17

Discussant's Response to "Neural Nets Versus Logistic Regression: A Comparison of Each Model's Ability to Predict Commercial Bank Failures"

Miklos A. Vasarhelyi
Rutgers University - Newark

Bell, Ribar and Verchio [1990] apply two alternative methodologies in the prediction of commercial bank failure. This discussion first examines the nature of the work; second, explores issues in neural network methodology; and third, concludes with the discussion of other relevant issues such as alternative approaches and paths for future work.

On the Nature of the Work

The bank failure problem has been extensively explored in the literature of accounting and finance. Consequently, there is a wide body of knowledge about the problem and substantial insight on analytical methods that help in the prediction of failure. This makes it an ideal arena for competitive methodological testing allowing for comparison, not only among the methodologies in question but also with an external body of literature.

The study uses an extensive sample from the 1983-1988 period for drawing failed banks and chooses through a sample estimation procedure. Part of the sample is held out for model testing purposes. The method is quite standard and has been used in many similar studies. Some more recent studies have used the jacknife/bootstrap method in order to avoid having to hold a large part of the data as a holdout sample. This approach could be adopted in this study leading to a different set of basic assertions. Nonfailed banks were chosen through a stratified sampling procedure for group pairing.

The authors used 28 prediction variables for failure prediction, very much in line with the literature. In these types of studies, you should always be concerned with two issues: over-fitting and missing variables. This study presents a relatively large sample, thereby decreasing some concerns with the first issue. The variables used are the standard financial variables that appear in most studies. These do not include any potential "soft" causes for failure (e.g. poor management, fraud) and/or macro variables.

In summary, the problem context and approach relate to a large set of studies in the literature and are an ideal setting for evaluating competing statistical methodologies. It might have been desirable that the authors further discuss the literature and the main results. The logit approach has been used extensively in research during the recent years while neural networks are a new and forthcoming area. Consequently, they are to be discussed next.

On Neural Nets

Much of the effort that started the neural network field, stems from the desire of scientists and researchers to understand the way the brain functions and to emulate its behavior. With a desire to obtain human-like capabilities (e.g. speech understanding, learning, vision), researchers have used computers as an alternative to the human being. A family of these devices is classified as neural computers and are based on what is called neural networks as noted by the following excerpt from Russo and Lewy [1989].

> Generally speaking, neural networks are an alternative, parallel computing architecture. Instead of being programmed like conventional algorithm computers, neural nets are trained and it is therefore arguable that they learn from and/or adapt to the nature of their input.

In essence, neural net software present a series of desirable characteristics:

- they are adaptive in nature,
- they can be retrained for improved performance,
- they are fault-tolerant due to their massive parallelism,
- their algorithms are typically non-parametric and therefore, they are more robust.

TABLE 1 drawn from Miller [1990] compares neural versus digital computers.

TABLE 1
NEURAL COMPUTERS VS. DIGITAL COMPUTERS

DIGITAL DATA PROCESSING	ANALOG DATA PROCESSING
YES & NO DECISIONS BASED ON MATHEMATICAL AND LOGICAL FUNCTIONS	WEIGHTED DECISIONS ON THE BASIS OF FUZZY, INCOMPLETE, AND, CONTRADICTORY DATA
DATA HANDLED IN A RIGID STRUCTURED SEQUENCE	INDEPENDENTLY FORMULATE METHODS OF PROCESSING DATA
FIND PRECISE ANSWERS TO ANY PROBLEM, GIVEN ENOUGH TIME	FIND GOOD, QUICK, ... BUT APPROXIMATE .... ANSWERS TO HIGHLY COMPLEX PROBLEMS
SORT THROUGH LARGE DATABASES TO FIND EXACT MATCHES	SORT THROUGH LARGE DATABASES TO FIND CLOSE MATCHES
STORE INFORMATION TO RETRIEVE SPECIFIC INFORMATION	STORE INFORMATION TO ALSO RETRIEVE RELATED FACTS

Nevertheless, in spite of the promise and potential of this technology, there are many myths and hype that surround its usage. TABLE 2 represents a partial list of hypes and realities in neural nets.

TABLE 2
HYPE AND REALITY IN NEURAL NETS*

HYPE (H):	MANY SUCCESSFUL NN APPLICATIONS EXIST
REALITY (R):	HARDWARE HAS NOT YET CAUGHT UP WITH THE APPROACH AND IS HOLDING THINGS BACK
H:	NEURAL NETS PERFORM BETTER THAN BAYESIAN CLASSIFIERS
R:	BAYESIAN CLASSIFIERS ARE OPTIMAL BUT CANNOT BE CONSTRUCTED WITH MANY VARIABLES
H:	VERY FAST PROTOTYPING
R:	ONLY FOR SMALL NETWORKS TRAINING AND DATA GATHERING TAKES TIME
H:	PARALLEL, THEREFORE FAST
R:	SERIAL IMPLEMENTATION, PRACTICAL VLSI 3-7 YEARS AWAY
H:	PATTERN RECOGNITION AND DETECTION IS EXTREMELY POWERFUL, ROBUST AND TOLERANT TO NOISE
R:	TRUE— > BUT MUST BE DESIGNED AND TRAINED

* Adapted from Russo and Lewy [1989], Miller [1990], and Widrow [1988].

A simple neural network will have three layers: 1) the input layer, 2) the hidden layer and 3) the output layer. The second layer has some form of internal representation that can be known or unknown.

There are two major types of neural networks: feedback and feed-forward. Feedback nets or associative memories can store many template patterns by presetting the network weights. A feed-forward multi-layer network is composed of rules (not templates) that are learned and stored as the weighted connections of the network.

Neural nets have been discussed as a family of technologies and/or implementations of similar nature. Among the types of neural network approaches to learning and/or propagation and conflict resolution, we find: A) Grandmothering [Caudill, 1990], B) delta or least mean square rules (LMS) [Widrow and Hoff, 1960], C) Back-propagation [Rumelhart, McClelland et al., 1986], D) Kohonen self-organizing network [Kohonen, 1984], E) Outstar learning [Grossberg, 1982], F) The avalanche model [Caudill, 1990], and G) Adaptive resonance theory [Grossberg, 1982].

The Bell and Ribar paper used NeuralWare's software (back-propagation) using 11 predictor variables in the input layer, and eight nodes in the hidden layer. It was trained using 102 banks from the 1984 estimation sample and 102 of the non-failed banks drawn randomly. The network was trained with about 300,000 iterations and after the examination of a Hinton diagram it was decided that two nodes were to be dropped. This model was compared with a rather different logit model.

A point to ponder, in this type of comparison is that both the neural model and the logit model are the product of substantial tinkering in the attempt to improve failure predictability but there is no assurance that they provide any form of best predictor. Second, considering the model structure and data differences, it is not clear if there is any information superiority in one of the approaches. The countervening approach would attempt to keep parity between the approaches at the cost of not "best" utilizing its features.

Some Further Issues

Researchers differ in approach and taste when dealing with particular problems. The authors should be encouraged to continue this line of investigation. It is seldom that a commercial firm endeavors in the examination of emerging technologies with the seriousness and methodological quality that is found in this study. A few suggestions however are of relevance.

This study lacks somewhat in the examination and comparison of its results with the extant literature. Substantial insights may be acquired by better positioning this methodological study in relation to the bank failure literature. Furthermore, the quality and nature of predictions by the neural net model should be compared not only with the logistic model but also with the nature of the findings in the literature. Secondly, as discussed in the introduction of this paper, considerable insights of a methodological nature can be acquired from an examination of better and worse performing models and the nature of the predictions obtained.

The authors may also consider, after some preliminary variable selection work, performing their comparison on the same set of variables as opposed to varying the basic variables set. The approach adopted by the authors may be confused as a methodological improvement while the variance may have been explained due to the inclusion of additional variables.

The most important issue is that this study is a comparison of the logit technology with a particular type of neural net approach and algorithm, specifically, the back-propagation method. Neural net, more than many other technologies, is a generic name for many, often non-similar approaches, and within each, dissimilar families of more similar types of approaches are found. Consequently, this work tested a particular type of neural net implementation versus the logit model. This study has not answered the question of what type of neural net approach or what type of structure would be the most promising for the bank failure problem.

In conclusion, the study found neural net technology to have some promise in a bankruptcy prediction context. From the nature of the findings and of the approach, many fascinating ideas for extension and application to other areas may follow. For example, can a model be developed and trained to evaluate a sample taken in an audit, based on some of its context variables, to increase the probability of finding values in error? Can the same model be extended, not only to predict failure but to rate and point out different levels of financial distress?

This paper deals with a very rich area and much was clearly learned just by manipulating the neural model to improve predictability of bank failure. An interesting extension would be to compare different neural algorithms among themselves and in relation to the logit model.

References

Bell, T. B., G. S. Ribar, and J. Verchio, "Neural Nets Versus Logistic Regression: A Comparison of Each Model's Ability to Predict Commercial Bank Failures," *Proceedings of the 1990 Deloitte & Touche/University of Kansas Symposium on Auditing Problems*, pp. 29-53.

Brownston, L., R. Farrell, E. Kant, and N. Martin, *Programming Expert Systems in OPS5: An Introduction to Rule-Based Programming* (Addison-Wesley Publishing Company, 1986).

Caudill, M., "Neural Nets Primer," (part I), *AI Expert* (The Fairmont Press, 1990).

Elliot, L.B., "Applications of Neural Networks," presented at the *Second International Symposium on Expert Systems in Business, Finance, and Accounting* (University of Southern California, 1989).

Grossberg, S., *Studies of Mind and Brain* (Drodrecht, Holland: Reidel Press, 1982).

King, P.M., "Audit Issues and Skills Involving Neural Net Technology," presented at the *Second International Symposium on Expert Systems in Business, Finance, and Accounting* (University of Southern California, 1989).

Kohonen, T., *Self-Organization and Associative Memory* (Springer-Verlag, 1984).

Miller, R.K., *Neural Networks* (The Fairmont Press, 1990).

Ney, H., and L.R. Rabiner, "On Training Criteria for Pattern Recognition in Statistical and Neural-Network Approaches," *AT&T Bell Laboratories, Private Communication* (1989).

Pike, R., D. Presotto, K. Thompson, and H. Trickey, "Plan 9 From Bell Labs," *AT&T Bell Laboratories, Private Communications* (1990).

Rumelhart, D., and J. McClelland, *Parallel Distributed Processing* (MIT Press, 1986).

Russo, A.P., and D.B. Lewy, "Neural Networks: What They Are; Where They Are; Why They Merit Company Support," *AT&T Bell Laboratories, Private Communication* (1989).

Sharma, R., "VLSI Electronic Neural Networks," *AT&T & Bell Laboratories, Private Communication* (1988).

Solla, S.A., "Supervised Learning and Generalization," *AT&T Bell Laboratories, Private Communication* (1990).

Widrow, B., and M. E. Hoff, "Adaptive Switching Circuits," *Institute of Radio Engineers WESCON Convention Record*, part 4 (August 1960), pp. 96-104.