](#)

Cross Validation Files

Cross Validation Files

Description

When this menu item is selected, files for Cross Validation Input and Cross Validation Desired are created using the data on the active worksheet (provided that these cross-sections are tagged). The files are named and located according to the following naming convention:

Path = [Workbook Path]

Name = [Workbook Name (w/o extension)].[Worksheet Name].[Corresponding Suffix].asc

where [Corresponding Suffix] = “CrossValidationInput” or “CrossValidationDesired”

Created with the Standard Edition of HelpNDoc: [Full-featured Kindle eBooks generator](#)

Testing Files

Testing Files

Description

When this menu item is selected, files for Testing Input and Testing Desired are created using the data on the active worksheet (provided that these cross-sections are tagged). The files are named and located according to the following naming convention:

Path = [Workbook Path]

Name = [Workbook Name (w/o extension)].[Worksheet Name].[Corresponding Suffix].asc

where [Corresponding Suffix] = "TestingInput" or "TestingDesired"

Created with the Standard Edition of HelpNDoc: [Write eBooks for the Kindle](#)

Production Input File

Production Input File

Description

When this menu item is selected, a file for the Production Input is created using the data on the active worksheet (provided that this cross-sections is tagged). The files are named and located according to the following naming convention:

Path = [Workbook Path]

Name = [Workbook Name (w/o extension)].[Worksheet Name].ProductionInput.asc

Created with the Standard Edition of HelpNDoc: [Easy CHM and documentation editor](#)

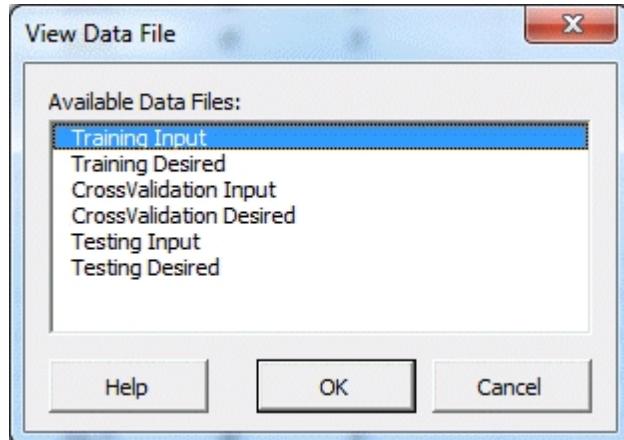
View Data File

View Data File

Description

When this menu item is selected, a dialog box is displayed showing all of the files that were previously created for the active worksheet. Choose the file you want to view then click OK. The selected file will then be opened within Notepad.

Dialog Box Options



Available Data Files

Choose the file that you want to view. When you click OK, the chosen file will be opened within Notepad.

Created with the Standard Edition of HelpNDoc: [Easily create Web Help sites](#)

Delete Data Files**Delete Data Files****Description**

Select this menu item to permanently delete all of the data files that were previously created for the active worksheet.

Created with the Standard Edition of HelpNDoc: [Free CHM Help documentation generator](#)

FAQ & Troubleshooting**Table of Contents**

- [**General Technology**](#)
 - [What is data preprocessing?](#)
 - [What is GPU \(Graphical Processing Unit\) processing?](#)
 - [What is Leave-N-Out?](#)
 - [What is the difference between NVIDIA CUDA and OpenCL?](#)
 - [What is the difference between Single vs. Double Precision?](#)
- [**NeuroSolutions for Excel**](#)
 - [Does my data need to be formatted any particular way?](#)
 - [What are the restrictions of the free trial?](#)
 - [What are the differences between NeuroSolutions & NeuroSolutions Pro?](#)
- [**Sales & Licensing Information**](#)
 - [Contacting NeuroDimension](#)
 - [Do you provide educational discounts?](#)
 - [Do you provide site licenses?](#)
 - [How to buy NeuroSolutions?](#)
 - [Related products & Modules](#)
 - [What is NeuroSolutions Infinity?](#)
- [**Additional Information**](#)
 - [System Requirements](#)

Created with the Standard Edition of HelpNDoc: [Easily create Help documents](#)

General Technology**Table of Contents**

- [**General Technology**](#)
 - [What is data preprocessing?](#)
 - [What is GPU \(Graphical Processing Unit\) processing?](#)
 - [What is Leave-N-Out?](#)
 - [What is the difference between NVIDIA CUDA and OpenCL?](#)
 - [What is the difference between Single vs. Double Precision?](#)

Created with the Standard Edition of HelpNDoc: [Full-featured Documentation generator](#)

What is data preprocessing?

Data preprocessing is an important step in the data mining process. The phrase "*garbage in, garbage out*" is particularly applicable to data mining and machine learning projects. Data-gathering methods are often loosely controlled, resulting in out-of-range values (e.g., Income: -100), impossible data combinations (e.g., Sex: Male, Pregnant: Yes), missing values, etc. Analyzing data that has not been carefully screened for such problems can produce misleading results. Thus, the representation and quality of data is first and foremost before running an analysis.

Created with the Standard Edition of HelpNDoc: [Full-featured Documentation generator](#)

What is GPU (Graphical Processing Unit) processing?

GPU-accelerated computing is the use of a graphics processing unit (GPU) together with a CPU to accelerate scientific, engineering, and enterprise applications.

 For help understanding GPU processing further, see [What is the difference between NVIDIA CUDA and OpenCL?](#)

Created with the Standard Edition of HelpNDoc: [Free iPhone documentation generator](#)

What is Leave-N-Out?

The Leave-N-Out algorithm trains the network multiple times, each time omitting a different subset of the data and using that subset for testing. This enables you to use all of your data for training and all of your data for testing. The testing results are still out-of-sample since the rows being tested are not used to update the network weights during that particular training run.

Created with the Standard Edition of HelpNDoc: [Easily create PDF Help documents](#)

What is the difference between NVIDIA CUDA and OpenCL?

CUDA™ is developed by NVIDIA and is proprietary for their product line of graphics cards. OpenCL is the open standard in parallel computing and supports not only AMD and NVIDIA graphics cards, but also AMD and Intel processors.

Created with the Standard Edition of HelpNDoc: [Single source CHM, PDF, DOC and HTML Help creation](#)

What is the difference between Single vs Double Precision?

Single and Double Precision refer to the floating-point format for Binary32 and Binary64 which allows for more decimal values (double precision).

What does that mean?

Accuracy. Double Precision supports more decimal values than Single Precision which makes the weights from the neural network more accurate.

For complete technical specification on Single vs. Double Precision please refer to this [Wikipedia Article](#).

Created with the Standard Edition of HelpNDoc: [Easily create Web Help sites](#)

NeuroSolutions for Excel

Table of Contents

- [NeuroSolutions for Excel](#)
 - [Does my data need to be formatted any particular way?](#)
 - [What are the restrictions of the free trial?](#)
 - [What are the differences between NeuroSolutions & NeuroSolutions Pro?](#)

Created with the Standard Edition of HelpNDoc: [Generate EPub eBooks with ease](#)

Does my data need to be formatted any particular way?

There are a few data requirements that must be met. These are as follows:

1. All columns of data must contain labels (see [Insert Column Labels](#)).
2. All labels for a particular column tag (“Input” or “Desired”) must be unique.
3. The labels must be located in the first row of the worksheet.
4. The data must be arranged in one contiguous rectangular block (with the exception of the Production dataset in which Desired data is not required).
5. There can be no empty cells within the tagged cross-sections (see [Clean Data](#)).
6. The data must start in the first column and the second row (directly after the labels).

	A	B	C	D	E	F	G	H
1	Specimen	Species	FrontalLip	RearWidth	Length	Width	Depth	Sex
2	142	0	20.6	14.4	42.8	46.5	19.6	Male
3	19	1	13.3	11.1	27.8	32.3	11.3	Male
4	169	0	16.7	14.3	32.3	37	14.7	Female
5	56	1	9.8	8.9	20.4	23.9	8.8	Female
6	164	0	15.6	14.1	31	34.5	13.8	Female
7	53	1	9.1	8.1	18.5	21.6	7.7	Female
8	113	0	14.1	10.5	29.1	31.6	13.1	Male
9	7	1	11.1	9.9	23.8	27.1	9.8	Male
10	73	1	12.8	12.2	27.9	31.9	11.5	Female
11	187	0	19.9	16.6	39.4	43.9	17.9	Female
12	173	0	17.5	14.7	33.3	37.6	14.6	Female
13	190	0	20.1	17.2	39.8	44.1	18.6	Female
14	188	0	19.9	17.9	40.1	46.4	17.9	Female
15	50	1	21.3	15.7	47.1	54.6	20	Male
16	34	1	16.4	13	35.7	41.8	15.2	Male
17	186	0	19.7	16.7	39.9	43.6	18.2	Female
18	72	1	12.8	12.2	26.7	31.1	11.1	Female
19	111	0	14	11.5	29.2	32.2	13.1	Male
20	127	0	17.4	12.8	36.1	39.5	16.2	Male
21	102	0	10.2	8.2	20.2	22.2	9	Male
22	28	1	15.7	12.6	35.8	40.3	14.5	Male
23	83	1	15	14.2	32.8	37.4	14	Female
24	138	0	18.8	13.8	39.2	43.3	17.9	Male
25	174	0	17.6	14	34	38.6	15.5	Female
26	122	0	15.4	11.1	30.2	33.6	13.5	Male
27	62	1	11.2	10	22.8	26.9	9.4	Female
28	175	0	18	14.9	34.7	39.5	15.7	Female
29	38	1	17.1	12.6	36.4	42	15.1	Male
30	52	1	9	8.5	19.3	22.7	7.7	Female
31	183	0	18.9	16.7	36.3	41.7	15.3	Female
32	59	1	10.4	9.7	21.7	25.4	8.3	Female

Example of valid data

Created with the Standard Edition of HelpNDoc: [Single source CHM, PDF, DOC and HTML Help creation](#)

What are the restrictions of the free trial?

There are two levels of NeuroSolutions (NeuroSolutions & NeuroSolutions Pro). If you have not purchased the Pro version, then you may encounter some restrictions as documented below. However, this will not keep you from experimenting with all of the features available within the Pro version.

Whenever a restricted operation is attempted, the program will ask if you would like to enter into evaluation mode. The evaluation mode allows you to use all of the features of the Pro version. However, once you have entered this mode, you will encounter the following restrictions:

Components	Restrictions
Axons & Synapses	The weights are not stored when the breadboard is saved
DataWriter	Cannot save the text from the display window to a file

	Cannot copy the text from the display window to the pasteboard
	Cannot redirect the probed data to a file
ImageViewer	Cannot save the probed image to a bitmap file
Criterion	Cannot automatically save the best network weights to a file
StaticControl & DynamicControl	Cannot save the network weights to a weights file
GeneticControl	The source code generation feature produces incomplete source code
All Neural Components	Cannot enable genetic optimization.
Macro Bars	Cannot create, compile, or debug a DLL
	Can only load DLLs created by NeuroDimension
	Cannot create a new Macro Bar or delete an existing one

Some of the tutorials within the on-line help require some advanced features available only in the higher-level versions. By allowing the program to switch to evaluation mode when prompted, you will be able to work through all of the tutorials within the text.

Note: If you have not [activated](#) your copy of NeuroSolutions within 14 days from the time of installation then the evaluation software will expire.

Created with the Standard Edition of HelpNDoc: [Full-featured multi-format Help generator](#)

What are the differences between NeuroSolutions & NeuroSolutions Pro?

Level-specific features:

Feature Level	NeuroSolutions ¹	NeuroSolutions Pro
Key Features		
Microsoft Excel Ribbon Control	x	x
Genetic Optimization		x
Advanced Input Optimization		x
Number of Pre-Built Neural Network Topologies	4	17
Number of Learning Paradigms	1	4
Maximum Number of Hidden Layers	2	Unlimited
Maximum Number of Inputs/Outputs/Neurons Per Layer	50	Unlimited
Number of Exemplars/Samples	Unlimited	Unlimited
Topologies		
Linear Regression	x	x
Multilayer Perceptron (MLP)	x	x
Generalized Feedforward	x	x
Probabilistic Neural Network (PNN)	x	x
Modular Network		x
Jordan/Elman Network		x
Self-Organizing Map (SOM)		x
Principal Component Analysis (PCA)		x
Radial Basis Function (RBF)		x
General Regression Neural Network (GRNN)		x
Neuro-Fuzzy Network (CANFIS)		x
Support Vector Machine (SVM)		x
Support Vector Machine Regression (SVM-R)		x
Hopfield Network		x
Time Delay Neural Network (TDNN)		x
Time-Lag Recurrent Network (TLRN)		x

	General Recurrent Network	x
Learning Paradigms		
Backpropagation	x	x
Unsupervised Learning		x
Recurrent Backpropagation		x
Backpropagation through Time (BPTT)		x
Optimization Techniques		
Genetic Optimization		x
Input Optimization: Greedy Search		x
Input Optimization: Back-Elimination		x
Input Projection Component		x
Gradient Descent Methods		
Step / Momentum	x	x
Delta Bar Delta	x	x
Quickprop	x	x
Resilient Backpropagation (RProp)	x	x
Conjugate Gradient	x	x
Levenberg-Marquardt	x	x
Advanced Features		
Weight Regularization	x	x
Exemplar Weighting	x	x
Sensitivity Analysis	x	x
Macros / OLE Automation	x	x
Iterative Prediction	x	x
User-Defined Neural Components (using DLLs)		x

1. The NeuroSolutions Student Edition is equivalent to the NeuroSolutions level.

Created with the Standard Edition of HelpNDoc: [Create HTML Help, DOC, PDF and print manuals from 1 single source](#)

Sales & License Information

Table of Contents

- [**Sales & Licensing Information**](#)

- [Contacting NeuroDimension](#)
- [Do you provide educational discounts?](#)
- [Do you provide site licenses?](#)
- [How to buy NeuroSolutions?](#)
- [Related products & Modules](#)
- [What is NeuroSolutions Infinity?](#)

Created with the Standard Edition of HelpNDoc: [Free Kindle producer](#)

Contacting NeuroDimension



NeuroDimension, Inc.
3701 NW 40th Terrace, Suite 1
Gainesville, FL 32606
www.nd.com

Sales and Information

Product Questions OR to Place Order	1-800-634-3327
Fax	352-377-9009
Email	info@nd.com
Calls Outside U.S.	352-377-5144

Technical Support

Technical Support	352-377-1542
Email	support@neurosolutions.com

Before contacting technical support, please attempt to answer any questions by first consulting the following resources:

- The NeuroSolutions Infinity Help
- The [NeuroSolutions Customer Forums](#)
- The on-line [Video Library](#)

Created with the Standard Edition of HelpNDoc: [Free Web Help generator](#)

Do you provide educational discounts?

Absolutely! NeuroDimension is a proud supporter of its ties with the educational community. We offer NeuroSolutions Student Edition for students and faculty members looking to get started with NeuroSolutions at a low price point. In addition, we offer academia bundles for individuals, research groups or large university labs with exceptional discounts. For more information on our academia bundles please refer to our website at: <http://www.neurosolutions.com/order/academia-bundles.html>

Created with the Standard Edition of HelpNDoc: [Easily create EBooks](#)

Do you provide site licenses?

Yes. We provide Corporate/Government Bundles as well as Academia Bundles with exceptional discounts. You can find out more information about these various offerings on our website at:

- [Corporate/Government Bundles](#)
- [Academia Bundles](#)

Created with the Standard Edition of HelpNDoc: [Free help authoring environment](#)

How to buy NeuroSolutions?

We offer a convenient checkout process on our website at: <http://www.neurosolutions.com/order/>

We accept most major credit cards including:

- American Express
- Discover Card
- MasterCard
- Visa

We also accept:

- Check
- PayPal
- Purchase Orders (for pre-approved customers)
- Wire Transfer (\$20 Bank Fee)

If you should have any questions about ordering please do not hesitate to [contact us](#).

Created with the Standard Edition of HelpNDoc: [Generate EPub eBooks with ease](#)

Related Products & Modules

Related Products

NeuroSolutions Infinity

NeuroSolutions Infinity neural network software offers reliable, scalable, distributed processing of large data across clusters of computers to create highly accurate predictive models for data mining and analysis. It is designed to scale up from a single computer to thousands of machines, each offering local computation. The easy-to-use interface allows you to set minimal conditions for preprocessing and neural network learning. Or, take the reins and set specific conditions to have total control over data preprocessing, training termination rules and neural network architectures. Perform sales forecasting, sports predictions, medical classification, and much more with NeuroSolutions Infinity.

NeuroSolutions

There are two levels of NeuroSolutions, all of which allow you to implement your own neural network models. NeuroSolutions features an Excel interface that boast exclusive features such as the Express Builder, Leave-N-Out Training and Vary a Parameter. It also comes with the classic NeuroSolutions wizards including the NeuralExpert for beginner users and the NeuralBuilder for more advanced users. NeuroSolutions supports traditional linear regression techniques as well as probabilistic and multi-layer perceptron neural networks.

The **NeuroSolutions Pro** level provides a three-times more neural network topologies for both static and dynamic pattern recognition applications, time-series prediction and process control problems. It also features powerful genetic optimization and search attribute methods for optimizing neural network parameters and input variables. NeuroSolutions Pro adds the capability to generate Custom Solution Wizard DLL's (Dynamic Link Libraries) that can be used in NeuroSolutions Infinity.

Modules

Agents (for NeuroSolutions Infinity)

Agents expand the computing capabilities of NeuroSolutions Infinity's distributed processing to additional computers for enhanced performance. By adding additional Agents, NeuroSolutions Infinity can perform more preprocessing and neural network training than any single machine can do alone!

NeuroSolutions Accelerator

NeuroSolutions Pro and NeuroSolutions Infinity users can now harness the massive processing power of multi-core CPU's and graphics cards (GPU's) from AMD, Intel and NVIDIA with the NeuroSolutions Accelerator module. NVIDIA CUDA and OpenCL enables training time improvements from hours to minutes when compared to traditional CPU's on neural networks using Levenberg-Marquardt.

Custom Solution Wizard

There are two levels of the Custom Solution Wizard, all of which allows you to take any neural network created with NeuroSolutions Pro and automatically generates and compiles a Windows-based DLL (Dynamic Link Library) for that network which can then be embedded into your own application. The **Custom Solution Wizard** allows you to create recall-only networks whereas the **Custom Solution Wizard Pro** level allows for the creation of both learning and recall networks.

C++ Code Generation for Windows

The C++ Code Generation for Windows allows users to generate ANSI C++ compatible code, all you to embed NeuroSolutions Pro algorithms into your own applications. It allows any simulation prototyped with NeuroSolutions Pro to be run on other platforms (e.g. faster computers or embedded real time systems).

C++ Source Code for All Platforms

The C++ Source Code for All Platforms allows you to compile the generated code using other Windows compilers or on other platforms such as Unix. Included with the license is the source code for the entire object library, enabling you to compile this library for your particular platform/compiler and link it with the generated code.

Created with the Standard Edition of HelpNDoc: [Easily create CHM Help documents](#)

What is NeuroSolutions Infinity?

 NeuroSolutions Infinity is leading edge distributed computing neural network software that streamlines the data mining process with advanced neural networks, artificial intelligence (AI) and cutting-edge input preprocessing. It creates highly accurate predictive models with an easy-to-use and intuitive interface that provides valuable insights that can be used to drive better decisions.

- Automatically mine your data for hidden relationships using distributed processing on large data sets from a single server to thousands of machines, each offering local computation.
- Harness the massive processing power of multi-core CPU's and graphics cards (GPU's) from AMD, Intel and NVIDIA through parallel computing with the NeuroSolutions Accelerator add-on.
- Powerful, automated data preprocessing to create highly accurate predictive neural network models with the click of the mouse.

Created with the Standard Edition of HelpNDoc: [Write EPub books for the iPad](#)

Additional Information

Table of Contents

- [Additional Information](#)
 - [System Requirements](#)

Created with the Standard Edition of HelpNDoc: [Free EPub producer](#)

System Requirements

Before installing NeuroSolutions, you should verify that the configuration of your system meets the following minimum specifications:

Operating System	Windows XP / Vista / 7 / 8 / 10
Memory	2GB RAM (4GB recommended)

Hard Drive	750MB free space
Video	800x600 (1024x768 recommended)
Internet Connection	Required for Activation

Excel Supported Versions

2007 (32/64)	2016 (32/64)
2010 (32/64)	
2013 (32/64)	

Created with the Standard Edition of HelpNDoc: [Create HTML Help, DOC, PDF and print manuals from 1 single source](#)
