

Detecting A Pattern

Insurers can use predictive modeling to segment risks, reveal new market potential and maximize profits.

by Roosevelt Mosley

Home mailboxes are filled with solicitation offers for a plethora of products—credit cards, mortgage refinances, even automobile or homeowners insurance. Grocery stores provide coupons that print out at the checkout lane based on specific purchases. These special offers stem from data-collection methods that service and retail companies use to learn consumers' particular preferences so the companies can maximize potential profitability.

Like other industries, property/casualty insurance companies are discovering the power of predictive modeling to reveal new market potential, improve underwriting choices and, ultimately, boost profitability. Predictive modeling helps better identify and segment insurance risks so executives no longer have to manage on instinct or "gut feel." Instead, they can use factual data to assist in making better business decisions. In contrast, insurers not applying predictive modeling techniques are missing competitive opportunities.

Predictive Modeling Defined

Predictive modeling is a form of data mining. Data mining is the analysis of historical data to find relationships between elements in the

Illustration by Angel Negron

Contributor Roosevelt Mosley is a senior consultant of Pinnacle Actuarial Resources Inc., based in Bloomington, Ill.

Key Points

- Predictive modeling can help insurers improve their rating plans by identifying mispriced risks.
- New information on market segments or an insured's relationships to the insurance process that provide a new perspective on risk may be revealed by using predictive modeling.
- Predictive modeling ultimately can boost profitability.

datasets. Modelers analyze specific subsets of data and then use the results to make predictions.

Predictive modeling often reveals new information regarding market segments or an insured's relationships to the insurance process that provide a new perspective on risk. By having more detailed information, executives can make more informed business decisions.

Specifically, predictive modeling can help insurers improve their rating plans by identifying mispriced risks. By analyzing distributional relationships in insurance databases in a multivariate framework, predictive modeling will reveal assumptions that have been generating mispriced risks. For example, when considering the relationship between insurance losses and age and the relationship between insurance losses and prior accidents, it is no surprise that younger drivers tend to cost more to insure and that drivers with prior accidents cost more to insure. These findings, however, don't account for the fact that a larger proportion of younger drivers tend to have prior accidents. Knowing this could make a

difference in the way an insurer chooses to surcharge younger drivers for prior accidents.

Predictive modeling helps insurers define groups that are more homogeneous for rating, underwriting and marketing. An insurance company may rate the city of Dallas and the areas surrounding the city the same. Predictive modeling may show that the risk of loss outside the city of Dallas is considerably different from the risk of loss inside the city. A company that successfully separates the Dallas area into more homogenous risk areas will gain a significant competitive advantage.

Further, by discovering new risk identifiers or new relationships between current risk identifiers, predictive modeling also can identify new ways to segment risks. Using credit history for generating insurance scores is a vivid example of predictive modeling. Insurance scores have been used not only for rating, but also for tiering, underwriting and marketing.

Step by Step

Successful predictive modeling begins with identifying the specific questions that need to be answered. After finding the appropriate data, begin mining the data and developing models to help understand it better. Finally, take the knowledge gained from predictive modeling and apply it to the insurance function.

Clearly defining the modeling project's purpose is critical. Otherwise, trying to do too much can become overwhelming. Among the first applications of predictive modeling has been to better analyze rating and tiering of insurance risks. Historically, the establishment of rating factor relativities for insurance has been based on one-way loss ratios or pure premiums. The problem with this approach was illustrated in the example of prior claim history and youthful drivers. There are many distributional biases in a dataset that cause the one-way approach to produce incorrect results. Using multivariate predictive modeling methods will account for the distributional overlap and correlation between risk factors and ensure

that the rating and tiering factors being used account properly for differences in risk.

Customer response modeling holds a great deal of potential for increasing profitability as well. Generally, the pricing of insurance is focused on the supply side of the economic equation, with little emphasis placed on the demand for insurance or the willingness of different market segments to insure at different prices. Given a set of risk characteristics, customer response modeling looks at items such as the likelihood of policyholder renewal and the likelihood of writing a new business policy. Understanding that the probabilities of renewal and new business conversion are going to differ according to the characteristics of the risk can help a company boost profitability.

Claims department functions can benefit from several applications, including estimation of claim-settlement value. Claims that are settled by insurance companies have characteristics associated with them, including claimant information, types of injuries and attorney involvement. A model can be developed from historical closed claims that estimates the value of the claim based on its characteristics. This model then can be applied to new claims to estimate their ultimate settlement value.

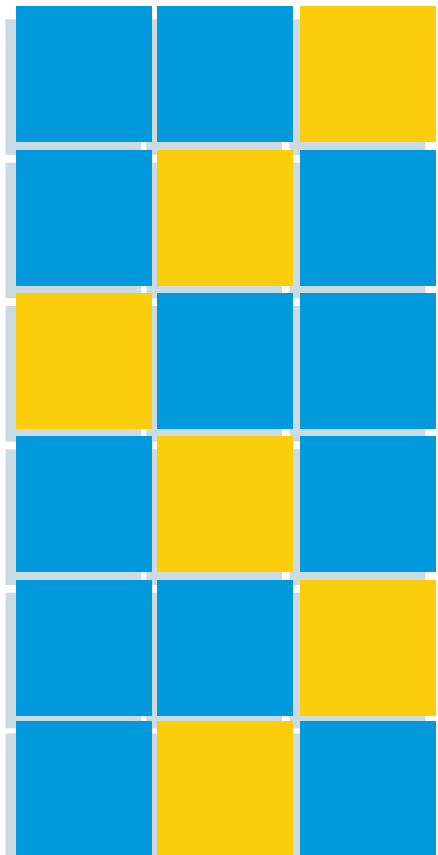
Predictive modeling comes in many other forms, including vehicle classification, fraud detection, agency evaluation and loss-reserve development factor modeling. Regardless of the approach being chosen, it will always be important to stay focused. Carefully defining and carrying out one of these applications will better prepare a company to handle the next application.

Getting the Right Data

Insurers have access to a wealth of information, internal and external, that can be used in predictive modeling. Data sources include traditional internal sources such as rating and underwriting; nontraditional internal sources, including agency, marketing and billing information; and external sources such as credit and allowable

demographic information. Regardless of where the data come from, the quality of the data is critical. The "garbage-in, garbage-out" rule certainly applies here. After determining what data are needed, extract, verify and cleanse the data as necessary. While it is important to be sure that the data being collected are relevant to the task at hand, be careful not to exclude potentially valuable information. Many times assumptions are disproved by the actual data, and dismissing data before modeling may undermine the process.

After identifying and extracting the data, ensure that they appear reasonable. One approach is to summarize the major statistics, such as premium, exposure, claim counts and claim amounts, by the independent variables. Doing so tests the reasonableness of the distributions of the independent variables. For example, there would be reason for concern if 95% of the auto drivers were coded as males. This process also forces the modeler to understand the levels of the independent variables being reviewed, which is



Technology

invaluable information once it is time to interpret the model's results.

Developing the Model

Several different types of models can be fitted to the data. The appropriate model will depend on the structure of the data and the application being developed. "Different Models for Different Data" below describes various models and their uses. While model types may vary, the predictive modeling process will be similar for whichever one is selected. It uses historical experience to attempt to predict future outcomes.

Before actually generating model results, hold back a random portion of the dataset for purposes of testing and validating the model once it has been developed. The size of the holdback will vary depending on the size of the dataset being used, but this holdback will help prevent overfitting the model to the data. To perform this validation, take the model developed on the largest portion of the data and apply it to the holdback dataset. To the extent that the results are significantly different, there could be an over-fitting problem.

Interpreting and Applying Results

People who understand the process being modeled and may see the data from differing vantage points

should be involved in interpreting the results. If modeling a rating and tiering plan, include actuarial, claims, underwriting, marketing and senior management professionals to interpret and apply results. Many perspectives can help in applying professional judgment to model results and coming up with a final product that is both powerful and practical.

This diverse team should consider many factors when applying modeled results to the real world. For starters, policyholder impacts can be a significant hurdle to implementing predictive modeling results. If an insurance company traditionally has not used these insurance pricing techniques, the modeling results can indicate large rate changes or disruptions that may make a company uncomfortable. Many times, predictive model results suggest that insurers should make significant changes in the way they do business. Computer systems may not be able to handle some changes. Potential systems impacts should be considered, and at times it may be necessary to adjust the application of the model results so they fit within the framework of what is possible in the current situation. This also can identify potential systems and infrastructure changes needed for the future.

Corporate culture also can be a potential hindrance to the application

of predictive modeling results. Many times, predictive models confirm that the assumptions a company is making about a risk are correct. In some instances, however, models go against conventional wisdom. The difficult choice of whether to follow the data or follow the "way we've always done things" will need to be made.

Lastly, public and regulatory acceptance must be considered. Just because data say the insurance industry should do something does not mean the public or the regulatory community will accept it. Therefore, explaining and implementing results should be done with caution. While the regulatory community generally has accepted predictive models, the results they have generated have not always been embraced. Educating and clearly communicating to regulators and legislators can help ease these concerns.

Effective predictive modeling can and does enhance underwriting, pricing and marketing decisions and boost insurers' profitability. As companies continue to take advantage of predictive modeling applications, find new rating variables and sources of data, and apply the results in new and innovative ways, it will likely become a way of life for all successful companies, much as it is in other industries, such as banking. **BR**

Different Models for Different Data

Generalized Linear Modeling—Enables a series of independent variables to predict the value of a dependent variable. This model is especially effective for determining the impact of class plan variables on loss costs or the impacts of different claim characteristics on an ultimate claim settlement value.

GLM also provides a framework for discovering the interactions of variables in an automated way. Interactions occur when two independent variables in a model do not have a constant relationship with each other.

Decision Tree Analysis—Attempts to separate a group of risks into homogeneous groups based on an identified response variable. The process begins by taking the entire population, and then analyzes each independent variable to determine which creates the largest degree of separation in the dependent variable. The dataset is then "split," or

branched off, into two or more groups based on this characteristic. Next, each branch is analyzed independently to determine which independent characteristic is most important in distinguishing between levels of the dependent variable for that branch.

Logistic Regression—Determines responses to certain situations. Logistic regression generally attempts to model questions with a "yes" or "no" answer. Examples include: "Does the policyholder renew?" or "Does the applicant quoted actually purchase a policy?" Based on a set of independent characteristics, logistic models determine the likelihood of obtaining a positive response.

Neural Networks—Attempt to model human responses to a set of stimuli.

Regression Splines—Build on multiple regression models by making the model structure more flexible.

Insurance is a multi-billion-dollar industry.
Who knows where the money goes?

You will.

The Best's Review[®] **2005 Guide To
*Understanding The Insurance Industry***

This new reference is packed with charts and graphs that explain—in a colorful, concise and easy-to-understand way—how insurers make and spend money.

The *Guide* showcases exclusive information about:

- The property/casualty, life, health and reinsurance industries
- How insurance is sold and regulated
- Solvency and ratings
- Background on well-known insurers
- Plus: Industry overview, calendar of meetings and events, and glossary of insurance terms

The *Guide* is an excellent source for industry facts and figures—perfect for reports, speeches and presentations—and a great gift for:

- New employees
- Board members
- Customers
- Anyone who wants to gain a better understanding of how the insurance industry works

The *Guide* is yours with a one-year subscription to *Best's Review* magazine. For only \$20—half the regular subscription price—you'll get 12 issues of *Best's Review* plus the *Best's Review 2005 Guide to Understanding the Insurance Industry* as a free gift.



Reserve your copy of the *Guide* today by calling (908) 439-2200, ext. 5742, or visiting www.bestreview.com/subscribe. Once you receive your *Guide*, you'll have the opportunity to purchase up to five additional copies for \$10 each by filling out the order form at the back of the book.