



November 2004

Welcome to Synapse

In This Issue...

Introducing NeuralSight – New Product Announcement

A Model A Day is not Enough – a NeuralSight Case Study

Greetings!

With this issue of Synapse, we are proud to introduce **NeuralSight®**, the latest member of NeuralWare's family of advanced neural network model development products based on the NeuralWorks Predict® platform. We also give you a "behind the scenes" look at how we employed NeuralSight to help one of our major customers quickly build and test thousands of models in order to find the best subset to include in a new product being released for a new market. We hope this issue will inspire you to see new opportunities for taking full advantage of your investment in NeuralWare technology - and if you've not yet experienced the power and ease-of-use that NeuralWorks Predict offers, we hope that what you read here will convince you to come aboard. As always, we look forward to your comments - keep those cards and letters coming to newsletter@neuralware.com.

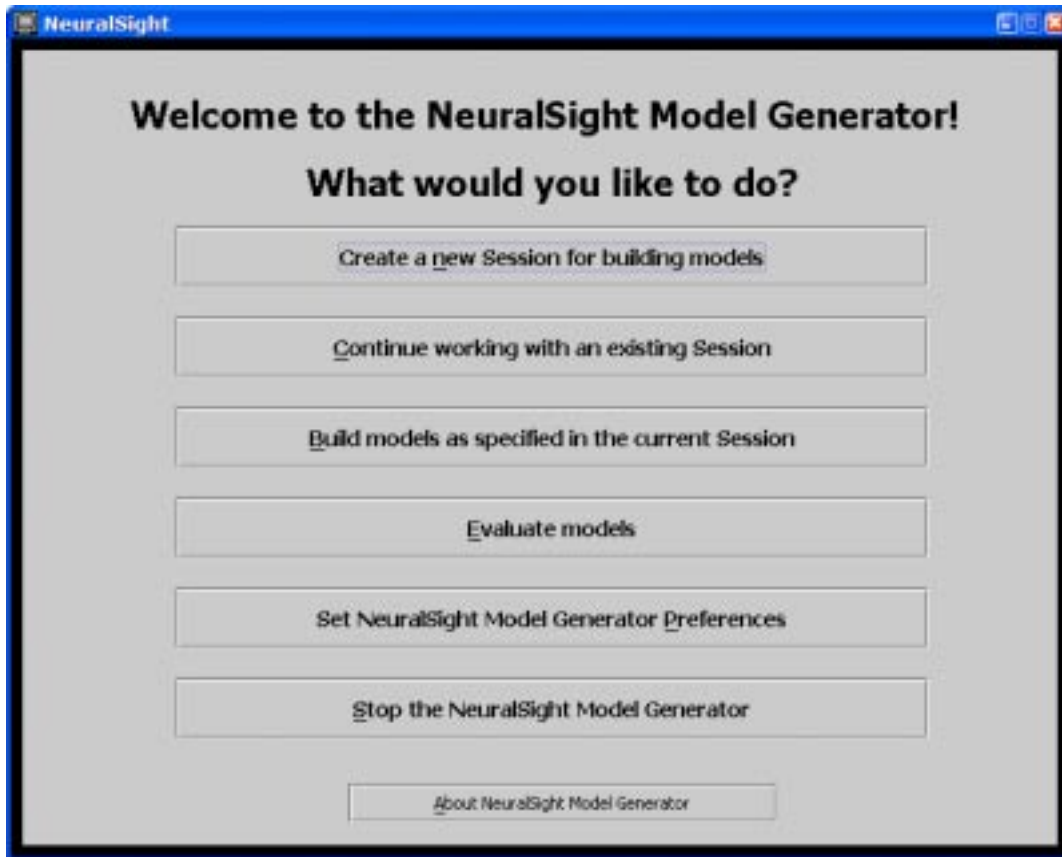
Introducing NeuralSight

If you currently are a user of NeuralWorks Predict, you already know how its comprehensive data transformation facilities and its powerful, genetic algorithm-based variable selection capability let you build robust neural networks for prediction, classification, and clustering problems right on your desktop. Now NeuralWare brings you NeuralSight - an advanced model development and evaluation environment that will help you greatly improve your model building productivity, so you have more time to devote to the business relationships and value propositions that your models will support (or, if you are an academic or corporate researcher, you'll have more time for preparing your publications!)

The NeuralSight Modeling Environment

If you currently are a user of NeuralWorks Predict, you already know how its comprehensive data transformation facilities and its powerful, genetic algorithm-based variable selection capability let you build robust neural networks for

prediction, classification, and clustering problems right on your desktop. Now NeuralWare brings you NeuralSight – an advanced model development and evaluation environment that will help you greatly improve your model building productivity, so you have more time to devote to the business relationships and value propositions that your models will support (or, if you are an academic or corporate researcher, you'll have more time for preparing your publications!)



NeuralSight extends the power of the Predict neural network engine by offering more flexible processing of large files with many fields, including filtering records based on any descriptive information contained in records, and by allowing you to specify model performance criteria which NeuralSight uses to automatically retain ("adopt") the best neural networks from a series of networks built by Predict during an unattended model building session (models which don't meet the criteria are discarded). You simply tell NeuralSight where to find the training data file, you specify performance requirements, you specify a small set of Predict options, and then you start the build process and let NeuralSight do the rest!

In conjunction with offering comprehensive training and testing dataset selection options, NeuralSight maintains an easy-to-access account of the particular dataset records used in training and evaluating models. This facility permits examining the characteristics of training and/or validation records that resulted in specific prediction or classification values.

In addition, during evaluation, all networks that have been adopted based on the selection criteria can be ranked by performance, to further aid in deciding which network or combination of networks are most suitable to deploy. A sophisticated Sensitivity Analysis facility can then be used to identify the network inputs that have the most effect on model performance.

The screen shots which follow illustrate how some of the key features of NeuralSight are implemented.

Configuring a Session

Logical model and folder naming conventions

Support for multiple types of models

Time or performance-based stopping criteria

Data-based (Classification) or User-defined categories

Session Configuration [Classification]

General | Filter by Fields | Select Data Records | Set Adoption Criteria | Set Build Options

Names
 Session Name: Example | Model Name: Model

Model Type
 Prediction
 Prediction Using Categories
 Classification

Model and Data Files
 Folder for Session: Select... C:\Modeling\Example\Sessions
 Folder for Adopted Models: Select... C:\Modeling\Example\Models
 Training Data File: Select... C:\Modeling\Example\Data\ClassificationData.csv
 Data Row Numbers from File: | Select...
 Text Excluded Records: | Select...

Other Files
 Create Predict Log File
 Keep Predict Command Files
 Keep Intermediate Training Data Files

Stopping Criteria
 Maximum Model Building Time: 4 [Hours]
 Maximum Number of Adopted Models: 4

Categories

Category	Upper Bound
1	0.1
2	0.2
3	0.3
4	0.4
5	0.5
6	0.6
7	0.7
8	0.8
9	0.9
10	

Number of Categories: 10
 Update Table

< Back | Next > | OK | Apply | Cancel

Models are organized in terms of Sessions (a collection of models which are all intended to model the same phenomenon). When a Session is configured, each tab in the illustration above is only accessible (enabled) after valid parameters have been entered on the prior tab. The **General** tab is the first tab that appears when Session configuration starts (in the illustration above, information for all tabs was entered, then the **General** tab was clicked to make it the active tab).

After a Session has been configured, model building commences. Models are trained and evaluated with respect to performance criteria specified by the analyst. The illustration below shows the information that is dynamically displayed while models are being trained and evaluated.

Building Models

While models are being trained, the form shown below is visible to permit monitoring progress (the form can be minimized but it cannot be closed while models are being built).

Model building can be performed on multiple threads if multiple-processor hardware is available

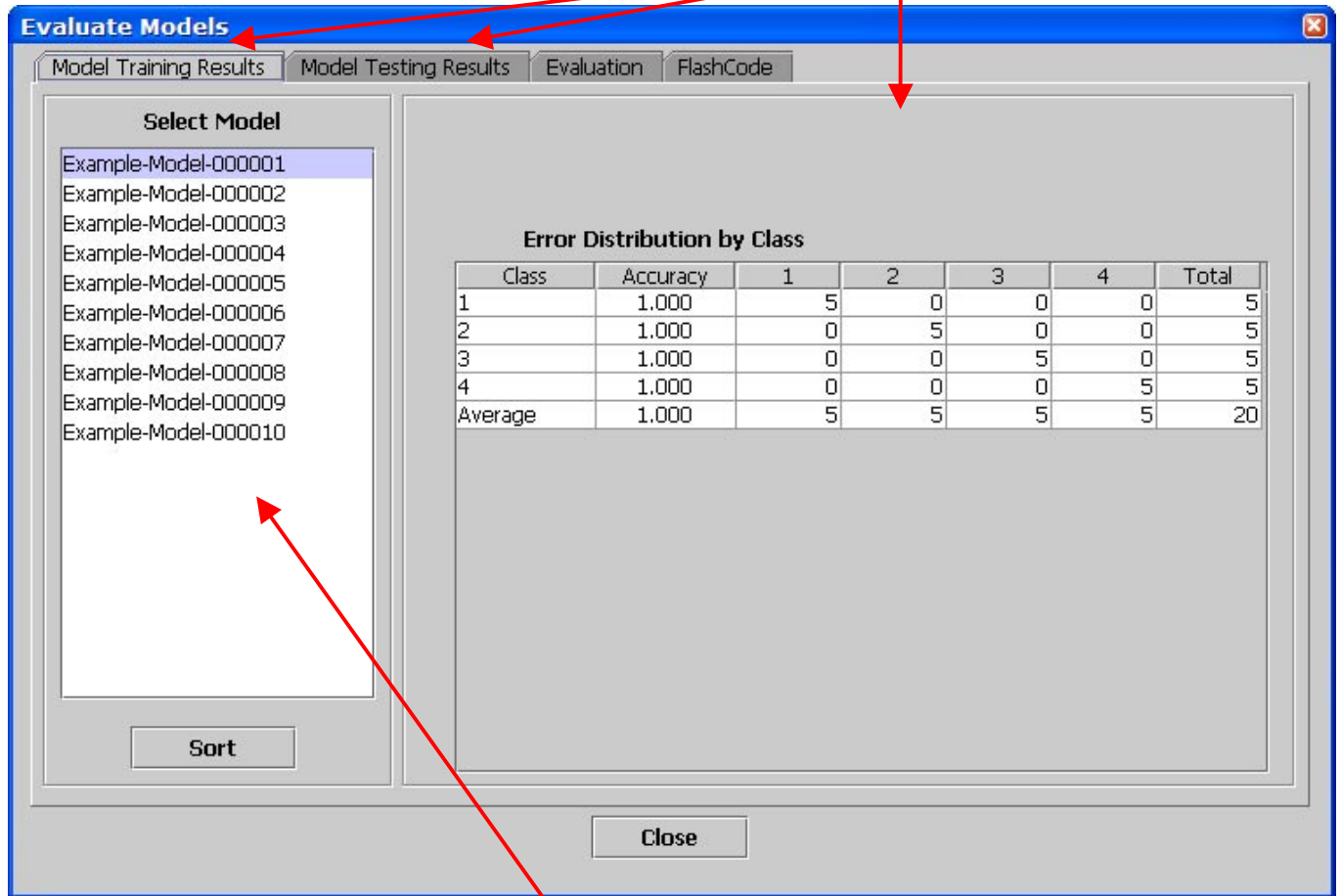
The screenshot shows a software window titled "Building Models [Example-Model]". It is divided into three main sections: "Session", "Training", and "Progress".

- Session:** Contains a "Time Model Building Started" field with the value "November 26, 2004 11:19:25 AM EST". Below it is a text input field for "Approximate Number of Models to Build" containing "216", with a red arrow pointing to it from the text box above. To the right of this field is the text "[Advanced Options Selected]". Below that is a "Maximum Number of Threads to Use" field containing "1". At the bottom of this section is a checkbox for "Sound Alert When Finished" which is checked.
- Training:** Features a "Start Training" button on the left and a "Stop Selected" button on the right. A central text area displays "000005 ST 11:19:28". Below the "Stop Selected" button is a "Stop All" button. A red arrow points from the text box below to this "Stop All" button.
- Progress:** Displays several metrics: "Number of Models Built" (4), "Number of Models Adopted" (4), "Elapsed Time" (00:00:03), and "Limits" (10). There are also fields for "Limits" (10) and "Time Model Building Ended" (empty). A "Close" button is located at the bottom center of the window.

Model building can be interrupted and re-started

Evaluating Models

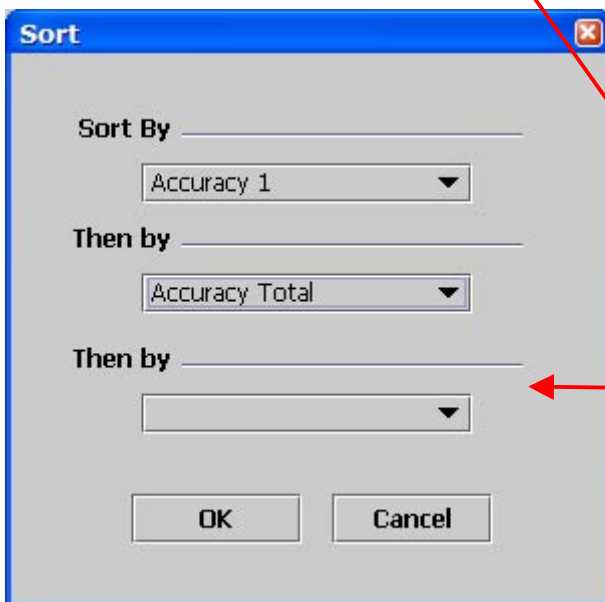
When training is finished a complete error distribution for training and testing datasets is created



The 'Evaluate Models' dialog box has four tabs: 'Model Training Results', 'Model Testing Results', 'Evaluation', and 'FlashCode'. The 'Evaluation' tab is active. On the left, under 'Select Model', there is a list of ten model IDs from 'Example-Model-000001' to 'Example-Model-000010'. Below the list is a 'Sort' button. On the right, the 'Error Distribution by Class' table is displayed:

Class	Accuracy	1	2	3	4	Total
1	1.000	5	0	0	0	5
2	1.000	0	5	0	0	5
3	1.000	0	0	5	0	5
4	1.000	0	0	0	5	5
Average	1.000	5	5	5	5	20

At the bottom of the dialog is a 'Close' button. Red arrows point from the text box above to the 'Evaluate Models' title bar, the 'Error Distribution by Class' table, and the 'Sort' button.



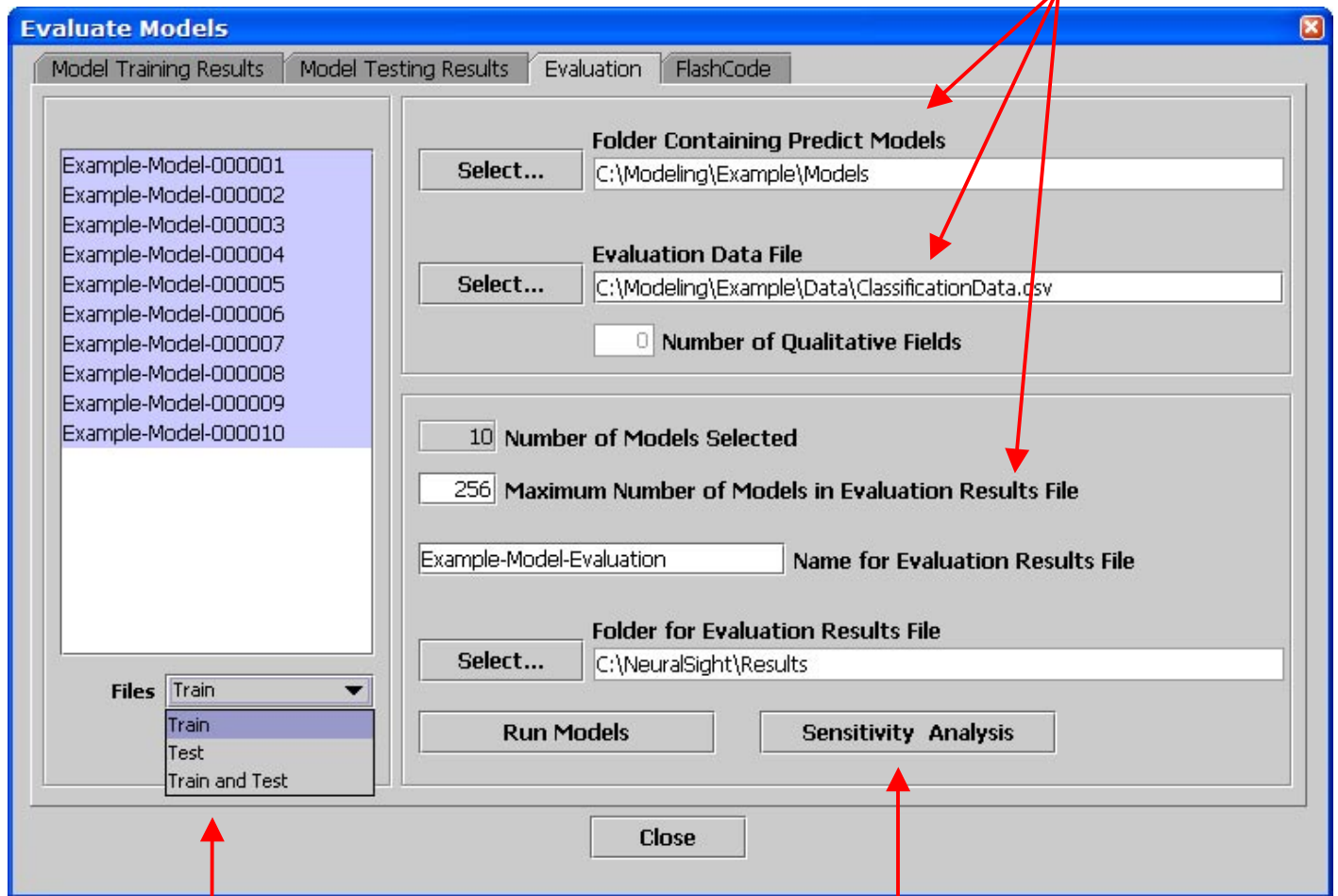
The 'Sort' dialog box contains three sorting criteria fields:

- 'Sort By' with a dropdown menu set to 'Accuracy 1'.
- 'Then by' with a dropdown menu set to 'Accuracy Total'.
- 'Then by' with an empty dropdown menu.

At the bottom are 'OK' and 'Cancel' buttons. A red arrow points from the text box on the right to the empty 'Then by' dropdown menu.

The list of models can be sorted by performance based on multiple criteria

The Evaluation tab provides additional facilities for evaluating model performance – models can be executed with other datasets and results written to Excel compatible .csv files

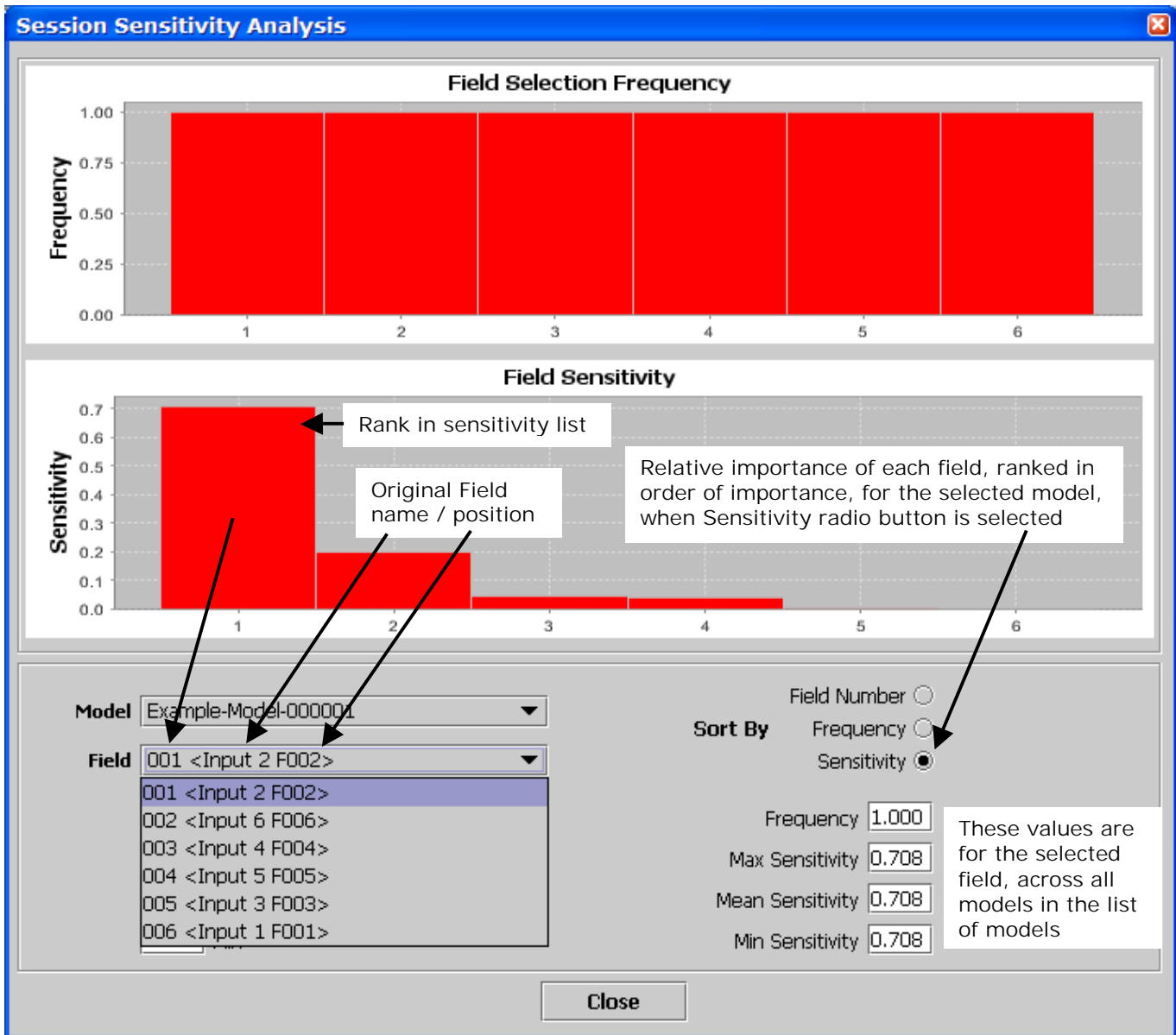


Models can be evaluated with respect to their performance on Train, Test, or both Train and Test datasets

Sensitivity Analysis is a powerful technique to identify which input fields have the most influence on model performance (refer to picture on next page)

Sensitivity Analysis

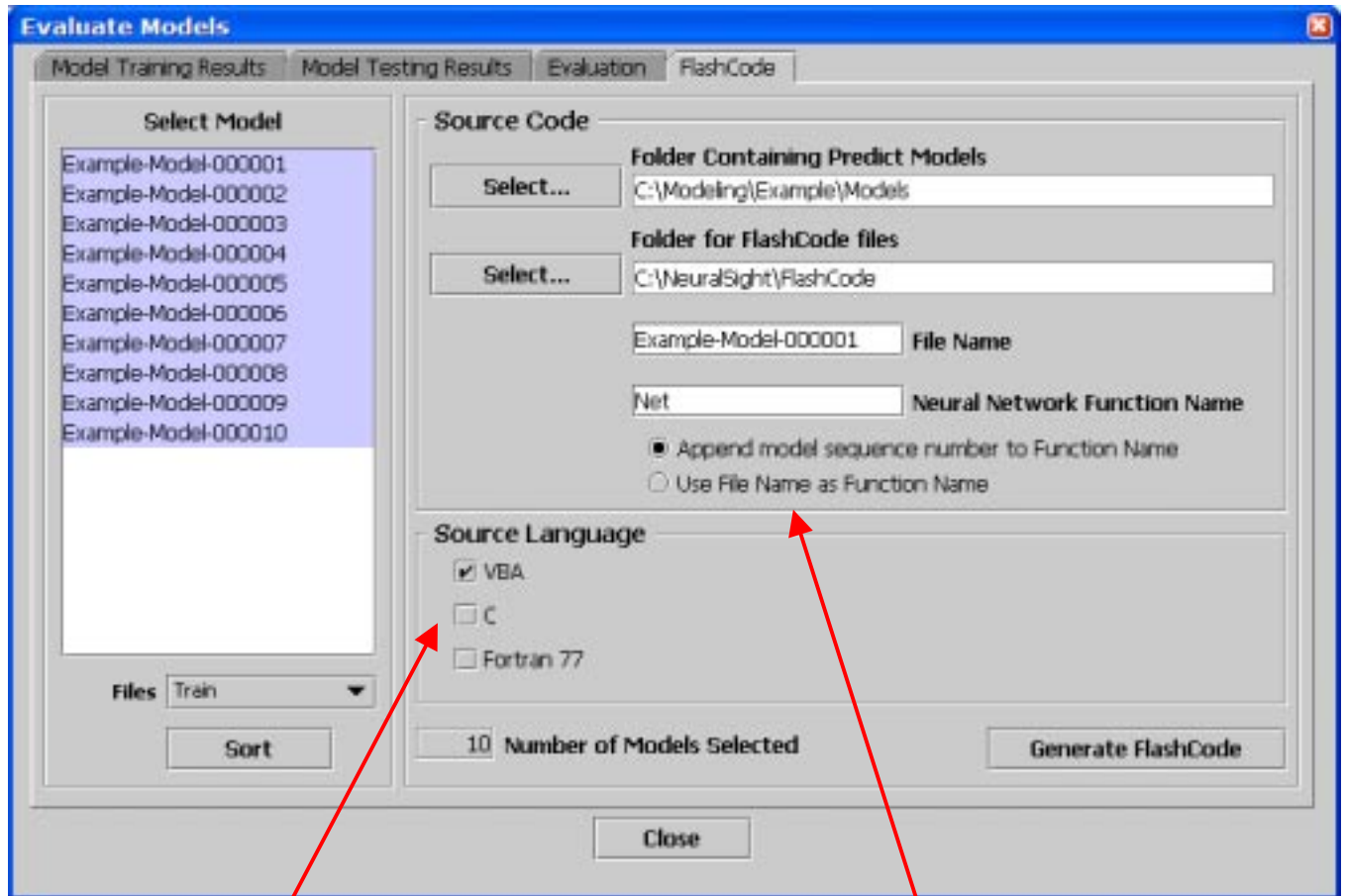
Sensitivity Analysis charts show the frequency that input fields were used in the models trained in a Session, and the impact of input fields on model performance for the selected model



At the same time that Sensitivity Analysis charts are generated, a .csv file is created that maps sensitivity values to variables used in each model, so that additional comparisons of model behavior can be easily performed in Excel.

Generating Flash Code

FlashCode™ is NeuralWare's powerful facility for generating a source code representation of a trained neural network so that it can be compiled and embedded in an application designed to execute in microprocessor-based equipment that may not have a standard operating system.



Multiple programming languages are supported

Application-specific file and function naming conventions are supported

This brief introduction has only touched on the many capabilities that NeuralSight brings to your empirical model development tasks!

Please contact us to learn how you can begin using NeuralSight today, and take advantage of special introductory pricing. Contact John Wavle at +1 412.278.6292, or john.wavle@neuralware.com.

A Model a Day is not Enough!

In one of the first applications of NeuralSight to a real world development challenge, NeuralWare worked closely for one month with scientists from DuPont Canada to build and test almost 7,000 models for a new grain quality assessment system. The application is the DuPont(TM) Acurum(TM) system that DuPont has developed with its strategic partners to enter a new market for grading wheat.

The Acurum system (www.acurum.com) is the world's most advanced grain quality assessment instrument today. It consists of an imaging subsystem that captures a visual image of grain, and an ensemble of neural networks to assess the quality of the grain, based on patterns of image attributes. Of course, it is a huge challenge to create sufficient labeled training datasets, and then to build sufficient models such that statistically sound judgments can be made about the best models. This is especially true given the many possible variations in the large number of features that can be extracted from images.

A model development process was established to ensure that DuPont Canada scientists primarily spent time identifying image features and harvest and commodity characteristics. As a result, the scientists could provide instructions to specify target (label) values for various grain conditions. NeuralWare personnel then employed [NeuralSight](#) to independently build and validate models for each of the assigned classification scenarios.

Since empirical modeling by its nature is an iterative process that ideally incorporates all available training data, a critical element of the development process was instituting a consistent data file naming convention that leveraged NeuralSight's built-in facility to use the concept of a model building *Session* to logically organize collections of models that are candidates for the same modeling function.

At the conclusion of the model development phase, over 5 million training data records had been processed, and almost 7,000 models had been trained and evaluated, before the final ensemble was selected and validated using independent real world data. All of the model building/evaluation occurred during a period of approximately one calendar month, using a mix of 4 standard Pentium® based computers operating at speeds from 1.4 to 3.06 GHz.

Future newsletters will offer other hints and tips for undertaking large-scale model development projects. But you don't need to wait for future newsletters. [NeuralWare](#) offers in-depth model development and deployment consulting that can help you get your project started on the right track – and finished with high quality models – as soon as you're ready! Contact jack.copper@neuralware.com or bob.everly@neuralware.com to discuss your next model development project.

And don't forget to check out NeuralWare's training schedule for the remainder of

2004 and into 2005. Come learn from NeuralWare experts, as well as your peers, how to take advantage of NeuralWare's ever expanding technology portfolio.

NeuralWare and NeuralSight are registered trademarks of NeuralWare.

DuPont™ is a trademark of E. I. du Pont de Nemours and Company. DuPont Canada is a licensee.

Acurum™ is a trademark of E. I. du Pont Canada Company

Special Offers on NeuralWare Products

If you already have a license for NeuralWorks Predict 3.1x, you are eligible for special introductory pricing on NeuralSight.

If you are not yet a Predict user, through November 30, commercial customers can take 15% off the standard price for NeuralWorks Predict and 1 year of TAP, PLUS become eligible for introductory pricing on NeuralSight.

For more pricing and other details concerning these special offers, please contact John Wavle, VP for Sales, at john.wavle@neuralware.com or +1 412 278 6292.