A BETTER WAY TO PREDICT CREDIT RISK

By Albert Fensterstock

Introduction

Accurate prediction of the future is the ultimate goal of credit risk analysis. How much risk should you take on an account? Will they pay on a timely basis? How can you determine when to pull the plug and not sell them anymore? These are some of the questions a credit manager is expected to have answers to. And now there is a better way to help answer these questions using a revolutionary technique that relies on more than one model to solve these problems, a so called ensemble model. This technique has far better accuracy than traditional statistical-based analysis and judgment-based models are left in the dust.

What is an Ensemble Model?

An ensemble model combines two or more models to enable a more accurate prediction of credit risk. If you have a current statistical model that you like it can become part of the ensemble, but its accuracy will be significantly enhanced by the other models in the ensemble. The goal of the ensemble is to predict which accounts will go BAD and which will not. Where BAD is your definition of BAD (for example, an account might be considered BAD if more than 20% of its total monthly outstanding balance goes 91 or more days past due).

Building an Ensemble Model

Like most scientific-based models an ensemble model, for your company, will be created from your historical data. Anything you know about your accounts is applicable. The table below shows some of the possible data that the model might use:

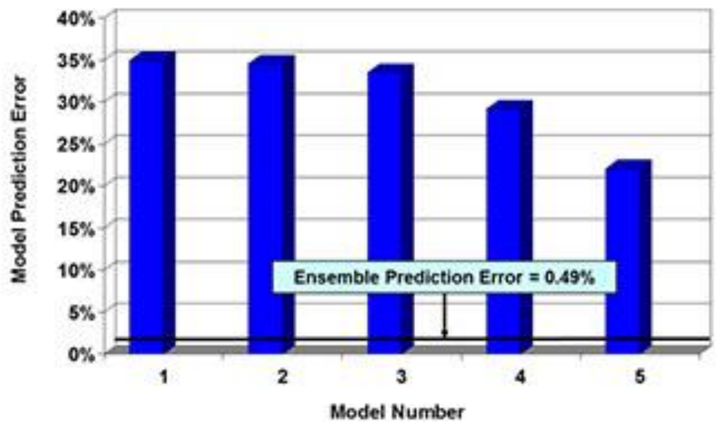
REPRESENTATIVE MODEL DATA AND ITS SOURCE			
INTERNAL	COMMERCIAL BUREAU	CONSUMER BUREAU	FINANCIAL STATEMENTS
Account Tenure	Various Bureau Predictive Indicators – Paydex; CCS; FSS; Intelliscore, etc.	Various Bureau Predictive Indicators – FICO, etc.	Leverage
Collection Effort			Working Capital
Credit Balance		Age of Newest Trade	Net Liquid Balance
Current Aging	Company History	Average Trade Balance	Net Worth
Date of Last Payment	Industry/Geography	Charge-Offs	Solvency Ratios
Historical Aging	Negative Payment	Collection Inquiries	Cash Position
Late Fees	Experiences	Credit Limit	Profit Returns
NSF Checks	Previous Bankruptcy	Current Balance	Industry Norm Information
Days Beyond Terms	Secured Financing	Delinquent Trade Lines	Total Liabilities
Payment Amounts	Size of Company	Inquiries	Gross Profit Margin
Write-Off Amounts	Suits/Liens/Judgments	Public Records	
Application Date	UCC Filings	Time On File	
Application Decision	Years in Business	Total Trades	
Funding Date	Trade Interchange Data		

The first step in the model building process is to determine which variables are the most predictive and should be considered for use in the final model. The off-the-shelf software used in this application has the capability to help you do this. In a given application you might start off with 100 possible variables and wind up with a fraction of that which are used in the final model.

Once the predictive variables are determined, multiple models will be developed with different variables receiving a variety of weights. Again, the software has the ability to vary the models and help you decide which ones to use in the ensemble. You might create 25 models, but use only five of them in the ensemble.

After the models to use in the ensemble are determined the weight to give each model’s prediction can be decided. A well-known statistical concept is used to determine the weights. Each of the individual models’ output (0 for BAD or 1 for GOOD) is multiplied by its weight where the sum of the weights equal 1. The weighted results of each model are added together to arrive at the ensemble’s prediction. Any output equal to 0.5 or less is considered BAD and an output greater than 0.5 is considered GOOD.

The graph below shows the individual predictiveness of five different models and the resultant predictiveness when they are combined into an ensemble.



In this application, the final model's prediction error rate was 0.49%, (i.e., the model misclassified less than 0.5% of the accounts that would go BAD) while the error rates for the five models in the ensemble ranged from 22.1% to 35.0%. In other words, the prediction error rates of the individual models were from 45 to 71 times higher than the ensemble's prediction error rate. This supports our statement that an ensemble model can produce a prediction error rate that is superior to the prediction error rate of any individual model in the ensemble.

Using the Ensemble to Determine a Credit Line

Individual account evaluation is performed in real-time. This allows a great deal of flexibility in the decisioning process, in particular with respect to determining a credit line. The model's output is either a GOOD or BAD decision, but all GOODs and all BADs are not created equal.

Model Output is GOOD: In any model that evaluates a request for credit, the amount requested is an input variable. Because the model's decision is in real-time you can replace the original amount requested with another amount and, instantaneously, get a new model evaluation. In the case of a GOOD decision you can increment the requested amount until the model's decision changes from GOOD to BAD. This is the maximum amount of risk (credit line) the model feels is applicable for a given account. If the model indicates a higher line than the account requested it might allow the sales department to generate additional revenue.

Model Output is BAD: In this instance, you want to determine if there is any amount of risk the model will accept. Companies don't make money by refusing to sell an account. Here, you'll decrement the amount requested to see if the model's decision will change from BAD to GOOD. If it does, this amount represents the amount of risk the model deems acceptable, and provides the possibility of profiting from an otherwise unacceptable account.

Using the Ensemble as an Early Warning System

Once a credit line is determined, it is important to continue to track the account to be sure that their circumstances do not change to the extent that their current risk of non-payment is not significantly greater than it was at the time the credit line was granted. Therefore, the model should be run against all of the open accounts, on a periodic basis (weekly or monthly), to determine whether an account's circumstances have changed. If so, remedial action can be taken immediately, and will very likely significantly reduce the number of accounts turned over for collection.

CONCLUSION

Ensemble models represent a very powerful method for determining credit risk. And, because they can be built on a PC based platform, at a reasonable cost, monitored and improved in-house without the need for outside consultants, it makes them a very attractive alternative to other methods of credit evaluation, and in particular to judgment-based models. In review, these type of models can aid in providing an easy to implement solution to the associated problems of credit evaluation, credit line determination and ongoing account tracking.

Albert Fensterstock is managing director of Albert Fensterstock Associates. He has more than 40 years' experience in financial and operations management and analysis. He specializes in the development and implementation of decision support systems based on various types of predictive analytics including statistical analysis, neural networks and simulation technology. His solutions are frequently used for improving risk analysis capability and collection department efficiency. He can be reached at 516-313-1020 or via e-mail at albie@afassociates.org