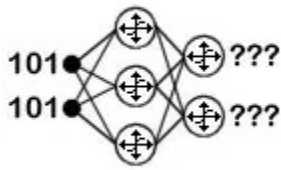


Testing Wizard Overview



After training a network, you will want to test the network performance on data that the network was not trained with. The Testing Wizard automates this procedure by providing an easy way to produce the network output for a new [data set](#) once the training phase has been completed.

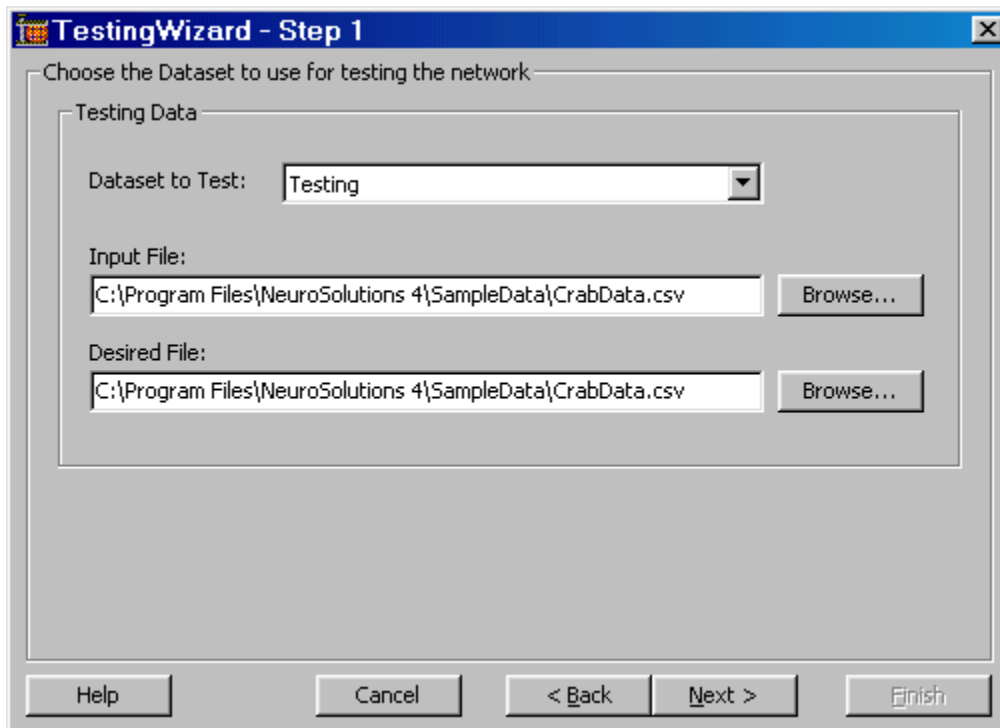
In order to test your network, the wizard must collect the following information:

1. The location of the file containing the input data to test your network
2. Optionally, the location of the file containing the desired output data to test your network. If this data exists, NeuroSolutions can not only compute the output data corresponding to your input, but also compute statistics such as the mean squared error between the network output and the desired network output.
3. How you want to view the data. The two current options are "display in a window" or "export to a file". If the desired output data is known, there is a checkbox to include the desired data in the display for easy comparison.

Important Notes:

- If the network you are testing already contains a testing or production [data set](#), then that data set will be the default data set and will appear automatically in the wizard
- All data files must be ASCII column formatted (ASCII with column headers).
- If there is no desired output, "Production" data set should be selected instead of the "Testing" data set.

Data Set Selection Panel



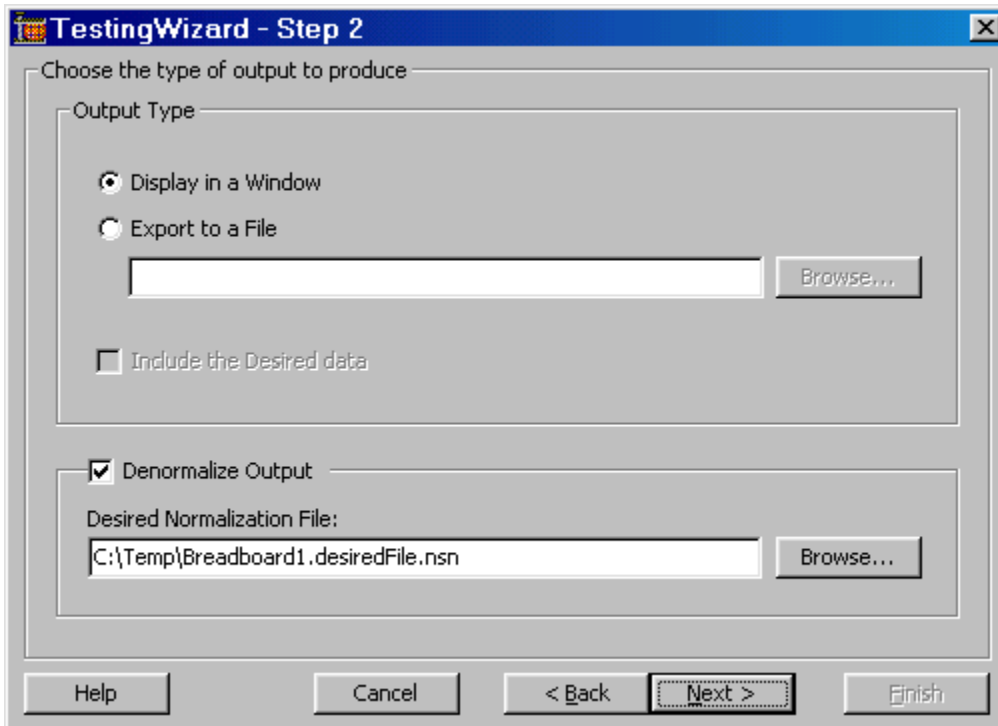
TestingWizard Data Set Selection Panel

This panel is used to specify the data to use for your test. The two most common data sets to use are Testing and Production. Testing is used if there is desired data that corresponds to the input data. Production is used if you do not know what the output is supposed to be – you want to neural network to tell you. You may also test the network using the Training or Cross Validation sets.

If the selected data set has already been defined, either manually or using one of the wizards, then the file path(s) will be filled in for you by default. If you have not yet defined the data set, or would like to override the existing file settings, click the Browse button(s) to define the input and/or desired files.

Note that the input file for the Testing/Production set must have the same column labels as the input file used for the training set. Likewise, the desired file of the Testing set must be of the same structure as the training desired file.

Output Production Panel



The image shows a Windows-style dialog box titled "TestingWizard - Step 2". The main heading inside is "Choose the type of output to produce". Below this is a section labeled "Output Type" containing two radio buttons: "Display in a Window" (which is selected) and "Export to a File". To the right of the "Export to a File" option is a text input field and a "Browse..." button. Below the "Output Type" section is a checkbox labeled "Include the Desired data". Another section below that has a checked checkbox labeled "Denormalize Output". Underneath this is a label "Desired Normalization File:" followed by a text input field containing the path "C:\\Temp\\Breadboard1.desiredFile.nsn" and another "Browse..." button. At the bottom of the dialog are five buttons: "Help", "Cancel", "< Back", "Next >", and "Finish". The "Next >" button is highlighted with a dashed border.

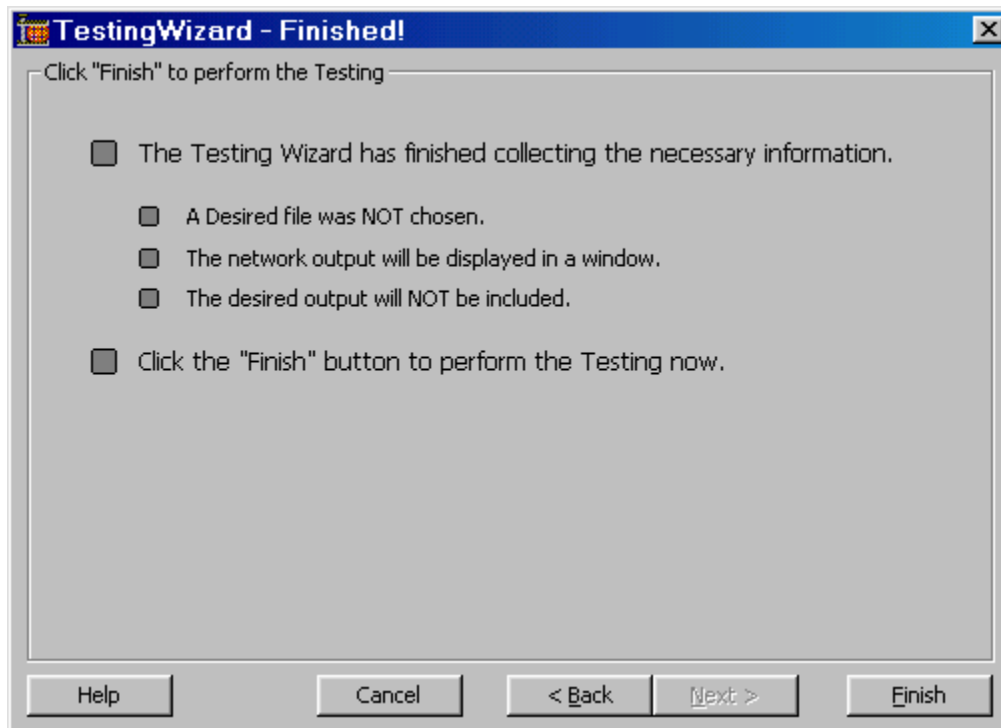
TestingWizard Output Production Panel

This panel is used to specify how the network output of the testing set should be produced. The DataWriter probe is the component used to extract the output data. You may either display the output data in a window or have the data written to an ASCII file. To do this, click the "Export to a file" radio button, click the "Browse" button and specify a file name and directory to write the output data to.

You may also include the desired data within the display/file in order to do a side-by-side comparison. If you only want to produce the output data, uncheck the "Include the Desired Data" box.

There may be times when you want to test a network that is a pure recall network – one that contains no error criterion or backprop components. In this case, you will have the option to denormalize the output data. Denormalization is the process of applying a scale and an offset to the data so that the values are in the same range as the desired output of the training set. The coefficients used for this operation are stored within a normalization file (*.nsn). If a normalization file exists for the desired File component, then its file path will be displayed by default.

Finish Panel



TestingWizard Finish Panel

This panel is used to summarize the selections that were made. Once you click the Finish button, the probes will be stamped on the breadboard and the testing data will be fed through the network.

Data Sets

NeuroSolutions provides a facility to divide your data into multiple data sets. This is typically used to ensure that the network is not overspecializing on the training data. The standard data sets are Training, Cross Validation and Testing (although others can be used).

Training data is the portion of the data used to actually train the network. This is normally the largest portion of your data.

Cross Validation data is used to intermittently validate the training. Periodically testing the network (no weight changes during cross validation) during training can help avoid overspecializing on the training data. Typically, the network is either stopped or its state is saved (save best weights) when the cross validation error begins to rise.

Testing data is used to further validate the results of a trained network. Although the network was not trained with the cross validation data, the training may have been stopped using it. Therefore, the cross validation data is not truly "out-of-sample". The testing data is data set aside to test the network after it has been trained and is truly "out-of-sample".

In NeuroSolutions, the Testing data set is used when you have both inputs and desired outputs for the Testing data. NeuroSolutions also has a Production data set which should be used when you have only testing input data but no desired output.

{ewl RoboEx32.dll, WinHelp2000, }

