



NEURALWARE

Synapse
September 2002

Welcome to the September issue of Synapse, a quarterly electronic newsletter to customers and associates interested in neural network technology.

Predict 3.0 is now shipping

What's new in 3.0? While retaining its power and flexibility to quickly construct robust neural networks without having in-depth knowledge of the underlying technology, Predict 3.0 represents a complete upgrade of its intuitive user interface. All command and parameter views follow a property page paradigm, with related features and functions collected in easily identified pages.

Also new in Predict 3.0 is a Kohonen Self Organizing Map (SOM) that allows development of sophisticated models that group data records sharing similar trends or patterns. Jack Copper, NeuralWare's CEO, says "the ability to 'cluster' data patterns opens Predict to a wide range of new applications. It also will improve existing applications by providing the ability to pre-process and group data patterns together prior to building specific prediction or classification models. We are excited about the 3.0 release of Predict because it is the foundation for our next generation of advanced technology software."

Predict 3.0 has also been developed to be readily localized for international markets. A Japanese version of Predict 3.0 will be available shortly, and other localized versions are in the works.

If you are a Predict Technical Assistance Program (TAP) subscriber, your 3.0 update should be arriving during the next few weeks.

NeuralWorks Predict named finalist

NeuralWorks Predict® 2.41 has been named a finalist in the Scientific Computing & Instrumentation™ 2002 Readers' Choice Awards! The award was in the Neural Network/AI category. Additional information is available at www.scimag.com

An Introduction to Clustering in NeuralWorks Predict 3.0

Clustering (which also is known as segmentation or partitioning, as well as several other terms) is the process of analyzing a collection of data records to identify those which are related or associated, based on some measure or metric of "similarity." Even though the attributes that comprise a data record typically vary widely both in type and in the range of values that each attribute can assume, the similarity measure is often characterized as a "distance" in a high dimensional "attribute space."

The range of applications that can benefit from clustering technology is as broad as the range of applications that are appropriate for empirical modeling. Knowing the characteristics of customers whose purchase habits are similar can provide a basis for developing targeted marketing campaigns. Knowing the characteristics of individuals or enterprises whose spending patterns have led to defaulting on loans is a critical element of managing financial exposure and risk. Segmenting scientific or technical data based can lead to valuable insights that aid in understanding and using the data from which the clusters were derived.

As with other aspects of empirical modeling, a variety of techniques and methodologies exist that can be used to identify associative patterns in data. One fundamental characteristic of any approach is whether it requires prior knowledge of the number and/or type of clusters. In short, does the technique rely on supervised or unsupervised learning in the attribute space in order to group data records appropriately? Clearly, if the number of clusters must be specified in advance, either luck or some non-trivial amount of "pre-clustering" analysis must be performed before clustering can be performed – yet in many situations the number of clusters is not known, and in fact is the reason for applying a clustering algorithm to the data in the first place.

A second characteristic of clustering technology (assuming that a "distance" metric is the basis for similarity judgments), is that it is necessary to first normalize the values for all attributes. This prevents attributes whose values in domain units are significantly larger than other values of other attributes from distorting the distance computation, which would lead to erroneous weight updates and meaningless clusters.

With the release of Predict 3.0, NeuralWare has incorporated a robust clustering facility to complement and extend the proven prediction and classification technologies that are the foundation of the Predict product line. The Predict clustering facility is based on Self-Organizing Maps, originally conceived by Professor Teuvo Kohonen and his colleagues at the Helsinki University of Technology in Finland. The power of a Self-Organizing Map (SOM) lies in its ability to project relationships in high-dimensional attribute space into a more understandable (to a human analyst) two or three dimensional cartesian space, while preserving proximity relationships. As the name implies, this projection yields information that can be presented in the form of "maps" that highlight attribute associations.

NeuralWare's implementation of Self-Organizing Maps addresses many of the practical problems that inhibit the use of clustering technology. First and foremost, Predict SOM networks rely on the powerful data pre-processing facilities already present in Predict – there is no need to manually normalize data before presenting it to the SOM network. Next, Predict provides direct access to critical SOM architecture parameters, including map size, neighborhood type (for nearest neighbor calculations), and learning coefficients, so that alternative networks can be easily developed and evaluated. Finally, after a SOM is created and trained, Predict provides access not only to the "winning" Kohonen Processing Element (PE) when a new record is presented to the SOM, but also to the weights of all PEs and to the raw distance from the new record to all PEs in the Kohonen layer. This greatly extends the possibilities for constructing hierarchical maps or using SOM output with other unsupervised or supervised learning architectures that can further refine the associative relationships that the SOM discovers.

We hope this brief introduction to clustering with NeuralWorks Predict has been helpful, and has stimulated you to think how you might begin to take advantage of this powerful new feature. Please contact sales@neuralware.com to discuss your application requirements in detail, and watch for more information about using Self-Organizing Maps in future issues of Synapse.

Training

Register by October 30th for a NeuralWare course from the November through December series and receive a 10% early-bird discount. Courses are taught in Pittsburgh at our conveniently located headquarters just a few miles from downtown Pittsburgh.

Our regularly scheduled training courses ensure that our customers are knowledgeable about neural network technology and using our software effectively. See the schedule below to select the course that best fits your schedule. Make sure to visit our website or talk with our sales department to get complete course information before making your final selection.

Class size is limited to offer participants personalized instruction when necessary. A well-organized yet informal structure allows flexibility and provides opportunities for real-world information sharing and problem solving. When you visit the training website at www.neuralware.com/training.jsp, look for comments from previous attendees.

If you would like to expand your knowledge of neural network technology using NeuralWare's state-of-the-art tools, plan to attend a course soon. Don't forget the 10% early-bird discount. Contact us today!

Current training schedule

Applying Neural Networks to Business, Industry and Government (4.5 days)

September 23 – 27

November 4 -8

Advanced Neural Computing (3.5 days)

December 9 – 12

Product

Did you know that the latest version of Predict is now available for evaluation by going to www.neuralware.com/products_prdemo.jsp. We hope you will give Predict a "test-drive"!

Personal License Extension Information

If you are someone who needs to use your NeuralWare software both at home and at work, you are a candidate for NeuralWare's personal license extension (PLE). With a PLE you can request a second license key for your home or laptop computer for the low one-time cost of \$250. Only users who are on the technical assistance program (TAP) are eligible for this benefit. To take advantage of the TAP/PLE program, contact our sales department for more information at sales@neuralware.com.

Submitting an article

If you wish to submit a case study, application note, or technical tip for publication, please contact NeuralWare sales. If we publish your contribution, you will receive a one-year Technical Assistance Program (TAP) subscription for the NeuralWare program that you use - a minimum \$375 value. Submit your article to newsletter@neuralware.com.

Canceling your subscription

If you wish to discontinue your subscription to the newsletter, send email to newsletter@neuralware.com and ask to be removed from the distribution list.

Additional Information

Additional information on NeuralWare is available at www.neuralware.com! We hope you will visit us!

About this eLetter

This eLetter is copyright 2002 NeuralWare. All rights reserved. This eLetter may be freely distributed as long as the above copyright notice is included in any redistribution, electronic or printed.