Pergamon

Expert Systems With Applications, Vol. 10, No. 2, pp. 233–241, 1996

0957-4174/96 $15.00 + 0.00

0957-4174(95)00049-6

Aggregating and Updating Experts' Knowledge: An Experimental Evaluation of Five Classification Techniques[1]

BRENDA MAK

Cleveland State University, Department of Computer and Information Science, Cleveland, OH 44115, U.S.A.

TUNG BUI

Naval Post-Graduate School, Department of Systems Management, Monterey, CA 93943, U.S.A.

ROBERT BLANNING

Vanderbilt University, Owen Graduate School of Management, Nashville, TN 37203, U.S.A.

Abstract—*Knowledge acquisition consists of eliciting expertise from one or more experts in order to construct a knowledge base. When knowledge is elicited from multiple experts, it is necessary to combine the multiple sources of expertise in order to arrive at a single knowledge base. In this paper, we present and compare five techniques for aggregating expertise. An experiment was conducted to extract expert judgments on new product entry timing. The elicited knowledge was aggregated using classical statistical methods (logit regression and discriminant analysis), the ID3 pattern classification method, the k-NN (Nearest Neighbor) technique, and neural networks. The neural net method was shown to outperform the other methods in robustness and predictive accuracy. In addition, the explanation capability of the neural net was investigated. The contributions of the input variables to the change in the output variable were interpreted by analyzing the connection strengths of the neural net when the net stabilized. We conclude by discussing the use of neural nets in knowledge aggregation and decision support.*

1. INTRODUCTION

BUSINESS LEADERS often face strategic decisions whose contexts have little to do with the past. They have little historical data (Tavana, Lee, & Joglekar, 1994) to learn from, and are forced to rely on experts' judgments. This was described by Simon (1960) as non-programmed decision-making, involving the use of heuristic problem solving techniques. In order to construct these heuristics, it is necessary to elicit information from experts and to present it to managers—either directly or in the form of an expert system. We address here the case in which multiple experts are available and it is necessary not only to obtain information from each of them but also to aggregate the information thus obtained. When the experts have different areas of expertise, each one is captured in a separate knowledge-base (Barrett & Edwards, 1995). But when they share the same area of expertise, it is necessary to aggregate their knowledge to a single knowledge-base. We are concerned here with this latter case.

Group decision techniques, from the simple majority rule to more elaborate techniques, are available for knowledge aggregation (e.g. Jessup & Valacich, 1993). Group voting algorithms, however, are not effective for the treatment of multiple (similar yet different) problems and multiple experts. It is theoretically unsuitable to combine vote results from different problems. Unlike historical data, data samples collected from multiple experts are often heterogeneous. The heterogeneity makes it difficult to use conventional statistical analyses to group or classify data. Unfortunately, there is no evidence that suggest which statistical method would be most appropriate to aggregate non-historical and subjective data.

[1]An earlier version of this paper was presented at WITS-94 and published in the proceedings of the conference.

The authors would like to thank Philip Kotler, Marvin Manheim, and Dirk Ruiz from Northwestern University for providing comments on the methodology adopted in this paper.

The objective of this study is to evaluate methods for aggregating heterogeneous judgment data across experts. We test the performance of the most popular classical classification methods, i.e., multivariate discriminant analysis (MDA), the logit method, and compare them with the ID3, the *k*-NN (Nearest Neighbor), and the neural net approaches. In particular, we compare the methods along the following criteria: robustness, predictive accuracy, adaptability, and explanatory capability. For this purpose, we conduct an experimental study soliciting marketing experts' judgments for new product entry decisions. The data are used to compare the effectiveness of the five classification techniques.

2. KNOWLEDGE AGGREGATION AND UPDATING

In this study, we assume that all experts consulted have a valid view of the problem. We borrow Arrow's Impossibility Theorem of Social Choice to contend that a consensual solution derived from all experts' judgments collected is a socially acceptable solution. This assumption is particularly useful when the experts possess different views of the problem. However, judgments collected from multiple experts are, more often than not, heterogeneous. The heterogeneity arises from two sources. First, different experts have different views on how to solve a problem. Second, the decision problem is often so fragmented that not all of the important variables are clearly understood. Under these conditions, it is not always suitable to use a group decision technique (such as the majority rule) to look for the answer, as this might rule out the "minority" but the correct solution.

The difficulty of implementing a formal decision technique calls for a more pragmatic approach. The proposed approach takes into consideration as many as possible all the factors that are in the decision problem. It also captures the knowledge of all available and qualified experts, and aggregates this knowledge and stores it in a form useful for decision support. The central issue is thus how to aggregate and update this "collective" knowledge. In a previous study (Mak & Bui, 1995), we proposed a knowledge acquisition and aggregation procedure consisting of three main steps. First, business cases are used as a knowledge acquisition vehicle to extract experts' problem assessments and decisions. Second, acquired knowledge is aggregated using statistical models. Third, if the knowledge aggregation process is successful and the model validated, the model can then be used as a decision aid. These three steps are briefly discussed below.

2.1. Knowledge Acquisition from Multiple Experts using Multiple Cases

The use of skillful experts is an effective means of assembling contingencies into a workable decision model. If more than one expert are available, one must either select the opinion of the best expert or pool the experts' judgments (Taylor, Weimann, & Martin, 1995; Barrett & Edwards, 1995). It is assumed that when experts' judgments are pooled, collectively they offer sufficient cues leading to the building of a comprehensive theory (Wielinga et al., 1993). The argument here is germane to the concept of a consensual model discussed in the Computer-Supported-Cooperative-Work literature (Bui, 1987). Anderson (1990) demonstrates that humans tend to solve problems by analogy. Thus, we contend that using cases that are similar to those that experts have dealt with is an appropriate way to extract knowledge from them (Srikanth, 1994). Further we assume that asking experts to analyze and solve many different cases will increase the robustness of the knowledge-based model (Falkenhainer & Forbus, 1991). Therefore cases are used to solicit experts' qualitative judgments.

2.2. Knowledge Aggregation

"Chunks" of knowledge collected from experts are still functionally fragmented. It is necessary to devise a method to aggregate them. Traditional ways of aggregating knowledge from multiple experts include Bayesian method decision theory, certainty factors, and fuzzy logic (Castillo & Alvarez, 1991). All these techniques require experts to assign probabilities or other numerical values to rules and outcomes. However, as experts tend to think qualitatively, it is often difficult for them to specify explicitly the quantitative values associated with choosing different strategies. Therefore, we use qualitative scenarios to retrieve knowledge from the experts.

In aggregating knowledge, we seek to identify the significance of each of the extracted factors and the functional inter-relationships among the relevant factors. Regression techniques are widely recognized as statistically the most robust techniques to derive functional relationships between dependent and independent variables. However, when the dependent variable is a descrete outcome (e.g. yes or no), conventional regression methods are inappropriate (Greene, 1993), and other techniques are required. In the following sections, we compare the performance of five such techniques in classifying the judgment data.

2.3. Decision Aid and Simulation

Once validated, the aggregated knowledge-based model can be used as a simulation tool, allowing decision makers to perform "what-if" analysis on various related scenarios, and update their business knowledge. Decision analysis based on a validated knowledge-based model can be useful both in solving specific problems presented to the user and also in training users and helping them to understand the class of real-world problems they are trying to solve (Kolodner, 1993).

3. CLASSIFICATION TECHNIQUES FOR KNOWLEDGE AGGREGATION

The neural net method as well as other statistical techniques have been applied in the analysis of business decisions (Wong, Bodnovich, & Selvi, 1995). In these studies, the neural net method has been found to be more robust (Subramanian, Hung, & Hu, 1993) than discriminant analysis. For example, the neural net method has been shown to be more robust and have better predictive ability than discriminant analysis in treasury bond classification (Walker & Devaney, 1993), stock price prediction (Swales & Yoon, 1992), corporate failure prediction (Odom & Sharda, 1990; Raghupathi, Schkade, & Raju, 1991), bank bankruptcy prediction (Tam & Kiang, 1992), and the prediction of the financial health of thrift institutions (Salchenberger, Mine, & Lash, 1992). However, in most of these studies, the focus has been the analysis of financial time series data. In this paper, we seek to analyze the performance of neural nets in a different context. We apply the neural net method to search for consensus among multiple experts. We then test the predictability and adaptability of the knowledge-based model developed.

3.1. Discriminant Analysis

Discriminating analysis was used for the analysis of two-group situations. It involves developing a linear combination of the set of group characteristics (e.g., income, lot size) that will maximally differentiate among the groups in question. Consider an arbitrary combination of income $x1$ and lot size $x2$:

$$Y = f1 \times x1 + f2 \times x2$$

where Y is the discriminant score, $f1$ and $f2$ are the arbitrary weights. The weights are derived so that the variation in the discriminant scores between the groups is as large as possible.

In the case of binary classification, the discriminant function is as follows:

$$D(X) = X'\Sigma^{-1}(\mu_1 - \mu_2) - 1/2(\mu_1 - \mu_2)^{\mathrm{T}}\Sigma^{-1}(\mu_1 + \mu_2)$$

where, μ_1, μ_2, and Σ^{-1} are the mean vectors and inverse of the common covariance matrix respectively. The classifiers are optimal in minimizing the expected cost of misclassifications, when the following conditions hold: (i) the costs of misclassifying an object into different groups are equal; (ii) the *a priori* probability of each group is equal; (iii) the distribution of the variables in the population is multinomial with equal and known covariance matrices. In reality, these assumptions are often violated. A possible solution is to transform variables, or to use a quadratic (instead of linear) discriminant function. However, the quadratic classifiers are difficult to interpret and possess poor predictive ability. To solve this problem, the probability of belonging to a group is sometimes used as the dependent variable. The following logistic regression is then adopted:

$$Y = 1/(1 + \exp(\Sigma c_i \times x_i))$$

where $x_0 = 1$, x_i, $i = 1..N$ are the independent variables, and c_i is the coefficient of the ith independent variable. Y is the dependent variable, and is the probability of a class outcome. The logit maximum likelihood estimator is more consistent and robust than the discriminant analysis estimator when normality is violated (Amemiya, 1985).

3.2. The ID3 Approach

Unlike classical statistical methods that use discriminant functions, the ID3 method develops a discrimination tree by induction for classifying the training sample. The ID3 approach is valid and useful when there is a body of data consisting of a large number of patterns with numerous attributes. The data are examined to find out the minimum combination of attributes required to determine class membership. The main advantage of the ID3 procedure is the ease with which it can be automated (Pao, 1989). However, the ID3 method has one main drawback, in that one cannot readily update the decision tree without having to rebuild the entire tree. The ID3 has been applied to a number of business situations such as credit scoring (Carter & Catlett, 1987), corporate failure prediction (Messier & Hansen, 1988) and stock portfolio construction (Tam & Kiang, 1990).

3.3. The *k*-NN classification method

The k-NN (Nearest Neighbor) method is a classification method based on distance. A set of previous cases is stored, and when a new case is encountered, the system calculates the k most similar cases. Generally, $k = 1$, and so the system finds the most similar case (called the nearest neighbor) and recommends that case to the decision maker. Kim et al. (1993) compared the classification ability between backpropagation neural network and k-Nearest Neighbor classification methods, and found that the average recognition rate of neural net is better than k-NN. In the Nearest Neighbor method, the system carries out match-based learning, and finds the category that best matches the input (Carpenter & Grossberg, 1992; Goodman et al., 1992). This is different from the error-based learning approach found in back-propagation neural networks described below.

3.4. Neural Networks

Neural networks are composed of many simple elements operating in parallel. These elements are inspired by biological nervous systems. The network function is determined largely by the connections between elements. The neuron model and the architecture of a neural network describe how a network transforms its input into an output. For example, consider a neuron with a single input p. The input p is transmitted through a connection that multiplies its strength by weight w, to form the product y, where $y = w \times p$. Here the weighted input $w \times p$ is the only argument of the transfer function F, which produces the output a. The w is an adjustable parameter of the neuron. The transfer function F is typically a sigmoid function, given by:

$$F(y) = 1/(1 + \exp(-y)).$$

The sigmoid function transforms the input (which may lie between plus and minus infinity) into the range of 0 to 1. It is commonly used in backpropagation networks, in part because it is differentiable. Two or more of the neurons may be combined in a layer, and a particular network might contain one or more such layers. A one-layer network with R inputs and S neurons is shown in Figure 1. Each element of the input vector P is connected to each neuron input through the weight matrix W. Each of the S neurons has a summer, and the summer outputs form an S element vector N. The input of the transfer function is the sum of weighted input $w \times p$. These sums are the arguments of the transfer functions. Finally, the neuron layer outputs to a column vector A.

One of the most commonly used algorithms in neural net training is backpropagation (Rumelhart et al., 1986). Backpropagation networks often have one or more hidden layers of sigmoid neurons followed by an output layer of linear (or sigmoid) neurons. The initial matrix of connection strengths for the backpropagation network is created with random elements between -1 and 1. This allows the connection strengths to be changed easily and yet leaves enough variation in them so that neurons in the network can update the connection strengths with the behavior of the training data. In the backpropagation learning rule, the sum of squared error of the network is minimized. This is done by continually changing the values of the network connection strengths in the direction of the steepest descent with respect to error. Backpropagation will minimize the sum of squared errors if the model is not trapped in a local optimal (Marquez et al., 1991).

Neural nets are found to perform best under conditions of high noise and low sample size (Marquez et al., 1991; Subramanian, Hung, & Hu, 1992), and when the data pattern is complex and nonlinear (Curram & Mingers, 1994). There are several limitations that restrict the use of neural network models for prediction (Salchenberger, Mine, & Lash, 1992). The development of a neural net requires expertise and the training is computationally intensive. Also there is no formal rules specifying how many hidden layers, and how many units per layer should be used. It has also been proven that only one hidden layer is sufficient to approximate any continuous function (Cybenko, 1989; Hornik, Stinchcombe, & White, 1989; Dutta & Shekhar, 1988). Salchenberger, Mine, & Lash (1992) has also verified empirically that more hidden layers did not improve the classification results. Besides the number of hidden layers, the network performance is also affected by the number of elements in the hidden layer (Watkin, Rau, & Biehl, 1993). Too few neurons in the hidden layer will lead to underfitting, when no sets of connection strengths can generate the desired target. On the other hand, too many neurons in the hidden layer will result in overfitting, in which case all training points are well fitted, but the fitting curve oscillates and fails to predict new data (Curram & Mingers, 1994). In the following section, we describe an experiment that compares the performance of the five techniques for knowledge aggregation.

4. AN EXPERIMENTAL STUDY—NEW PRODUCT ENTRY DECISION

4.1. Research Design: Experimental Problem, Subjects, and Procedure

An experiment was conducted in which experts gave their recommendations on new product entry strategies when different scenarios were specified. This method of scenario analysis has been used by researchers in strategic planning (Porter, 1980). Experts' judgments were constructed based on their assessment and decision of entry of new consumer electronics products. The products included a TV phone, a beeper watch, a computerized piano learning system, a notebook com-

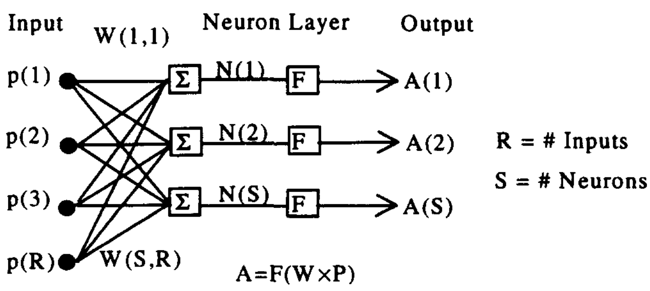

FIGURE 1. A neuron with R inputs and S neurons.

puter, a VCR with compact discs, a car with automatic check, a voice-driven computer, a Hi-Fi-television set, a daily reminder, and a scanner. Magazines on high-tech developments were used as a reference basis to identify the ten hypothetical products. To increase real-life complexity to decision situations each case consisted of a product script.

The script was provided in a way that the expert could assess the situation based on seven scenario variables derived from key findings in the marketing literature (Mak & Bui, 1995). These variables were previously identified in a protocol analysis (Simon & Ericsson, 1986). For computational simplicity, one of the two possible values was given for each variable. The variables and their values are as follows: the position of the firm (dominant vs small), the financial strength of the firm (strong vs weak), the expected demand growth (high vs low), the product's life cycle (long vs short), diffusion across competitors (fast vs slow), cannibalization (high vs low), and the cost of market development (high vs low). With the products and the scenario variables, we generated a case by selecting a new product and a scenario template for each variable. Each subject chose the new product entry strategy (Go/No-Go, denoted as 1, 0 of the choice outcome) for the product described in the case.

The experts were 36 MBA students with previous work experience and were in their senior year at the Kellogg Graduate School of Management. Each subject was given a total of ten cases. The first five cases were used as the *training set*, and the next five were used as the *test set*. To reduce the effect of uncertainty in knowledge acquisition, each subject was given 10 different scenarios. The more the evidence collected, the higher was the confidence on the solicited judgment (expert consistency). Moreover, to test the consistency of the aggregated model, we sought to generalize the model as much as possible, by providing different sets of scenarios for different experts (group consistency).

4.2. Data Analysis Procedures and Results

To demonstrate the iterative nature of our methodology, we used the first five decisions entered by the experts as the training set to develop the model. We then tested the predictive ability of the model with the sixth and seventh decisions, and checked the adaptability of the model by finding its on-going predictive ability for the sixth to tenth decisions of the experts. The details of the methods used to analyze the data are summarized in Table 1.

5. EVALUATION FOR ROBUSTNESS, PREDICTIVE ACCURACY, ADAPTABILITY, AND EXPLANATORY CAPABILITY

The five methods of aggregating experts' knowledge are compared. Given below are the tables for validation, prediction, and adaptation for the five methods. We measured the three evaluation criteria as follows: (i) Robustness: the rate of correct prediction for the original training data set; (ii) Predictive Ability: the rate of correct prediction for a test data set (6th and 7th decision); (iii) Adaptability: the on-going rate predictive

TABLE 1
Methods Used to Analyze the Judgment Data

Method	Tools Used	Details of Operation
Discriminant analysis	Discriminant analysis procedure on the SPSS package	Fisher's linear discriminant function was used to discriminate among the four strategies and to predict choices
Logistic model	Multinomial logit model in TSP	Probabilities for choosing timing strategies were computed, and the strategy with the highest probability was selected
ID3	PASCAL programs	Partitioning iterated until cost of further dividing the data set outweighed benefits of additional partitioning
k-NN	PASCAL programs	1-NN is used to find the nearest neighbor(s) in the training set. The predicted strategy for the test item is the strategy most frequently used for the nearest neighbor(s). *k*-NN ($k=2..N$) is used to break ties
Neural network	PASCAL programs (Backpropagation algorithm)	A neural net of one hidden layer (3 elements) was used. Backpropagation with gradient descent algorithm was used. A variable learning rate was used to fine tune to the error minimum. The training results showed that minimum error was attained when the number of elements in the hidden layer was three, and adding a fourth element did not reduce the error. Based on the principle of parsimony (Watkin, Rau, & Biehl, 1993; Yeung, 1993; Hertz, Krogh, & Palmer, 1991), we used a model with three elements in the hidden layer.

accuracy for an additional test data set (6th to 10th decision).

5.1. Robustness, Predictive Accuracy, and Adaptability

Tables 2–6 report the performance of the logit, the discriminant analysis, ID3, *k*-NN, and the neural net model. Table 7 summarizes the results of the comparison of the five classification techniques. Overall, the neural net method gives the best robustness (about 90% for neural net, 80% for logit and ID3, and 70% for *k*-NN and discriminant analysis). The predictive ability and adaptation are roughly the same for the five methods (about 70%).

The predictive ability and adaptability of the neural net method only average about 70%. This may be due to

TABLE 2
Logit Model

	Robustness		Predictive ability		Adaptability	
	1	0	1	0	1	0
1	58	9	32	3	75	13
0	13	25	10	6	23	17
Hit rate	79.05%		74.51%		71.88%	

TABLE 3
Discriminant Model

	Robustness		Predictive ability		Adaptability	
	1	0	1	0	1	0
1	49	18	25	10	75	13
0	10	28	4	12	23	17
Hit rate	73.33%		72.55%		71.88%	

TABLE 4
ID3 Model

	Robustness		Predictive ability		Adaptability	
	1	0	1	0	1	0
1	58	9	30	5	70	18
0	8	30	10	6	17	23
Hit rate	83.8%		70.59%		72.66%	

TABLE 5
***k*-NN Method**

	Robustness		Predictive ability		Adaptability	
	1	0	1	0	1	0
1	53	14	26	9	69	19
0	18	19	7	9	20	20
Hit rate	68.6%		68.6%		69.5%	

TABLE 6
Neural Net Model

	Robustness		Predictive ability		Adaptability	
	1	0	1	0	1	0
1	61	5	29	9	71	19
0	6	33	6	7	17	21
Hit rate	89.52%		70.59%		71.88%	

the difference in the data patterns of the training and test set. As shown in past studies (Patuwo et al., 1993), if the training and test data sets have different patterns, then the procedures better at classifying training sets are not as good at classifying test sets. The tradeoff between generalizability and validity is also observed in our study. Thus a more powerful method better able to classify a given data set might be less generalizable.

5.2. Explanatory Capability

In spite of its better ability to validate and predict, the neural net method has been criticized for its poor explanatory capability. It is difficult to explain why a particular conclusion is reached (Yoon Guimaraes, & Swales, 1994). A neural net may give accurate predictions, but it does not explain how these predictions are reached and why they should hold (Garson, 1991). A neural net is also limited if one wants to test the significance of individual inputs (Tam & Kiang, 1992), and no confidence interval can be assigned to the estimated connection strengths. There is no reported formal method to derive the relative importance of an input from the connection strengths of a neural net. Expert systems, however, are known for their ability to explain how they reach conclusions (Castillo & Alvarez, 1991). With a rule-based system there is a chain of reasoning that can be made explicit.

However, with the techniques used here, the most useful explanatory information is the relative contributions of the various input variables to the conclusion. Garson (1991) suggested a method to evaluate the contributions of the inputs to the outputs. Garson's measure is based on the connection strengths of the net. Garson's measure of relative impact of input I on output k is:

$$\mathrm{Imp}_{ik} = \frac{\sum_{j=1}^{J} \frac{|w_{ji}|\,|v_{jk}|}{\sum_{i=1}^{I} |w_{ji}|}}{\sum_{i=1}^{I} \sum_{j=1}^{J} \frac{|w_{ji}|\,|v_{jk}|}{\sum_{i=1}^{I} |w_{ji}|}}$$

where w_{ji} are the connection strengths from ith input to jth element in the hidden layer, v_{jk} are the connection strengths from jth element in the hidden layer to kth output, I = total number of inputs, and J = total number of elements in the hidden layer.

Table 8 gives the rank order of the importance of the variables for the five methods. In the logit method, the size of the coefficients of the significant variables (t-stat) is used to determine the importance of the variables. The size of the coefficients of the significant variables (as determined by the F-statistic at the 0.05 level of significance) is used to find the rank order of importance for the variables in the discriminant analysis method. The amount of entropy reduction is used to determine the order for the ID3 method. In the k-NN method, the importance of variables is determined by removing individual variables and checking the corresponding reduction in robustness. Finally, in the neural net method, Garson's (1991) measure, based on the size of the connection strengths, is used to find the rank order of the variables. It can be observed from the data presented here that variables "expected demand growth" and "position of the firm" are the key factors affecting the Go/No-Go decisions. Additional measures of relative importance for the neural net inputs can be found in Mak & Blanning (1995).

6. CONCLUSION AND FUTURE RESEARCH

In this paper, we have applied five methods to aggregate expert judgments, and performed a comparative evalua-

TABLE 7
Summary of Results

	Logit model	Discriminant analysis	ID3	k-NN ($k=1$)	Neural net-model
Robustness	79.05%	73.33%	83.8%	68.6%	89.52%
Predictive ability	74.51%	72.55%	70.59%	68.6%	70.59%
Adaptability	71.88%	68.75%	72.66%	69.5%	71.88%

TABLE 8
Rank Order of the Importance of Variables*

	Logit model	Discriminant analysis	ID3	Neural net (Garson's method)	*k*-NN ($k=1$)
Position of the firm	2	2	2	1	2–3
Financial strength of the firm	7	7	7	2	1
Expected demand growth	1	1	1	3	2–3
The product's life cycle	5	4	3	6	7
Diffusion across competitors	6	6	4	4	5–6
Cannibalization	4	5	6	5	4
The cost of market development	3	3	5	7	5–6

* Detailed data for logit, discriminant, ID3 and neural net methods appear in Mak & Blanning (1995).

tion based on robustness, predictive accuracy, adaptability, and explanatory capability. We also propose an extension of the analysis of neural net connection strengths to interpret the contributions of the input variables. Unlike previous studies that made use of time series data, we applied the neural net method to aggregate heterogeneous judgment data. In our study, the neural net method scored better than the classical statistical methods in data classification. The finding is consistent with much of the literature comparing neural nets with statistical methods. We also found that neural net algorithms are more robust in analyzing qualitative data. A main drawback of using neural nets as a decision aid is the difficulty in interpreting the solution compiled by the method.

The findings of this study may have far-reaching implications for the development of knowledge-based systems. Systems based on neural nets can be used as a powerful decision support and training tool (Medsker & Turban, 1994) for decision makers who want to learn about their own cognitive thinking and the thinking of other decision makers and experts. Such an approach has not been developed in detail in the literature on neural nets based on time series data.

Another extension of this study involves validating the findings of this experiment with more case studies. The robustness of the neural net method in knowledge aggregation should be demonstrated for other types of cases in addition to new product entry timing. Further, the effect of the decision-making style on the performance of the classification methods should also be explored. Through understanding the fit between the decision-making style of the experts with the algorithmic characteristics of the classification techniques, better knowledge aggregation tools for use in decision support can be designed.

REFERENCES

Amemiya, T. (1985). *Advanced econometrics.* Harvard University Press.

Anderson, J. (1990). *Cognitive psychology.* Freeman.

Barrett, A. R. & Edwards, J. S. (1995). Knowledge elicitation and knowledge representation in a large domain with multiple experts. *Expert Systems With Applications*, **8**, 169–176.

Bui, T. (1987). *A group decision support system for cooperative multiple criteria group decision making* (Lecture Notes in Computer Science, 290). New York: Springer.

Carpenter, G. & Grossberg, S. (1992). A self-organizing neural network for supervised learning, recognition, and prediction. *IEEE Communications Magazine*, September, 38–49.

Carter, C. & Catlett, J. (1987). Assessing credit card applications using machine learning. *IEEE Expert*, Fall, 71–79.

Castillo, E. & Alvarez, E. (1991). *Expert systems: Uncertainty and learning, computational mechanics publications.* Co-published with Elsevier Applied Science.

Curram, S. & Mingers, P. J. (1994). Neural networks, decision tree induction and discriminant analysis: An empirical comparison. *Journal of the Operational Research Society*, **45**, 440–450.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems*, **2**, 303–314.

Dutta, S. & Shekhar, S. (1988). Bond-rating: A non-conservative application of neural networks. *Proceedings of the IEEE International Conference on Neural Networks*, Vol. 2, pp. 443–450.

Falkenhainer, B. & Forbus, K. D. (1991). Compositional modeling: Finding the right model for the job. *Artificial Intelligence*, **551**, 95–144.

Garson, G. D. (1991). Interpreting neural-network connection strengths. *AI Expert*, April, 47–51.

Goodman, P. H., Kaburlasos, V. G., Egbert, D. D., Carpenter, G. A., Grossberg, S., Reynolds, J. H.., Rosen, D. B., & Hartz, A. J. (1992). Fuzzy ARTMAP neural network compared to linear discriminant analysis prediction of the length of hospital stay in patients with pneumonia. *1992 IEEE International Conference on Systems, Man, and Cybernetics*, Vol. I, pp. 748–53.

Greene, W. H. (1993). *Econometric analysis*, 2nd edn. New York: Macmillan.

Hertz, J. Krogh, A., & Palmer, R. G. (1991). *Introduction to the theory of neural computation.* Reading, MA: Addison-Wesley.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, **2**, 359–366.

Jessup, L. M. & Valacich, J. S. (1993). *Group support systems.* New York: Macmillan.

Kim, E., Wu, J., Tamura, S., Sato, Y., Close, R., Taketani, H., Kawai, H., Inoue, M., & Ono, K. (1993). Comparison of neural network and k-NN classification methods in vowel and patellar subluxation image recognitions. In I. Guyon & P. S. P. Wang (Eds), *Advances in pattern recognition systems using neural network technologies.* World Scientific.

Kolodner, J. L. (1993). *Cased-based reasoning.* San Mateo, CA:

Morgan Kaufmann Publishers.

Mak, B. & Blanning, B. (1995). Sensitivity analysis in neural net models. Hong Kong University of Science and Technology, Working Paper.

Mak, B. & Bui, T. (1995). Modeling experts consensual judgments for new product entry timing. *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 26, Number 9, 1996.

Marquez, L., Hill, T., Worthley, R., & Remus, W. (1991). Neural network models as an alternative to regression. *Proceedings of the IEEE 24th Annual Hawaii International Conference on Systems Sciences*, Vol. VI, pp. 129–135.

Medsker, L. & Turban, E. (1994). Integrating expert systems and neural computing for decision support. *Expert Systems With Applications*, **7**, 553–562.

Messier, W. F. & Hansen, J. (1988). Inducing rules for expert system development: An example using default and bankruptcy data. *Management Science*, **34**, 1403–1415.

Odom, M. D. & Sharda, R. (1990). A neural network model for bankruptcy prediction. *Proceedings of IEEE International Conference on Neural Networks*, pp. II163–II168, San Diego.

Pao, Y. (1989). *Adaptive pattern recognition and neural networks.* Reading, MA: Addison Wesley.

Patuwo, E., Hu, M., Hung, H., & Ming, S. (1993). Two-group classification using neural networks. *Decision Sciences*, **24**, 825–845.

Porter, M. (1980). *Competitive strategy.* The Free Press.

Raghupathi, W., Schkade, L. L., & Raju, B. S. (1991). A neural network application for bankruptcy prediction. *Proceedings of the IEEE 24th Annual Hawaii International Conference on Systems Sciences*, pp. 147–155.

Rumelhart, D., McClelland, J., & the PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition, Vol I: Foundations*, p. 329. MIT.

Salchenberger, L. M., Mine, C. E., & Lash, N. A. (1992). Neural networks: A new tool for predicting thrift failures. *Decision Sciences*, **23**, 899–916.

Simon, H. (1960). *The new science of management decision.* Harper & Row.

Simon, H. & Ericsson K. A. (1986). *Protocol analysis.*. MIT Press.

Srikanth, R. (1994). Constructing conceptual models for knowledge capture and delivery systems. *Journal of Organizational Computing*, **4**(2), 199–217.

Subramanian, V., Hung, M. S., & Hu, M. Y. (1993). An experimental evaluation of neural networks for classification. *Computers & Operations Research*, **20**, 769–782.

Swales, G. & Yoon, Y. (1992). Applying artificial neural networks to investment analysis. *Financial Analysts Journal*, **48**(5), 78–80.

Tam, K. Y. & Kiang, M. (1990). Predicting bank failures: A neural network approach. *Applied Artificial Intelligence*, **4**, 265–282.

Tam, K. Y. & Kiang, M. (1992). Managerial applications of neural networks: The case of bank failure predictions. *Management Science*, **38**, 926–947.

Tavana, M., Lee, P., & Joglekar, P. (1994). The development and validation of a campus recruiting expert system using expert opinions and historical data. *Expert Systems With Applications*, **7**, 305–321.

Taylor, W. A., Weimann, D. H., & Martin, P. J. (1995). Knowledge acquisition and synthesis in a multiple source multiple domain process context. *Expert Systems With Applications*, **8**, 295–302.

Urban, G. L. & Hauser, J. R. (1993). *Design and marketing of new products.* 2nd edn. Englewood Cliffs, NJ: Prentice-Hall.

Walker, W. & Devaney, M. (1993). Enhancing performance. *Pensions & Investments*, **21**(9) 34–35.

Watkin, T. L. H., Rau, A., & Biehl, M. (1993). The statistical mechanics of learning a rule. *Reviews of Modern Physics*, **65**(2).

Wielinga, B., Van de Velde, W., Schreiber, G., & Akkermans, H. (1993). Towards a unification of knowledge modelling approaches. In J.-M. David, J.-P. Krivine, & R. Simmons, *Second generation expert systems.* New York: Springer.

Wong, B., Bodnovich, T., & Selvi, Y (1995). A bibliography of neural network business applications research: 1988–September 1994. *Proceedings of The Third International Conference on Decision Support Systems*, Hong Kong, Vol. 1, p. 1.

Yeung, D. Y. (1993). Constructive neural networks as estimators of Bayesian discriminant functions. *Pattern Recognition*, **26**, 189–204.

Yoon, Y., Guimaraes, T., & Swales, G. (1994). Integrating artificial neural networks with rule-based expert systems. *Decision Support Systems*, **11**, 497–507.

Yoon, Y., Swales, G., & Margavio, T. M. (1993). A comparison of discriminant analysis versus artificial neural networks. *Journal of the Operational Research Society*, **44**, 51–60.