

## Fighting Fraud Through Predictive Analytics

### MBA (3/10/2006) McAfee, Jamie

To assist lenders in preventing fraud—which has been likened to finding a needle in a haystack—many have implemented fraud products that rely on data verification. A new product from **BasePoint Analytics**, Carlsbad, Calif., called **FraudMark** uses neural network technology that combines two techniques for finding fraud: behavior history for applicants, brokers, appraisers, loan officers and other entities and historical patterns of both fraudulent and non-fraudulent loan applications.

The basis of neural technology is that uses machine-learning techniques to understand what is going on in data. “It looks at good loans and bad loans and differentiates between the two,” said **Tim Grace**, president of BasePoint Analytics. “It uses mathematical and statistical techniques to make a statistically sound assessment of whether it looks more like a good or more like a bad.”

FraudMark uses several different techniques, one of which is a supervised model. This model uses good and bad tags or markers. “It learns the characteristics contained in the loan applications marked as bad and it learns what the characteristics contained in the applications or loans marked as good. The new application information is evaluated against this criteria and a determination is made of which one it looks most like,” Grace explained. “Given we know the bad characteristics, the model is trained to tell us which ones are bad on the new applications.”

The other models used are called unsupervised models. These examine changes in behavior. “In FraudMark the application model is our supervised model so it looks at application data and tries to determine whether the information provided is more like a good application or more like a bad,” Grace said. “The unsupervised models also influence the score by alerting the model to changes in behavior.”

An example of an unsupervised model is one that monitors brokers. “Historical broker information is stored and then on an on-going basis, we look at things like average loan amount that they have submitted, or average number of loans that they have submitted to a particular lender,” Grace said. “If there is a change in that going forward, that influences the model. It recognizes there is a change in behavior on one of the entities and that change influences the score. The entities that unsupervised models utilize include things like brokers, appraisers or loan officers. Each of those are called in through the unsupervised models and are influencers into the score.”

FraudMark also enables the lender to score loan applications multiple times during the loan origination process as additional information is accumulated. “Once a loan is submitted; all of the application information such as application information, lender specific information such as the broker information, the appraiser information and all

other information considered by the FraudMark model, an assessment is made of whether or not the loan has possible fraudulent activity involved in it," Grace said. "It gives it a score from 1 to 999, with 999 being the riskiest. It also gives risk indicators of what some of the patterns were that caused the score to be higher. These things could be high income amount for age bracket or risky geographic area combined with high income to age bracket."

There can be five to 15 different risk indicators that accompany a score. "What happens at the risk manager level is that a score threshold is set, which accounts for a given amount of loan volume they are going to review," Grace said. "Let's say they set the score at 840 and above and want to review all the loans that score 840 and above. That usually corresponds to a 5 percent review rate and that corresponds to a detection level of the model. What differentiates what BasePoint does is that we provide a performance report for the model so lenders can choose their risk management goals. They can choose to detect 40 percent of their fraud, which corresponds to a 5 percent loan volume. If a lender wants to detect 50 percent of their fraud, they will review 10 percent of my loans and that's what the score represents."

Compared to database verification tools, predictive analytics could reduce the amount of loan reviewed by risk managers, Grace said. "With a database validation tool you typically have very high false positives sometime between 200- to 300-to-one, meaning that you will have to review 200 to 300 different loan files to detect one fraud," he noted. "With our solution, we are able to command false positives of 10- to 20-to-one. You have to review far less loans in order to detect a fraud."