Artificial Neural Network
Portion of Coil Study

LTC William M. Crocoll
School of Systems and Acquisition Management
US ARMY LOGISTICS MANAGEMENT COLLEGE
Bld 12500, 2401 Quarters Road, FT Lee, VA 23801-1705
email: crocollw@lee.army.mil, DSN 539-4231; Comm (804) 765-4231
http://www.almc.army.mil

INTRODUCTION TO ARTIFICIAL NEURAL NETWORKS

A neural network is a massively parallel system comprised of many highly interconnected, interacting processing elements (also called nodes or neurons) based on neurobiological models of the brain (Dayhoff, 1990). A major task for a neural network is to learn and maintain a model of the world consistent enough to achieve the goals of the application of interest (Haykin, 1994). These systems act as non-linear, non-parametric function estimators that learn to map inputs to outputs on a non-linear, multidimensional surface to fit general non-linear, multivariate functions (Zurada, 1992).

Neural networks exhibit many advantageous properties for solving complex problems. The following characteristics of neural networks emphasize their pattern recognition capabilities, making them particularly attractive for solving complex, data rich problems (Sheppard & Gent, 1991). They:

(1) can learn from examples and adapt to new situations;
(2) are robust (i.e., can abstract information from noisy, incomplete, and inaccurate data);
(3) can generalize from examples (i.e., can provide correct solutions from data similar to but not exactly like training data);
(4) can construct solutions quickly with no reliance on domain knowledge;
(5) can approximate any complex (ill-defined or unknown) multivariate function and form a classification decision from the recognition of discriminating patterns;
(6) are computationally efficient (i.e., have the speed to operate in real-time); and
(7) can implicitly account for the relative importance of input sources.

Since neural networks are adaptable (i.e., learn from experience) there is no need for an *a priori* mathematical model for input-output transformation. They learn correct responses by observing real world examples. With sufficient examples, neural networks exhibit strong generalization capabilities. Generalization is defined as the capability of producing reasonable responses to inputs similar to, but not exactly like, the previous inputs encountered during training. This adaptability also facilitates relatively easy retraining to deal with minor changes in the environment being modeled (Haykin, 1994). Neural networks are not programmed but trained. Because of this characteristic, they are best applied to problems whose solution requires knowledge which is difficult to specify but for which there is an abundance of examples (Sheppard & Gent, 1991). This makes neural networks ideal for data rich, theory poor applications. They learn by adjusting the interconnection strengths between the artificial neurons using computationally efficient algorithms for discovering appropriate non-linear transformations between the problem space and the solution space (Sheppard & Gent, 1991). The objective of learning is to determine the optimal interconnection strengths (i.e., weights) that provide the best approximation of the desired continuous, multivariable function based on a set of training examples (Zurada, 1992). A properly trained network should then be able to respond correctly to input not previously seen. This would indicate that the neural network has generalized the knowledge of the domain and not merely 'memorized' the training data.

CASCADE CORRELATION NETWORK OVERVIEW

The Cascade Correlation network (CCN) (Fahlman and Lebiere, 1990) is an example of the network growing approach. The procedure begins with a minimal network consisting of the appropriate application specified input nodes and output node(s) but no initial hidden neurons. Hidden neurons are added to the network one by one resulting in a multilayered structure. Each new hidden neuron receives a connection from each of the input nodes and each pre-existing hidden neuron. Thus the structure is a deep net of cascaded hidden nodes as opposed to a network with a wide hidden layer. As each new hidden neuron is added, its input weights are trained first. Then all of the weight connections to the output nodes are trained while leaving the other weights unchanged. Weights are adjusted to reduce overall network error typically using either the LMS rule or the Quickprop learning algorithm. This process continues until performance is judged adequate.

WHY ARTIFICIAL NEURAL NETWORKS ARE APPROPRIATE FOR FORECASTING CARAVAN INSURANCE POLICY PURCHASING.

The pattern recognition capabilities of artificial neural networks (e.g., ability to learn from examples, approximate complex non-linear multivariate functions, handle noisy, incomplete, and inaccurate data, and generalize) lend themselves nicely to the characteristics of this insurance policy purchasing problem. This appears to be a complex, data rich, theory poor type of problem with substantial historical data depicting the characteristics of each person as well as whether that person purchased a Caravan insurance policy or not. However, there is little theory to rely upon to develop a specific functional (read mathematical) form relating a person's characteristics (input variables) to whether or not they purchased a Caravan policy. It can be assumed that some complex non-linear, multivariate function maps the input variables to policy purchase or not, thus this becomes a pattern recognition task. The task is to recognize the pattern or relationship between the input variables and policy purchase. An artificial neural network will accomplish that pattern recognition by approximating the unknown and unknowable function mapping inputs to output and use that function approximation to make future predictions about a person's propensity to purchase a Caravan insurance policy.

MODEL DEVELOPMENT

The initial task was to define the problem and objectives as well as select the variables to consider in the model. In this study, the problem was defined by the CoIL Challenge contest specifications. Likewise, CoIL provided the historical data. The data was used as provided and no further data transformations were conducted. The data was then divided into segments. The 5822 records in the ticdata2000 was divided into two sets. One set consisting of 4322 records (278 who purchased a Caravan policy and 4044 who did not) was used for model development. The other set consisted of 1500 records (70 who purchased a Caravan policy and 1430 who did not) and was used to test the neural network models. Various Cascade Correlation neural network models were investigated. All neural network models were developed and tested on a 400 Mhz Compaq Desk Pro with Windows NT. A commercial software package (NeuroShell Classifier from Ward Systems Group) was used to model and test the neural networks.

NeuroShell Classifier implements an algorithm called TurboProp 2 which is a proprietary variant of the cascade correlation algorithm.

MODEL PERFORMANCE RESULTS (AKA. MODEL VALIDATION)

A split sample validation technique was used to evaluate each neural network model's performance. As mentioned previously, the data was divided into a model development segment and test segment. Once each neural network was trained on the model development set, it was used to provide predictions on the test set. Since the correct answers to the test set were known, each neural network model's performance was evaluated. The best performing network utilized 12 input variables and produced an accuracy of 71.43% for those, in the test set, that actually purchased the Caravan insurance policy. A classification and regression tree technique was used for the variable reduction.

CONCLUSION

The major conclusion that can be drawn from this study is that, given the input variables provided, an artificial neural network is a feasible technique for forecasting a person's propensity to purchase a Caravan policy. This can be used for marketing decision making. In this limited investigative proof of concept study, the Cascade Correlation network, represented as a proprietary variant called TurboProp 2 in NeuroShell Classifier, produced satisfactory results.

FORECASTS (IMPLEMENTATION)

The actual implementation of a neural network for future forecasts is relatively straight forward. I simply presented the input variable values for the 4000 records in the prediction set (ticeval2000) to the best network model and the network produced the forecast as its output. The forecast was the probability that an insurance policy would be purchased. I then rank ordered the 4000 records in the prediction set and selected the top 800 as those most likely to purchase an insurance policy. The only drawback to the use of a neural network for this application is the lack of an explanation for the forecast. The forecasts are a function of the network weights which were developed during the training process. These weights, collectively, represent the knowledge of the relationship between the input and output variables but is nearly impossible to translate into a semantic representation.

REFERENCES

Beastall, W. D. (1989). Recognition of radar signals by neural networks. In *Proceedings of the IEE Conference on Artificial Neural Networks*, 139-142.

Dayhoff, J. E. (1990). *Neural Network Architectures: An Introduction*. New York: Van Nostrand Reinhold.

Fahlman, S. E. and Lebiere, C. (1990). The cascade-correlation learning architecture. In D.S. Touretzky (Ed.), *Advances in Neural Information Processing Systems 2* (pp. 524-53). San Mateo, CA: Morgan Kaufmann.

Fredrickson, S. E., & Tarassenko, L. (1995). Text-independent speaker recognition using neural network techniques. In *Proceedings of the IEE 4th International Conference on Artificial Neural Networks*, 13-18.

Haykin, S. (1994). *Neural Networks: A Comprehensive Foundation*. New York: MacMillan Publishing Company.

Leung, H., Blanchette, M., & Haykin, S. (1993). Clutter cancellation and sea ice detection using artificial neural networks. In *Proceedings of the SPIE-The International Society for Optical Engineering: Applications of Artificial Neural Networks IV*, **1965**, 312-321.

Moody, J., & Darken, C. J. (1989). Fast learning in networks of locally-tuned processing units. *Neural Computation*, **1**, 281-294.

Paez, T. L. (1993). Neural networks in mechanical system simulation, identification, and assessment. *Shock and Vibration*, **1**(2), 177-199.

Sheppard, C. P., & Gent, C. R. (1991). A neural network based sonar classification system. In *Proceedings of the Europe 1991 MILCOMP Conference: Military Computers- Systems and Software*, 181-186.

SPSS, Inc. and Recognition Systems Inc. (1998). *Neural Connection 2.1*. Chicago, IL.

Ward Systems Group, Inc. (1998). *NeuroShell Easy Classifier*. Frederick, MD.

Zurada, J. M. (1992). *Introduction to Artificial Neural Systems*. New York: West Publishing Company.