Selecting The Appropriate Artificial Neural Network To Minimize Audit Costs When Assessing The Financial Viability Of Audit Clients

Harlan Etheridge, (E-mail: Harlan@louisiana.edu), University of Louisiana at Lafayette
H. Y. Kathy Hsu, University of Louisiana at Lafayette

INTRODUCTION

As part of an audit engagement, auditors must assess the financial viability of the audit client (SAS Nos. 58, 59, 96). If the auditor thinks that the client is financially healthy, no action is required.

However, if the auditor believes that the audit client will fail within a year from the financial statement date, then the auditor must include explanatory language stating that belief as part of the auditor's report.

The assessment of an audit client's continuing financial viability is important to an auditor because (1) it is required by generally accepted auditing standards, and (2) a mistake is making this assessment can result in severe financial consequences to the auditor. If an auditor mistakenly believes that the audit client will fail within a year and issues an audit report stating that belief, that action may, in fact, contribute to the firm's failure. Creditors may refuse to extend credit terms to the firm and sources of capital may dry up. This situation is known as the "self-fulfilling prophecy" and may result in lawsuits against the auditor if the client or its shareholders believe that the auditor's report contributed to the client's financial demise. On the other hand, if an auditor fails to issue an audit report with the appropriate explanatory language for an audit client that does fail within the next fiscal year, then the auditor may be sued by investors and creditors who were financially harmed by the client's failure.

On many audits, the state of the client's financial condition is readily discernable—the client is obviously financially healthy or financially distressed. However, sometimes the client's financial condition is not readily apparent. In these cases, the auditor should use additional analytical techniques in an attempt to determine whether the client will fail within the next fiscal year (SAS No. 56). If the auditor believes that the client is seriously financially distressed, then, due to the gravity of the situation, the auditor also may wish to use additional techniques in an attempt to determine with a higher degree of certainty whether or not the firm will fail by the next financial statement date.

Statistical techniques, such as multivariate discriminant analysis (MDA) and logistic regression (logit), are proven methods of developing models to determine whether or not a firm will fail within a year. However, these techniques suffer from several limitations that restrict their usefulness in real-world situations. Artificial neural networks (ANNs), on the other hand, do not have these limitations and also typically outperform MDA and logit in classifying firms as financially healthy or failing.

Artificial neural networks, a type of artificial intelligence, are patterned after the human brain. They are composed of numerous simple processing elements (analogous to neurons in the brain) that are highly connected. ANNs also learn like the human brain—from experience.

Auditors may mistakenly believe that all ANNs are "created equal." However, several types of ANNs exist, each designed to accomplish a certain type of task. For example, some ANNs are best used to categorize observations into discrete groups, e.g., financially healthy or financially distressed, while others are designed to forecast a value given a set of data, e.g., forecast a firm's stock price given a set of financial and economic data.

Intuitively, the ANNs designed to categorize observations into separate groups would seem well suited for assessing the financial viability of an audit client. However, many ANN software packages do not offer these types of ANNs, and instead focus on ANNs designed for forecasting, typically backpropagation or one of its variants. What would happen if an auditor uses an ANN not designed for categorization, such as backpropagation, to assess client financial viability? That is the question that we address in this article.

AUDIT COSTS

Auditors should attempt to design their audits to minimize their overall cost. Audit cost consists of many factors, some obvious and some not so obvious. Obvious costs include the cost of personnel and the cost of computer software and hardware. Not so obvious costs include the loss of an audit client, harm to the auditor's reputation, and increased exposure to litigation due to audit failure. Audit failure occurs when any significant part of the audit process does not achieve its intended purpose. Therefore, failing to correctly assess the financial viability of an audit client can be considered an audit failure and may greatly increase the cost of an audit.

An auditor can minimize the audit costs associated with evaluation of a client's financial viability by using a method to assess financial viability that minimizes misclassification costs. Misclassification costs of a financial viability assessment method are determined by two factors: (1) the probability of making misclassifications using a specific classification method, and (2) the cost of making a misclassification error.

Auditors can make two types of misclassifications when evaluating the financial viability of an audit client: (1) classifying a firm that will fail within the next fiscal year as financially healthy, and (2) classifying a firm that is financially healthy as one that will fail within the next fiscal year. To simplify further discussion of these errors, we will refer to them as (1) Type I errors, and (2) Type II errors, respectively.

The probability of Type I and Type II errors occurring during an audit depends on the method that is used to assess client financial viability. For example, a model developed using MDA will have different Type I and Type II error rates than a model developed using logit. If an auditor wishes to minimize the overall error rate (Type I and Type II errors combined), then the following equation can be used to calculate the estimated overall error rates of financial viability models developed using different methods:

ESTIMATED OVERALL ERROR RATE = [TYPE I ERROR RATE × .02] + [TYPE II ERROR RATE × 0.98].

The reason that the Type I error rate is multiplied by .02 is that, on average, 2% of firms fail every year. Consequently, the Type II error rate is multiplied by .98 because, on average, 98% of firms do not fail. Of course, an auditor can adjust the probabilities to reflect industry and general economic conditions. Adding the products of these calculations together gives an estimate of the overall error rate that can be expected from a specific financial viability model.

Using the estimated overall error rate to select a model to use in assessing client financial viability poses a problem. The cost of a Type I error and the cost of a Type II error are not the same. Most auditors think that Type I errors (categorizing a firm that will fail within a year as financially healthy) are more costly than Type II errors (categorizing a healthy firm as one that will fail within a year). Consequently, auditors need to consider the relative costs of Type I and Type II errors when selecting a financial viability model to use in assessing financial viability.

We can use the following equation to estimate the misclassification cost associated with a financial viability model:

MISCLASSIFICATION COST = (PROBABILITY OF TYPE I ERROR × COST OF ERROR) + (PROBABILITY OF TYPE II ERROR × COST OF ERROR).

Determining the average dollar-cost a Type I or Type II error can be extremely difficult, if not impossible, because certain costs are difficult to quantify, e.g., the negative impact on an auditor's reputation due to an audit failure. However, we can express the costs of the errors relative to each other by using a ratio:

COST OF TYPE I ERROR/COST OF TYPE II ERROR

Since the relative cost of a Type I error compared to that of a Type II error cannot be determined with precision and probably varies across audit situations, we can vary the relative cost to examine how the misclassification cost of a financial viability model behaves in response to changes in the Type I/Type II error cost ratio. Since the cost of a Type I error generally is thought to be higher than the cost of a Type II error, we should increase the cost of a Type I error relative to that of a Type II error in the cost ratios. For example, we can vary the cost ratios from 1:1 to 50:1 to see how the overall cost of misclassification of a specific financial viability model behaves as the cost ratios change.

Using relative cost ratios also allow us to directly calculate and compare the estimated relative costs of financial viability models using the following equation:

ESTIMATED RELATIVE COST (RC) = (PI X CI) + (PII X CII).

Where PI is the probability of a Type I error, CI is the relative cost of a Type I error, PII is the probability of a Type II error, and CII is the relative cost of a Type II error. Comparing the relative costs of financial viability models as a method of determining the desirability of the models is superior to simply comparing the model error rates, because the relative cost calculation explicitly considers the costs of Type I errors relative to that of Type II errors. Consequently, if an auditor chooses the model with the lowest RC, then the costs of using a financial viability model should be minimized.

ARTIFICIAL NEURAL NETWORKS (ANNS)

ANNs are types of computer hardware or software that attempt to mimic, in a highly simplified manner, the physical structure of the human brain. Thus, they are composed of a large number of highly connected processing elements (artificial neurons) that usually are organized into layers (see figure 1). They "learn" through experience, the same way that humans do.

Development of a financial viability ANN usually consists of two phases: (1) the training phase and (2) the testing phase. During the training phase, the ANN is presented with a large number of examples (observations) of financially healthy and failed firms. Each firm observation should contain relevant financial (and possibly nonfinancial) data about the firm and a digit that indicates the firm's state of financial health (usually 0 and 1, to indicate financial health and financial failure). As the ANN processes the data, it "learns" the relationships between the data items and the firm's financial health. These relationships are learned for each firm observation in the training data set. Consequently, the ANN learns that certain data patterns are associated with financially healthy firms and other data patterns are associated with failed firms. If these relationships change with time, the ANN can learn the new relationships by processing new data.

During the testing phase of ANN development, the ANN is presented with firm observations that were not used during the training phase. If the ANN effectively learned the relationships between firm information and financial health, then the ANN should have few misclassifications. If the misclassification rates are too high, then the ANN did not adequately learn the relationships, and the ANN must be retrained or scrapped.

TYPES OF ARTIFICIAL NEURAL NETWORKS

The actual types of ANNs are too numerous to mention in this article. However, we think that three types of ANNs may be useful to auditors in assessing client financial viability: (1) categorical learning neural networks (CLN), (2) probabilistic neural networks (PNN), and (3) backpropagation neural networks (BPN). Each of these ANNs represents a different approach to problem solving. We do not know which type of ANN will minimize misclassification costs, but suspect that CLN and PNN will have lower misclassification costs than BPN because CLN and PNN are designed to categorize observations into separate groups, while BPN is not. However, BPN is a very popular and powerful ANN that is included in most ANN software packages, and we examine its ability to categorize firms' financial viability for that reason.

Categorical Learning Neural Networks (CLNs)

CLNs learn to distinguish between financially healthy and financially failing firms by using competition between processing elements. During the training phase of ANN development, the CLN is presented with the financial data patterns of hundreds or thousands of firms. Each processing element in the middle layer of the CLN learns the data pattern of one firm only and learns to associate that data pattern (and similar data patterns) with either financial health or financial failure, depending on whether the data pattern it learns comes from a financially healthy or failed firm. For example, if we construct a CLN to distinguish between financially healthy firms and those that will fail within a year, then each processing element will "compete" will all other processing elements to learn the current firm's financial data pattern; there can be only one "winner" of each competition. So after the first firm's data pattern is learned by one processing element, then the remaining processing elements "compete" to learn the remaining data patterns in the training data set. This cycle continues until all of the firm's data patterns have been learned. Each data pattern also includes a code, usually a 0 or 1, to indicate whether that data pattern belongs to a financially healthy or financially failed firm.

Once training of the CLN is completed, when the CLN is presented with the data pattern of a new firm, it categorizes the firm either as financially healthy or financially failing by determining to which of the previous learned data patterns the current firm's data pattern is closest. If the closest previously learned data pattern is associated with financial health, then the current firm is classified as financially healthy. Otherwise, it is categorized as financially failing.

Probabilistic Neural Networks (PNNs)

PNNs use probability theory as the basis for categorizing firms either as financially healthy or financially failing. A PNN is trained using a training data set in a manner similar to that of a CLN. Thus, during the training phase of ANN development, each processing element in the middle layer of the PNN learns the data pattern of either a financially healthy firm or a financially failing firm. Then, once training is completed, when the PNN is presented with the data pattern of a new firm, it classifies the firm either a financially healthy or financially distressed by considering (1) to which of the previously learned data patterns is the current data pattern closest, and (2) what is the probability that the current firm is financially healthy (e.g., .98) or financially failing (e.g., .02).

What happens if the data pattern of a new firm is "close" to the data patterns of both a financially healthy firm and a financially failing firm? Then the PNN would use the probability of failure (.02) and nonfailure (.98) to categorize the firm. Since the probability of nonfailure is much higher than that of failure, we should expect to see a substantial number of Type I errors (incorrectly classifying a financially failing firm as a financially healthy firm) compared to the number of Type II errors (incorrectly classifying a financially healthy firm as a financially failing firm) when using a PNN to evaluate client financial viability. Increasing the probability of failure will result in a lower Type I error rate; however, the Type II error rate will increase more rapidly than the Type I error decreases since 98% of most firms do not fail.

Backpropagation Neural Networks (BPNs)

BPNs use what is called a "gradient-descent learning law" to minimize the mean squared error (MSE) in the ANN. When the MSE is minimized, then the errors made by the BPN should be minimized as well. During the training phase of ANN development, the gradient-descent learning law attempts to assign the responsibility for errors to the processing elements in the BPN. Processing elements with a greater responsibility for errors are modified by the ANN to a greater extent than those processing elements with less responsibility for errors. Training is completed when the MSE cannot be reduced further. At this point, the BPN is optimized to predict the item of interest, e.g., financial health of a firm. Then when the BPN is shown the data pattern of a new firm, the BPN outputs a value, usually between 0 and 1, which the auditor must interpret as meaning either "financially healthy" or "financially failing." In order to make this decision, the auditor must establish a cutoff point that separates the two categories of firms. For example, if the BPN is trained to represent financially healthy firms with a zero and financially failing firms with a one, then the auditor may wish to use a cutoff point of .5—an output from the BPN of .5 or less would be interpreted to mean "financially healthy" while an output greater than .5 would be interpreted to mean "financially failing." Modifying this cutoff point will result in different Type I and Type II error rates and different misclassification costs for the same BPN.

THE STUDY

In order to investigate the impact of different types of ANNs on classification accuracy, we develop a CLN, PNN and BPN to categorize firms as either financially healthy or financially failing. The sample of firms in our study was supplied by KPMG Peat Marwick and is composed of 1139 banks (991 healthy and 148 failed) in various regions of the U.S. Our sample contains 57 financial variables for each firm for the years 1986 to 1988, a time period with a high rate of bank failures. We use the FDIC definition of failure to operationalize failed banks as assisted mergers and liquidated banks (FDIC Statistics on Banking, 1992). The FDIC assumed the operations of the failed banks in 1989, so we use the years 1986, 1987, and 1988 to represent three years, two years, and one year prior to failure.

Each ANN is trained with a subsample of the original data set and then tested using a holdout sample, which consists of 192 healthy and 23 failed banks for each of the three years prior to failure. The error rates for each ANN during testing are presented in table 1.

The estimated overall error rate (EOER) is low for all three ANNs. However, BPN and PNN have lower EOERs (ranging from 2.4% one year before failure to 6.57% three years before failure) than CLN. Consequently, if EOER is used to determine the desirability of a financial viability model, then both BPN and PNN would appeal to auditors. However, overall error rates are not sufficient to determine the adequacy of financial viability models because the EOERs do not incorporate the rates of misclassifying failed and nonfailed banks (Type I and Type II errors). Therefore, we also compare the ANNs on the basis of their Type I and Type II error rates from the testing phase.

Both BPN and PNN perform better than CLN in categorizing nonfailed banks (see table 1). The Type II error rate for BPN is lower than 4% one year before failure and lower than 6% three years before failure. For PNN, the Type II error rate is lower than 2% one year before failure and lower than 4% three years before failure. However, neither BPN nor PNN correctly classify failed banks as well as CLN. CLN has Type I error rates ranging from 0% to 22% one to three years before failure, while BPN and PNN have Type I error rates ranging from 13% to 52% one to three years before failure.

Because auditors should use the ANN financial viability model that minimizes their audit costs, we also calculate the estimated relative costs (RCs) for each of the ANNs using costs ratios ranging from 1:1 to 50:1. The performance rankings (1, 2, or 3, with 1 representing the lowest relative cost and 3 representing the highest) of the three ANNs for the various cost ratios are presented in table 2. To compare the performance of the ANNs, we compute a simple sum of the ranks (rank-sums) of the models for each of the three years prior to failure and for different RC ratios. Because the costs of Type I errors are believed to be greater than the costs of Type II errors, we exclude the rankings for the RC of 1:1 from the calculations of the rank-sums.

The rank-sum measure is 18 for CLN, 31 for BPN, and 41 for PNN. A lower rank-sum indicates lower relative costs associated with misclassification errors. Consequently, auditors should expect CLN to minimize the costs of misclassification, followed by BPN and PNN.

We also test whether the differences in the model RCs are statistically significant (not due to chance). Table 3 presents these results. Panel A lists the yearly means of the ANN RCs across cost ratios, while Panel B presents the one-tailed t-tests and p-values by year across model means.

The t-tests support the results of the simple rank-sum scores. The year immediately prior to failure, CLN has the lowest average relative cost, followed by BPN and PNN. Also, BPN has a lower average relative cost than PNN one year before failure. Two years prior to failure, CLN has a lower average relative cost than either BPN or PNN; however, the relative cost of BPN is not significantly different from that of PNN. Three years prior to failure, the relative costs of all of the ANN models are approximately the same (even though CLN appears to have a lower relative cost than the other two ANN models, the differences are not statistically significant).

CONCLUSIONS

Auditors wishing to minimize their audit costs must consider the type of financial viability model to use when assessing whether or not a client can be considered a going concern. ANN financial viability models perform as well or better than traditional statistical models, e.g., multiple discriminant analysis and logit. However, auditors must choose the appropriate type of ANN to use when evaluating client financial viability because different types of ANNs have different expected error rates and, thus, different costs of use. Choosing an appropriate ANN to use when evaluating client financial viability can help auditors minimize audit costs.

Expected error rates alone cannot be use to judge the desirability of financial viability models. Since Type I errors (categorizing a financially failing firm as financially healthy) generally are considered to be more costly than Type II errors (categorizing a financially healthy firm as failing), the relative costs of these types of errors must be considered when evaluating model desirability.

When relative cost ratios are low, the probabilistic neural network (PNN) and backpropagation ANN (BPN) perform well in our tests. As relative cost ratios increase, the categorical learning network (CLN) has a lower expected relative cost than those of either BPN or PNN. Based on the results of our tests, auditors should choose CLN over either BPN or PNN to minimize the costs associated with an incorrect assessment of a client's financial health.

However, BPN is believed by many to be an appropriate ANN method to use in evaluating client financial viability, given its widespread use in financial distress studies. Although it is more suited for solving forecasting problems, it appears to have performed reasonably well in classifying failed and nonfailed banks. Auditors choosing to use a BPN to evaluate client financial viability, however, should be aware that better alternatives exist.

REFERENCES

1.   FDIC Statistics on Banking. The Federal Deposit Insurance Corporation, Division of Research and Statistics, Washington, D.C., 1992.
2.   Statement of Auditing Standard 56: *Analytical Procedures.* Auditing Standards Board. 1988.
3.   Statement of Auditing Standard 58: *Reporting on Audited Financial Statements.* Auditing Standards Board. 1988.
4.   Statement of Auditing Standard 59: *The Auditor's Consideration of An Entity's Ability to Continue As a Going Concern.* Auditing Standards Board. 1988.
5.   Statement of Auditing Standard 96: *Audit Documentation.* Auditing Standards Board. 2001.

Table 1
Estimated Error Rates (Percentages)

ANN Model	Year	Type I Errors Fail as Nonfail	Type II Errors Nonfail as Fail	Overall Error Rate
BPN	1988	13.04	3.65	3.83
	1987	43.48	4.69	5.46
	1986	47.83	5.73	6.57
CLN	1988	0.00	7.29	7.15
	1987	17.39	7.81	8.00
	1986	21.74	11.98	12.17
PNN	1988	43.48	1.56	2.40
	1987	47.83	3.13	4.02
	1986	52.17	3.13	4.11

Table 2
ANNs Ranked by Estimated Relative Cost

ANN Model	Cost Ratio	1988	1987	1986	Total of Ranks
BPN	1:1	2	2	2	
	10:1	1	3	2	
	20:1	2	2	3	
	30:1	2	2	2	
	40:1	2	2	2	
	50:1	2	2	2	**31**
CLN	1:1	3	3	3	
	10:1	2	1	3	
	20:1	1	1	1	
	30:1	1	1	1	
	40:1	1	1	1	
	50:1	1	1	1	**18**
PNN	1:1	1	1	1	
	10:1	3	2	1	
	20:1	3	3	2	
	30:1	3	3	3	
	40:1	3	3	3	
	50:1	3	3	3	**41**

Table 3
Annual RC Means

Panel A

Model	1988	1987	1986
BPN	0.1140	0.3068	0.3431
CLN	0.0715	0.1809	0.2478
PNN	0.2762	0.3176	0.3437

t-statistics of Differences Between Means

Panel B

Year	Model	CLN	PNN
1988	BPN	-2.3057	-2.5266
		(0.0412)	**(0.0264)**
	CLN		-3.3295
			(0.0146)
1987	BPN	-1.9011	-0.1179
		(0.0579)	(0.4545)
	CLN		-1.8990
			(0.0580)
1986	BPN	-1.2823	-0.0056
		(0.1235)	(0.4978)
	CLN		-1.1989
			(0.1421)

Figure 1
Typical Artificial Neural Network Architecture

Output Layer
(provides the ANN results)

Hidden Layer
(# of PEs determined
by ANN developer)

Input Layer
(1 PE/data item)

Input Data