



The Bottom Line on Data Mining

There are many reasons given current market conditions for lenders to do a better job of mining their own data.

IT'S HARD TO IMAGINE MANY INDUSTRIES HAVING ACCESS to the kinds of data that mortgage lenders have concerning their customers. By the time the loan has been processed, lenders know more about their borrowers than just about anyone – with the possible exception of federal law enforcement.

So what are we doing with all of this information as an industry? Not much.

Even though data mining has been a hot topic in the mortgage lending business for a number of years, few lenders dig into their databases in search of elusive information. Typically, those lenders that mine their customer information do so for cross-selling purposes and to isolate existing customers who will either need a new mortgage loan or a second mortgage in the near future.

When information technology professionals – those highly educated folks working on the cutting edge of data processing – talk about data mining, they mean to look into databases in an effort to identify trends within data that

go beyond simple analysis and can be used to target business opportunities. Typically, sophisticated algorithms are employed to sift through large quantities of information in search of trend data. Unlike simple queries, these analytical tools look for multiple relationships in the data, often finding connections that were not anticipated.

Generally, data mining is the process of analyzing data and identifying correlations or patterns. The knowledge discovered can then be used to predict behavior that can be encouraged to increase revenue, cut costs, avoid losses, and to make operational decisions.

The process of data mining can be separated into two functions. The first is knowledge discovery, wherein we try to learn something from the massive collection of data elements. The second is using that new knowledge to make a prediction that will point to a business opportunity.

A simple example is given below. A lender sorts through existing information to determine what kinds of borrowers also opted for a second mortgage (knowledge discovery) and then uses the resulting profile to predict which other customers may do the same in the near future. Both of these functions can become very complicated as more powerful analytics are used to mine the data and make the forecasts.

Before data mining can be performed effectively, the data must be prepared or stored to ensure that the various records all contain the expected information in the proper format. This is more important in industries that do not have tight controls or standards over the methods by which data is gathered. In our



business, we are standardized by document type, with the 1003, 1004, GFE and HUD-1 to keep us all on the same page. Furthermore, industry efforts like MISMO are providing XML schemas that may one day ensure every lender uses the same field for the same data in the same format.

The benefit that has been most successful in getting lenders to investigate data mining has been its application to cross-selling. Share-of-wallet, the percentage you control of all the financial services business your client does, is an important industry metric, especially for banks and other diversified financial services industries. Monoline mortgage bankers benefit by selling additional mortgage products to their existing customers.

However, nationwide lenders should leverage the information in their databases for the benefit of all of their lending partners. If you depend upon third-party originators and have such strong lending partnerships, you want to be able to look for relationships in the data that might increase the sales of the types of products your lending partners sell to you. It's important to know how differences in borrower profiles, geographic location and loan program influences the business that comes in from your sources and to use that knowledge to benefit everyone.

It is key to look for connections and commonalities and also comb through data for anomalies that might indicate problems such as fraud, or opportunities such as new techniques or best practices that can be leveraged by other sources in other locations.

Typical results can include new approaches to mortgage planning for certain borrower profiles, suggestions for new product innovation and, of course, additional opportunities to deepen relationships with mortgage customers through cross-selling.

In order to take full advantage of data mining, lenders must provide an internal group with access to the

database (or some subset or archive) and then set a deadline for reporting. It is not necessary to set expectations for the kinds of information to be returned. In fact, it can be counterproductive to direct the group toward a certain result. The process works best when the analysts are free to return any relationships found in the data as long as the deadline is in place.

Many executives are uncomfortable with the ambiguity built into the aforementioned process, but, the results can justify an excursion from your comfort zone. The good news is that the hurdles to effective data mining are largely attitudinal barriers. The technology is not an obstacle.

// The real power available through data mining is the ability to analyze both internal factors ... and customer demographics. //

In the past, only organizations with large IT budgets and very powerful computer hardware could use the sophisticated analytical software required for thorough data mining. Lenders have been satisfied with simple database queries, which don't require much hardware or software.

Typically, for data-mining applications less than 50 gigabytes, relational database storage and management technology is sufficient. Lenders can go much further with their data-mining efforts without having to make additional investments in technology. Usually the current hardware and software you use to originate and process mortgage loans will suffice.

Lenders should be aware that there have been some privacy concerns raised in regard to data mining. These typically apply to applications that involve data sources outside the enterprise. Since lenders deal with public records on a daily basis, most companies are familiar with the limitations

imposed by Gramm-Leach-Bliley. As long as the data-mining effort is performed on existing customers and sensitive customer information is not shared over public networks, we have not found the need to take any special precautions in regards to customer privacy.

The real power available through data mining is the ability to analyze both internal factors such as program offerings, pricing and originator skills, with external factors, such as economic indicators, competition and customer demographics. To take advantage of this, lenders have to think beyond the cross-sell benefit. Lenders will need to think about discovering broader,

more macro knowledge that will help them with predicting the next surge in mortgage lending.

In the future, I suspect the industry will make better use of all the financial data available due to the loan origination process, to explore additional factors that will lead to customer satisfaction, bottom line profitability and overall partner success. **MT**

Paul Decoff is the chief lending officer of Thornburg Mortgage, a single-family residential mortgage lender focused principally on the jumbo segment of the adjustable-rate mortgage market. In this role he has oversight over information technology, operations development, financial engineering and organizational management in mortgage lending. With more than 15 years of experience in the mortgage industry, Mr. Decoff joined Thornburg Mortgage after 13 years at Countrywide Financial Corp., where he most recently served as a managing director and chief operations officer at Countrywide Bank NA, Countrywide's banking subsidiary.