Predictive Analytics In Property and Casualty Insurance: Real Value or More Empty Promises

In today's economic climate where budget reductions are common, executives are under pressure to deliver profitable growth. Business leaders must identify and implement the critical items that will enable the enterprise to remain competitive. Methods and techniques long utilized in the actuarial department are being leveraged to improve many areas of the insurance operation. Will these methods really deliver value, or are they just empty rhetoric?

About the Author

John Wilt
John heads property and casualty consulting for TCS' North Global Consulting Practice in North America. John's background includes more than 28 years of industry and consulting experience as a leader and hands-on implementer of sophisticated and transformational insurance solutions. He has held executive positions in property and casualty, and life and annuity insurance companies, as well as leadership positions within global consulting, system implementation, integration and outsourcing firms. John has served as a technology and business advisor for leading international and domestic financial corporations. John's prior experience includes building and managing his own consulting firm, as well as Senior Executive, CIO and consulting positions with 25 of the top 100 insurance companies. John holds a Master in Business Administration from Case Western Weatherhead School of Management, and a Bachelor of Science in computer science from Cleveland State University.

Table of Contents

Introduction

In current economic climate where budget reductions are common, executives are under pressure to deliver profitable growth. Business leaders must identify and implement the critical items that will enable the enterprise to remain competitive. Methods and techniques long utilized in the actuarial department are being leveraged to improve many areas of the insurance operation. Will these methods really deliver value, or are they just empty rhetoric? Today's soft market and other economic challenges make it more critical than ever that insurance company leaders maximize every dollar invested. Carriers, facing myriad challenges, are seeking opportunities to improve operating efficiencies and gain a better understanding of distribution channels, customer behavior, product and pricing efficiencies, risk selection and earlier recognition of problems before the impact is felt.

Despite the challenges, the current environment also creates the opportunity for insurers to outrun their competitors by improving enterprise agility and optimizing operating costs. The winners at the end of this economic cycle will be those who have taken the time to make the necessary improvements to quickly capitalize on opportunities provided when the cycle turns up.

Having the ability to accurately forecast performance in activities as diverse as operations, budgets, supplies, product demand and performance are crucial for business success. However, developing a better understanding of existing processes, customer behavior, identifying opportunities that may have been unexpected and anticipating problems before they happen, are also critical to the success of any company.[2] Predictive analytics can provide critical insight into these and other areas. It is effortless to see why using predictive analytics well is imperative to insurance organizations, which are particularly reliant on predicting future activities. An insurer's ability to forecast a policy's ultimate cost determines how accurately it prices its product and, in turn, the extent to which it can avoid adverse selection. In the fight for market share going on today, accurate pricing based on policy performance is one of the critical areas.[2]

It might be said that insurance carriers have always relied on forecasting, and that would be a correct statement. Initially, insurers simply guessed at appropriate premiums. Subsequently, in most cases led by the actuarial team, they determined premiums by analyzing a single factor, such as the age of an insured building or the piloting history of an insured ship's captain, both examples of *univariate analysis*. As insurance operations became more technologically advanced, multiple factors such as the age of the insured building, its type of construction, its usage, and so forth were used to determine an appropriate premium, *multivariate analysis*.

Insurers use techniques known as predictive analytics to determine many critical items. Internal and additional external data - information such as credit scores or local economic conditions that may be relevant or correlated with a potential insurance outcome are used to deliver key elements of an insurance solution. The use of predictive analytics has quickly become an insurance industry best practice. Insurers use predictive analytic techniques to target potential clients, identify the way they approach an insurance purchase, determine more relevant products for a specific market as well as more accurate product pricing and to proactively identify potentially fraudulent claims. This whitepaper will provide an overview of predictive analytics and present some drivers of its growth, its uses in the property and casualty insurance industry and the advantages for insurers who use it.

Overview of Predictive Analytics

There are several perspectives on predictive analytics but most will agree that it is a broad term describing a variety of statistical and analytical techniques used to develop models that predict future events or behaviors. The form of these predictive models varies, depending on the behavior or event that they are endeavoring to predict. Most predictive models generate a score (a credit score for example), with a higher score indicating a higher likelihood of the given behavior or event occurring. Others see predictive analytics as an mysterious set of techniques and technologies that puzzle many business and IT managers. Used by special groups with informal processes and producing questionable results often do not correspond with other enterprise information sources and reports.

These days, predictive analytics represents the blending of statistics, advanced mathematics, analytics tools, business intelligence with artificial intelligence and a heavy dose of data strategy and management to create a "potent brew that many would rather not drink!" Most recently, a series of technology breakthroughs in an area known as machine learning have added another element to this predictive analytics brew.[8]

Some are unsure if predictive analytics is a legitimate business endeavor or an ivory tower science experiment run wild. However, most insurance companies will agree that predictive analytics is a set of Business Intelligence (BI) technologies that uncover relationships and patterns within large volumes of data that can be used to predict behavior and events. But unlike other BI technologies, predictive analytics is forward-looking, using past events to anticipate the future and the demand for it is growing and becoming an imperative for success.[2]

Predictive Models: Credit Score

The most prevalent and widely understood examples of predictive models are those used by the three credit bureaus (Experian, Equifax, and TransUnion) to develop credit scores for individuals. Each credit bureau uses a variety of information about an individual (income, credit history, outstanding loan balances, and so forth) to develop a credit score that predicts the likelihood that he or she will repay current and future debts. The higher the credit score, the more likely the individual is to pay his/her debt.[2]

Data mining is a component of predictive analytics that entails analysis of data to identify trends, patterns, or relationships among the data. This information can then be used to develop a predictive model. Predictive analytics, along with most predictive models and data mining techniques, rely on increasingly sophisticated statistical methods, including multivariate analysis techniques such as advanced regression or time-series models. Such techniques enable organizations to determine trends and relationships that may not be readily apparent, but still enable it to better predict future events or behaviors. A critical element in leveraging predictive analytics as a strategic advantage is finding new attributes, factors or patterns that will provide greater insight or accuracy in forecasting future events.[7]
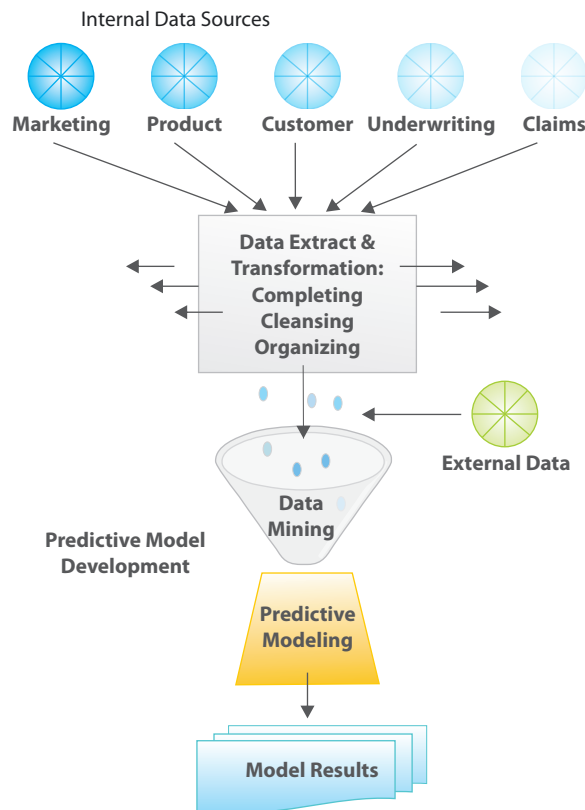
The Predictive Analytics Process

When using predictive analytics, an insurer begins with data. Data is an asset crucial to the success of a modern insurance company. Property and casualty insurers can reduce uncertainty caused by a lack of available information by transforming internal data and tapping into external sources. Not only is data the key to predictive analytics, but also a key driver to improving key processes and products. Data management techniques that include elements of governance, quality, reliability and credibility are required to convert the collection of legacy data sources into a useful asset. If and when this is completed the insurer can initiate work with new and existing analytics capabilities to identify and gather the attributes desired.

Data Transformation

If the desired attributes exist within the enterprise, the data sources must be identified. Often data sources will have missing, incomplete or even conflicting data and the "source of record" must be identified. When that is finished, the insurer must begin aggregating and "cleansing" its internal data for use in the analytics models. Cleansing entails scouring records to identify those with missing or incomplete data. Records with missing or incomplete data can have an impact on the accuracy of the predictive model. These records must be completed and/or corrected so that the ultimate predictive model is as accurate as possible. For example, an insurer developing a predictive model for auto insurance claims would start with its own marketing, underwriting, and claims records for auto policies it has sold. Some of the records may have missing information, for instance the age, sex, or marital status of the insured, or may not contain the complete details of the claim (not noting whether a police report was filed or subrogation efforts were made). For some insurers, determining this missing information may involve a long and costly process, multiple legacy systems that may have accumulated through an insurer's mergers and acquisitions add to the difficulty of converting the data to a usable format. Once the data have been aggregated and cleansed, sound statistical practices dictate that they be divided into an in-sample group that will be used to develop the predictive model and an out-of-sample group that will be used to test the model. External data sources can also be identified and incorporated, if required.

The Predictive Analytics Process

Internal Data Sources

Marketing   Product   Customer   Underwriting   Claims

Data Extract &
Transformation:
Completing
Cleansing
Organizing

External Data

Data
Mining

Predictive Model
Development

Predictive
Modeling

Model Results

In addition to utilizing best practices in data management across the enterprise to treat and realize data as the critical corporate asset it is, sophisticated tools may also be used to infuse clarity and integrity into the every phase of the data management process. Some of these include enterprise data models, metadata repositories, data dictionaries and data and test mining tools.[7]

Data Mining

Data mining is the analysis of data to identify underlying trends, patterns, or relationships. It is a necessary first step in predictive analytics, because the data that the mining process identifies as relevant can then be used to develop the predictive model. One can think of data mining as gathering knowledge about relationships, and the resulting predictive analytics model as applying that knowledge. One distinct advantage to data mining is that it catalogs all relationships (or correlations) that may be found among the data, regardless of what causes that relationship. For example, data mining may discern a relationship between age and gray hair, or age and number of auto accidents, but it does not imply that age causes gray hair or auto accidents.[7]

Model Development

Predictive models can vary greatly and are found in many shapes and sizes, depending on their complexity and the target purpose for which they are designed. This section introduces some of the statistical methods that may be used to develop a predictive model. Data mining can be

thought of as obtaining the information, and modeling as applying it. Unlike data mining, many of the statistical procedures that are employed in predictive models search for one specific relationship. This may require, for example, specifying during model development that age does cause gray hair or that age may reduce the likelihood of auto accidents (at least up to certain ages).[2]

Regression Basics

Ideally, to explain relationships between variables (such as age and losses in auto), insurers would examine the entire population (in this case, the entire population of drivers). However, insurers can only draw data from a sample of the population (such as only the drivers they insure), so they must do the following: [1] Build the best model they can to determine the "true" relationship between variables,[2] Analyze confidence in the model (mathematical description of its accuracy). Predictive models may use a variety of regression models. The most basic regression models are linear regression models such as ordinary least squares (OLS) regression. Some of the most complex regression models are multivariate adaptive regression splines called MARSplines.[10] There are other models applicable to the insurance business. Some are listed below:

- Linear Regression
- Regression Splines
- Advanced Models

What Drives Insurers' Use of Predictive Analytics

Though insurance businesses have employed predictive analytics techniques for many years, the application and prevalence has recently increased throughout the industry. The primary drivers pushing this growth include the following:

- Drive for cost reduction
- Desire for growth in soft market, slow-growth conditions
- Search for competitive advantage

The growing number of uses for predictive analytics within property and casualty insurance make the use of analytics seem more and more promising to ever larger segments of the insurance carrier population. From fighting fraud to modeling future human resource needs, to risk selection and policies pricing; the benefits of well planned and used analytics demonstrate value. There are a few that seem to be more critical in the dogfight for market share that is ongoing currently that we will address. The following discussion is limited to three applications of predictive analytics on which insurance carriers focus today.

Marketing

Insurance marketing has often relied on referrals and other traditional marketing approaches. Predictive modeling utilized in insurance marketing represents a revolutionary approach to what has commonly been perceived as a relationship-based business. Predictive analytics is used in the marketing of many products and services. Financial services organizations use predictive analytics to identify potential customers for mortgages, annuities, loans, and investments. Property-casualty insurers can use predictive analytics to analyze the purchasing patterns of insurance customers. This information can be used to increase the marketing function's hit ratio and retention. As technology moves to eliminate the barriers between agent and direct channels, marketing

analytics will be a critical element for planning the differentiators relevant to each market segment. Making sure that investments for high-touch guidance have been made and communicated correctly will be a important message.[5]

Hit Ratio

Hit ratio is a measure of how often the marketing function generates a sale for each contact made with a potential customer. If an agent makes one sale for every ten potential clients, his or her hit ratio is one in ten (ten percent). Predictive analytics used to analyze purchasing patterns may allow carriers to focus on the correct channel: direct internet, phone or agent to focus on customers more likely to buy, thereby increasing his or her hit ratio. For example, if predictive analytics identifies the two customers out of every ten potential customers who are least likely to purchase a policy, the elimination of those potential customers from a channels sales agenda will raise the hit ratio to one in eight (12.5 percent).[5]

Underwriting and Pricing

Smart underwriting is receiving a bigger share of funding for analytics. The use of predictive analytics in underwriting is more evolutionary than revolutionary. Underwriting has always attempted to accurately predict future losses and price the products that protect against those losses. Predictive analytics represents the next generation of underwriting tools available to achieve those goals. Predictive models can be used in expert underwriting systems to remove the human error factor from the underwriting process by streamlining the "normal" underwriting cases and only referring the "exceptions" to the human underwriters. Insurers can use predictive analytics to filter out applicants who do not meet a pre-determined model score. This type of screening can greatly increase an insurer's efficiency by reducing the employee hours it may have spent researching and analyzing an applicant who ultimately is not a desired insured. If an applicant's model score is sufficient for consideration, then the model score can be used as a rating mechanism on which the insurer can base a variety of price/ product points.[4]
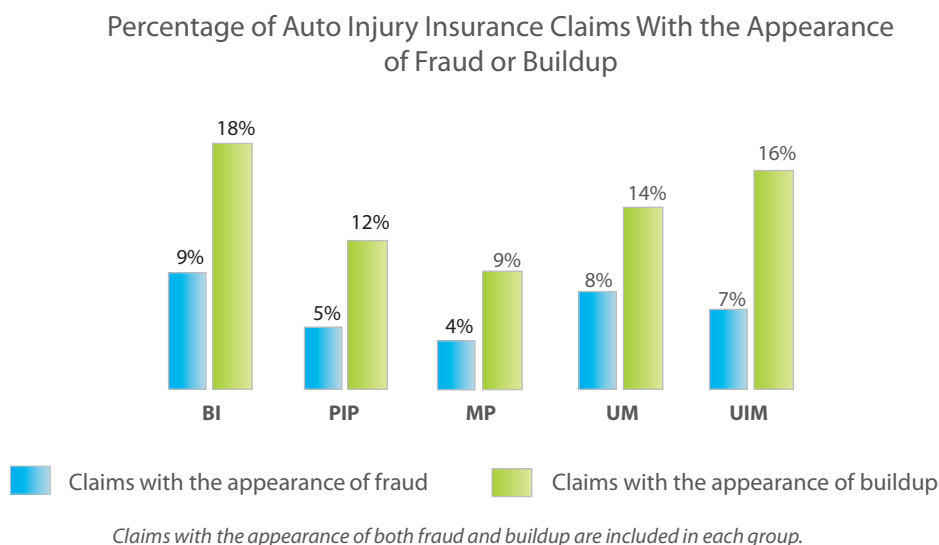
Claims

Property and Casualty insurance fraud was estimated to be a 1.3 billion dollar business in 2009. So why are insurance companies not doing more to detect and prevent fraud? Often the answer is "cost effectiveness". But progress in predictive analytics is changing the game. Predictive analytics is more of a revolutionary concept in claims handling than it is in underwriting. Insurers can use predictive analytics to help identify potentially fraudulent claims, segment them and identify the most likely candidates for action. It also can be used to score claims based on the likely size of the settlement, enabling an insurer to more efficiently allocate resources to higher priority claims.[6]

Identifying Potentially Fraudulent Claims

Insurers have struggled for years to develop methods to identify potentially fraudulent claims. The Insurance Information Institute estimates that insurance fraud costs property-casualty insurers over $30 billion annually. It is difficult to estimate the actual percentage of all claims that are fraudulent, but that is a staggering number no matter how you derive it. Fraud can take many forms, from staged accidents to the padding or building up of claims (inflating the value of the claim) for accidents that have already occurred. The Property-casualty insurers traditionally have

had difficulty identifying the relatively small number of fraudulent claims (the needle) made among the millions of claims filed every year (the haystack). Predictive analytics can help insurers more accurately determine claims that need additional review for fraud by increasing the likelihood of discovering fraudulent claims and helping it to refine the claims marked for review. This is known as limiting the occurrence of type I and type II errors. A type I error occurs when a legitimate claim is identified as possibly fraudulent. A type II error is the failure to identify a fraudulent claim and paying it as if it were legitimate, illustrated as follows:

Percentage of Auto Injury Insurance Claims With the Appearance of Fraud or Buildup

Claims with the appearance of fraud
Claims with the appearance of buildup

Claims with the appearance of both fraud and buildup are included in each group.

For a large North American Insurance Carrier, we utilized multivariate analysis predictive modeling to effectively detect fraudulent and classify claims where each claim had hundreds of attributes.[9]

Both types of errors can be costly. Identifying legitimate claims as fraudulent may anger policyholders and result in litigation or accusations of bad faith in claims practices, but failing to identify fraudulent claims results in higher claims costs and therefore higher premiums for all policy holders. Any tools that can aid in the accurate identification of fraudulent claims reduces both types of errors and significantly improves the claims process.

Fraud Reduction

TCS developed and delivered a proof of concept case study for a large insurance company in the US to demonstrate that statistical predictive modeling techniques can effectively detect fraudulent claims in personal automobile insurance.[9] Several dimensions of each claim, such as accident, damage, repairs, injury, treatment etc. were analyzed to detect "red flags" and effective classifiers were built to classify claims into suspicious / not suspicious. The study examined about 100,000 claims, each having hundreds of attributes. Work is now progressing towards building a large-scale claim fraud detection system.[2]

Predictive Analytics' Advantages for Insurers

If knowledge is power, then the advantages of predictive analytics are clear. Predictive analytic techniques allow insurers to better understand their data and how to use it to predict future events. Proper implementation of predictive analytic techniques can improve an insurer's consistency and efficiency in products development, help to define target markets and market selection, increasing the number of policy price points "rating", underwriting and claims processing, fraud detection and segmentation for proper attention and action;

- Helps marketing department more precisely identify potential policy sales through analysis of customer purchasing patterns
- Reduces the employee hours underwriters may have spent researching and analyzing an applicant who ultimately is not a desired insured
- Provides predictive modeling scores for applicants that can be used as a rating mechanism for determining a variety of policy price/product points
- Helps identify potentially fraudulent claims
- Scores claims based on the likely size of the settlement, enabling an insurer to more efficiently allocate resources to higher priority claims

Conclusion

Initially the question was asked, "Will these methods really drive value, or are they just empty rhetoric. It seems clear that predictive analytics, since the early days have had a positive impact on business. It is also clear that as data, methods, models and machine learning improve, based on maturity, the impact will increase exponentially. As the market leaders put forth the effort to transform data and apply analytics with increasing sophistication, the demand for predictive analytics will sweep the property and casualty insurance industry to new heights.

References

[1] Source: Author's calculations and Insurance Information Institution www.III.org

[2] Nyce, C. (2007). Predictive Analytics. American Institute for Chartered Property Casualty Underwriters/Insurance Institute of America, 2007, 1-14. Retrieved from www.aicpcu.org

[3] Mumtaz, A. (2009). Predictive Models Can Help Cut Costs Without Sacrificing Underwriting Control. National Underwriter/Property and Casualty, Technology Trends November 16, 2009, 36-39. Retrieved from www.property-casualty.com

[4] Thompson, K. B. and Ellingsworth, M. (2009). Using Analytics for Pricing Risk and managing Decisions. Florida Underwriter, Feature Story, August 2009, 21-24. Retrieved from www.floridaunderwriter.com

[5] Cusano, J. and Mahotra, R. (2010). Direct Sales, Agent Model Might Converge In PandC Distribution Future. National Underwriter/Property and Casualty, Agency Management, New Realities October 4, 2010, 30-31. Retrieved from www.property-casualty.com

[6] Wes, G. (2009). Fighting Fraud with Advanced Analytics. Canadian Underwriter September, 2009, 36-39. Retrieved from http://www.canadianunderwriter.ca

[7] Banfield, C. J. (2009). Data Quality Can Make Or Break Insurers. National Underwriter/Property and Casualty, Technology Trends June 8, 2009, 25, 31. Retrieved from www.property-casualty.com

[8] Leibensperger, M. (2010). Analytics 2.0. Canadian Underwriter, December 2010, 50-52. Retrieved from http://www.canadianunderwriter.ca

[9] Unknown (2010). Case Studies in Data Mining. TATA Consultancy Services, 20010, 1-2. Retrieved from https://knowmax.ultimatix.net

[10] Burns, R. (2009). Commercial Underwriters Following The Data Trail Blazed By Personal Lines. National Underwriter/Property and Casualty, Technology Trends June 8, 2009, 25, 31. Retrieved from www.property-casualty.com

TATA CONSULTANCY SERVICES

Experience certainty.

About TCS' Global Consulting Practice

TCS' Global Consulting Practice (GCP) is a key component in how TCS delivers additional value to clients. Using our collective industry insight, technology expertise, and consulting know-how, we partner with enterprises worldwide to deliver integrated end-to-end IT enabled business transformation services.

By tapping our worldwide pool of resources - onsite, offshore and nearshore, our high caliber consultants leverage solution accelerators and practice capabilities, balanced with our knowledge of local market demands, to enable enterprises to effectively meet their business goals.

GCP spearheads TCS' consulting capacity with consultants located in North America, UK, Europe, Asia Pacific, India, Ibero-America and Australia.

Contact

For more information about TCS' consulting services, email us at **global.consulting@tcs.com** or visit **www.tcs.com/consulting**

Subscribe to TCS White Papers

TCS.com RSS: http://www.tcs.com/rss_feeds/Pages/feed.aspx?f=w
Feedburner: http://feeds2.feedburner.com/tcswhitepapers

About Tata Consultancy Services (TCS)

Tata Consultancy Services is an IT services, consulting and business solutions organization that delivers real results to global business, ensuring a level of certainty no other firm can match. TCS offers a consulting-led, integrated portfolio of IT and IT-enabled, infrastructure, engineering and assurance services. This is delivered through its unique Global Network Delivery Model^{TM}, recognized as the benchmark of excellence in software development. A part of the Tata Group, India's largest industrial conglomerate, TCS has a global footprint and is listed on the National Stock Exchange and Bombay Stock Exchange in India.

For more information, visit us at **www.tcs.com**

IT Services
Business Solutions
Outsourcing

TCS Design Services | M | 0311