



Drive Your Business with Predictive Analytics

By Olivia Parr-Rud

Predictive Analytics Expert and Principal at OLIVIAGroup

Contents

The Nuts and Bolts of Predictive Analytics	2
Applications in Marketing	5
Applications in Risk	7
Applications in Operations	8
The Role of Leadership and Culture	9
References	10

Brought to you compliments of:



As global competition continues to shrink profit margins, companies are seeking ways to increase revenue and reduce expenses. At the same time, businesses are being deluged with data from every action, operation, and touch point. While this exponential increase in information is presenting challenges for many companies, others are leveraging it to drive their businesses to higher profits. During these turbulent times, predictive analytics is how smart companies are turning data into knowledge to gain a competitive advantage.

Businesses across a wide range of industries are beginning to realize what the large financial institutions have known for years: predictive analytics has the power to significantly improve the bottom line. From better targeting and risk assessment to streamlining operations and optimizing business decisions in all areas, predictive analytics is the next big step toward gaining and maintaining a competitive edge. So what is predictive analytics?

The Nuts and Bolts of Predictive Analytics

Predictive analytics is a blend of tools and techniques that enable organizations to identify patterns in data that can be used to make predictions of future outcomes. In business, predictive analytics typically take the form of predictive models that are used to drive better decision-making. They unveil and measure patterns to identify risks and opportunities using transactional, demographic, web-based, historical, text, sensor, economic, and unstructured data. These powerful models are able to consider multiple factors and predict outcomes with a high level of accuracy.

Why Use Predictive Analytics?

Advances in technology over the past two decades have led to a highly volatile, global economy that enables instant communication and connection worldwide. The result is increased competition through broader access to global markets, shrinking business cycles, and changing rules. In other words, there is no more “business as usual.”

The biggest shift has been in the power of consumers to influence their buying experience. With access to global products and services, they can demand better quality, lower prices, and faster delivery. As a result, companies are under pressure to achieve more with fewer resources. The challenge is for businesses to deliver the right offer, at the right time, at the right price, through the right channel.

The only way to survive in this new economy is to embrace and leverage the power of information. Knowledge gleaned from accurate, accessible, actionable information is essential for survival. Companies must now have up-to-the-minute information about their customers, operations, suppliers, competitors, and markets.

Fortunately, the same advances in technology that have empowered consumers have also enabled companies to obtain, organize, analyze, store, and retrieve huge amounts of information about their markets and customers and automate many marketing and operational tasks. Once these systems are in place, companies can leverage predictive analytics to gain a competitive edge. The first step is to understand how to connect the tools and techniques with your business goals.

What is Your Goal?

Predictive models can serve a variety of goals. But they should not drive the analysis. The best place to start is by asking what business problem you are trying to solve.

Attract new customers? Targeted response modeling on new customer acquisition campaigns will bring in more customers for the same marketing cost.

Avoid Risk? Risk of loss or claim models will identify customers or prospects that have a high likelihood of defaulting on a loan or filing an insurance or warranty claim. Businesses and governments protect their assets with models that predict security breaches.

Make unprofitable customers more profitable? Cross- and up-sell targeting models are used to increase profits from current customers.

Retain profitable customers? Retention/loyalty or attrition/churn models identify customers with a high likelihood of lowering or ceasing their current levels of activity. By identifying those customers before they leave, you can take action to retain them. It is often less expensive to retain customers than it is to win them back.

Reduce expenses? Better targeting through the use of models for new customer acquisition and customer relationship management will reduce expenses by improving the efficiency of your marketing efforts.

Identify fraud? Predictive models can unveil patterns that assist in identifying credit card, insurance, and health care fraud and money laundering operations.

Avoid process failure? Companies in a variety of industries are modeling data collected from production lines, customer feedback, hospital error, et cetera to predict breakdowns and take corrective actions.

Analyze health treatments? Predictive models in the life sciences are saving money and lives. Models that predict disease outbreaks, for example, are used to calculate inventory in high-risk geographic regions.

Sources and Types of Relevant Data for Maximum Insights

Historically, the primary use of predictive analytics was for scientific or medical research. This involved small data sets that were collected to prove a hypothesis. Over the past few decades, advances in computer power and server capacity have spurred the growth of data collection and usage. With the emergence of direct marketing, data about individuals and businesses started to be collected, combined, and aggregated from a variety of public and private sources, such as census statistics, phone directories, warranty cards, and public records. A long history of credit behavior was also available for limited use.

As the financial benefits of leveraging these sources became known, entire companies formed for the sole purpose of collecting, cleansing, aggregating, enhancing, and selling data about individuals and businesses. Today, the explosion of social media and mobile devices has unleashed enormous amounts of data, and each new technology, process, or activity is a potential source.

At 2.6 quintillion bytes of data being created per year 90% all the data in the world has been created in the past two years. Experts are calling this “Big Data.” Data earns its “bigness” in four ways:

- **Volume** – managing the collection and storage is the first challenge
- **Velocity** – leveraging sources and updates from technology with 24/7 usage
- **Variety** – combining a multitude of sources, types, and formats
- **Complexity** – interacting the first three requires advanced analytic processes

Tools created to leverage the new types of data from social media can analyze text to uncover patterns and make connections. Social network analysis is also being used to influence members of a social group. Unstructured data appears to be the next frontier of data sources. Powerful

software is now able to decipher and analyze handwritten text, audio files, photos, images, and videos. The translation of unstructured data and its integration with structured data are projected to be the primary data sources for analysis in the coming years.

Predictive data can be grouped into a few major categories.

- Personal data is based on individual characteristics, which includes demographics (such as age, income, dwelling type, and marital status) and behavioral (such as purchase behavior and click patterns), and psychographic (attitudes and opinions typically collected through surveys).
- Business data, or firmographics, includes both static and behavioral data such as annual revenue, number of employees, growth rate, industry, and number of sites. Web-savvy businesses are developing predictive models using online patterns of behavior to estimate preferences, industry changes, untapped markets, and future trends. Using estimates from these models, business can measure sentiment and activities that can inform the timing and messaging of marketing or other actions.
- Social network data includes size and density of a social group, its stability, and within-group similarities. Members of a social network are evaluated on characteristics such as connectedness, social role, and traffic role.
- Web behavior data is collected by analyzing the sequence and timing of a visitor to a website or series of websites. This is used to assist with search engine optimization and increase website “stickiness,” or the time a visitor lingers on a page or site.

Predictive Modeling Techniques

Over the past 20 years, advances in technology, especially computer processing power, have enabled the development of predictive models that are not only powerful but also able to leverage large volumes of data very quickly. Common predictive modeling techniques such as classification trees and linear and logistic regression leverage underlying statistical distributions to estimate future outcomes. New, more CPU-intensive techniques, such as neural networks, mimic the way a biological nervous system, such as the brain, processes information. In many cases, industry-specific techniques, such as rate setting in insurance and credit scoring in banking, are suitable to quickly solve specific business issues.

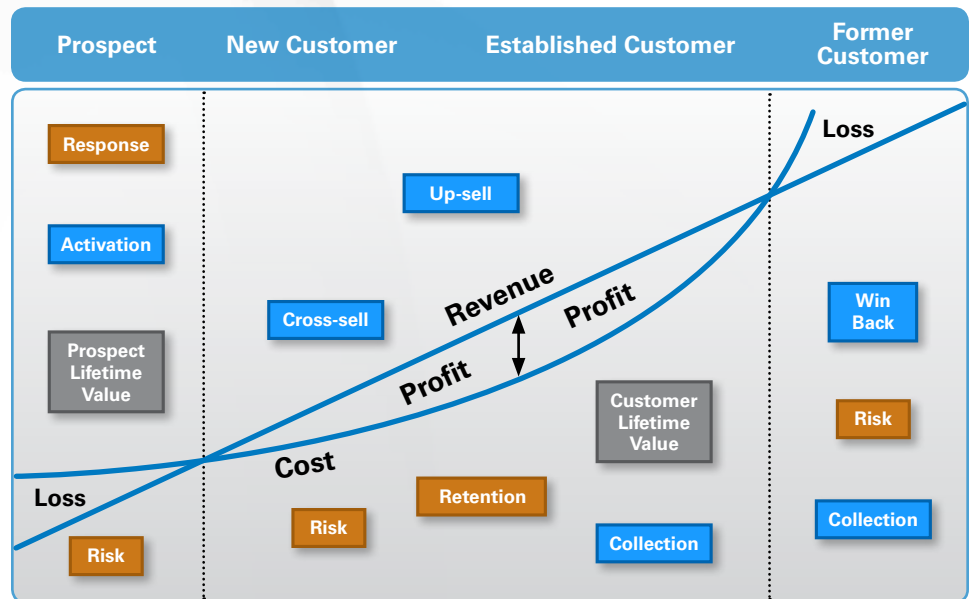
Model Deployment and Management

The standard output from a predictive model is a scoring algorithm (equation) or, in the case of a classification tree, a set of rules. Once the model is developed and validated, the final step is the deployment of the model scores or rules into a decision management platform that feeds key business processes and/or applications. Effective predictive model deployment requires a process that manages the inflow of data, the scoring, and the resulting actions, while monitoring data quality and model performance. For example, in a CRM application, customer response data is fed into a model deployment system from emails, call centers, or web activity. Each response is matched to an internal customer database and scored. Based on the resulting score or combination of scores, a next-best promotion or set of treatments is prescribed. For a predictive model to be useful in the long term, its performance needs to be monitored to determine whether the model has to be retired or updated or whether a new one is needed to reflect ongoing market changes.

Applications in Marketing

Response/Purchase

A majority of businesses earn profits through the sales of goods and services. Models that predict the response to a low-rate offer have been used by credit card banks for decades, and catalog companies have collected and pooled data to build response models for many years as well. With the recent explosion of online commerce, businesses are using online behavior and profile information to target offers.



Models to predict response and purchase have wide applicability to a variety of other industries.

- To take advantage of deregulation, gas and electric companies are using predictive models to increase their market share.
- Telecommunications entities such as cable, Internet, and phone companies use predictive models to attract and retain profitable customers based on usage and premium services.
- Gaming organizations such as casinos use highly sophisticated predictive models to identify, attract, and retain profitable patrons who enjoy gambling and star-quality entertainment.
- Nonprofits leverage predictive models to solicit donations through direct mail or telemarketing campaigns.
- Colleges and universities develop predictive models to target prospective students based on their likelihood to accept admission and graduate.
- Technology companies and business product manufacturers use predictive models to target companies that have a high likelihood of purchasing their products.

Response/purchase model scores are used to increase the number of responders/buyers, reduce expenses, or both. One national credit card bank saved tens of millions of dollars a year by

targeting prospects with a high likelihood to respond and carry a balance. By selecting prospects based on a predictive model, it reduced its direct mail by 50% while still capturing 85% of the profits.

Customer Segmentation

Classification trees or logistic models can be used to help companies understand their customer base by segmenting customers into groups based on measures that matter to the business. Three of the most common measures are response, revenue, and risk.

A large technology company was preparing to launch a campaign for a combination printer/fax/copier targeting a list of current business customers. The goal was to identify three distinct marketing segments based on a combination of model scores. The customer segment with the highest potential profit would be assigned a salesperson for a personal visit. The moderately profitable group would be mailed a promotion, with a phone call to follow. The least profitable group would receive an email with a link to the company's website.

To determine the groups, two models were developed. A logistic regression model estimated the likelihood that the customer would buy the printer. A linear regression model estimated the amount the customer would spend, given a purchase is made. The product of the two scores gave the customer an expected value. The company divided its customer list into groupings that allowed the marketing department to assign a strategy based on the customer's expected value.

The firmographic profile of the highest value customers was also used to identify new target markets for acquisition marketing.

Loyalty/Retention/Attrition/Churn

A familiar mantra in marketing is: It costs less to retain a customer than it does to replace one. For highly competitive industries where markets are close to saturation, this is particularly relevant. Industries that count on renewals, such as insurance, telecommunications, and publishing, pay close attention to retention and churn.

A basic predictive model with a binary outcome works well for predicting attrition or churn. However, a specific application of logistic regression, using time intervals as one of the input variables, can produce a model that estimates the probability of renewal or churn at future time intervals. For example, if you build a model on customers who cancelled their cellular service (totaled monthly) over the past three years, you can score current customers and estimate their likelihood to cancel in each month going forward. This allows you to prioritize your retention efforts to those most likely to churn in the next few months.

Social network analysis is used effectively to predict churn by examining the group behavior. By measuring the density of connection in a mobile phone network, cellular companies are able to predict patterns of churn and take proactive steps to retain their customers.

Customer Profitability

Some companies take a very comprehensive approach to customer management by employing a series of models to predict and manage customer profitability. There are many variations, depending on the industry, market, and product or service.

For example, given the competitive nature of interest rates in the credit card industry, one bank decided to examine its portfolio each quarter to determine the effect of changes in terms on customer profitability. The bank defined customer profitability as a combination of expected balances, payment behavior, and risk of default over the next three years. Predictive models were built that estimated the change in balance and payment behavior if either the interest rate or the credit limit was increased or decreased. A standard risk model was used. Based on quarterly scoring, the models estimated three-year profitability for each scenario based on the potential change in balances, payment behavior, and likelihood of default for each customer. These quarterly adjustments in terms for each customer allowed the bank to maximize rolling three-year profitability.

Applications in Risk

Default

A well-known application of predictive modeling is credit scoring. Based on credit history, financial profile, demographics, and other information, a lending institution or other business is able to determine the likelihood of an individual or business meeting a loan obligation. Predicting the likelihood of default or bankruptcy has been a main profit driver of the loan industry for many years. Whether for a mortgage, credit card, car loan, or even a utility service, many companies rely on a credit score to determine approval and terms of a contract. Banks also use aggregations of expected loss to meet regulatory requirements for loan reserves.

Claim

Insurance companies use predictive model scores to determine approval and optimize pricing for life, health, auto, and homeowners insurance based on demographics, claim history, and other risk factors. Risk pricing plays a very important role in auto insurance, where the rates are highly regulated. Warranty companies often use product information and safety records to model the likelihood of a claim.

Collections

Customers who fall behind in their payments and are sent to collections can be segmented much like those being considered for a marketing program. Combining the likelihood of default with the amount of the balance allows the collection department to determine the severity of the potential loss. In one example, customers with the highest expected loss were called immediately, while those with moderate risk were sent a letter and those with the lowest risk were sent an email reminder.

Fraud

With the increase in electronic and online transactions, fraud is increasing at an alarming rate. According to fraud research by the Association of Certified Fraud Examiners, fraud cost the global economy more than \$2.9 trillion in 2009. From stolen credit card purchases, insurance claims (including life, health, and auto), cellular usage, tax returns, online tracking, account/ACH fraud, and money laundering, fraud is increasing costs for businesses and consumers alike.

Powerful models are currently in use and in development to thwart these costly actions.

- Credit card banks use predictive models to identify the types of purchases that are typically made with a stolen card. For example, if the card is used to purchase expensive jewelry, furs, or firearms, the bank will freeze the card until the cardholder can verify the charges.
- Insurance claim fraud is also very costly to insurers and policyholders. Historically, every claim was personally inspected for fraud. Today, predictive models can estimate the likelihood that a claim is fraudulent based on common characteristics of past fraudulent claims. This allows insurers to reduce expenses by automating the approval of low-risk claims and placing a priority on claims with a high likelihood of fraud.
- Government institutions, including the IRS and social service agencies such as Medicare and the Social Security Administration, use predictive models to identify false tax returns or fraudulent requests for benefits.

Security

Models using social network data are being used by government agencies and private firms to detect security risk. Predictive models are being developed to predict the timing and location of breaches in security.

Process Failure

Process failures can be both costly and dangerous. Predictive models that can estimate the likelihood of a process failure at any given time by location not only save money, but also save lives. This is directly related to Asset Maintenance; see the next section for more detail.

Applications in Operations

Resource Planning

Models that predict marketing activities (response, purchase, and return) or activities specific to other industries are very useful for resource planning. Catalog and online retail companies use predictive models to manage inventory and optimize call center and distribution center resources. Airlines use models that predict air travel demand to optimize pricing, fuel purchases, and staffing. Companies in the hospitality and gaming industries rely on predictive models to optimize staffing by location and function.

Human Resources

The single greatest cost for a majority of businesses is employee compensation. Retail, travel, hospitality, and health care are among the top industries in which optimal staffing is a main driver of profitability. The goal is to have the right person, in the right place, at the right time, with the right skills, for the right price.

Predictive models using market measures such as sales, usage, and attendance combined with skill level, location, and mobility are increasingly being used to predict human resources demand and employee retention.

Asset Maintenance

Maintaining assets to avoid breakdowns and process failures is critical to financial survival in many industries.

- Hospitals and other medical treatment facilities use predictive models to identify patterns showing where medical errors occur. By predicting points of likely failure, managers can improve processes and establish checks and balances, thereby improving safety and reducing costs.
- Transportation systems (rail, air, truck, and auto fleet) use predictive models to estimate the likelihood of system interruptions and mechanical failures based on patterns of past events. This allows them to optimize the location of maintenance service and parts as well as assist in labor planning.
- Manufacturers in a variety of industries, including pharmaceutical, automotive, technology, agriculture, and equipment, avoid the expense of failure and protect workers by modeling patterns of past failures to predict the likelihood of breakdowns or other costly errors.
- Energy companies are building smart grids to collect data from suppliers, users, and strategically placed sensors. The data is used to develop models that predict usage patterns and equipment failure.

Government

National and local governments around the globe are using predictive analytics to estimate economic and social trends, enhance security, and reduce government fraud. As many countries in the developed world face tightening budgets, predictive models are able to predict costs in areas such as social support of aging populations, security infrastructure, law enforcement, and utility consumption.

Life Sciences

There are myriad uses for predictive analytics in medical research, pharmaceuticals, biotech, and genomics. Regression and neural network models are used to predict the occurrence of disease as well as mortality rates from diabetes, heart failure, cancer, and more.

Doctors and pharmacies are able to better manage their inventory for certain medicines through the use of predictive analytics. Models are being developed using over-the-counter and prescription sales within specific geographic areas. The scores use purchase patterns to estimate the likelihood of an outbreak within a geographic area for a variety of common ailments. This allows local healthcare providers to manage staffing and inventories for optimal care.

The Role of Leadership and Culture

To truly leverage the power of data in our high-tech, global economy, organizations require a new kind of leadership — one in which everyone is encouraged and empowered to think like a leader. The old hierarchical, top-down leadership model is no longer viable. Markets are shifting too quickly for data to be synthesized and communicated to top management, who is then responsible for creating directives to send back down the chain. In today's volatile global economy, speed to market provides a real competitive advantage.

Analytic talent will be in high demand for many years to come. Managers have the unique challenge of hiring “whole brain” thinkers who understand analytics as well as how the results drive the business. Top executives need to develop a “culture of analytics” that encourages an analytic-centric business model where everyone feels empowered to think like a leader. As Thornton May of the IT Leadership Academy puts it, “The question facing our industry is not ‘Do leaders understand analytics?’ but rather ‘Do analysts understand leadership?’”

Leadership in times of unprecedented complexity works best if top executives define the goals, articulate them to their teams, provide ample resources, and empower their teams to set strategies to reach those goals. Collaboration must be encouraged at every level. Skill development in effective communication and emotional intelligence are critical.

Success in today’s high-tech, fast-paced global economy will be enjoyed by those organizations that drive their business with predictive analytics with the agility of empowered leadership.

References

Parr-Rud, Olivia, *Data Mining Cookbook* (Wiley 2001)

Market Innovations Inc., *Multi-Stakeholder Collaboration, Innovation & Simulation*
[http://marketinnovationsinc.com/White_Papers/MCIS_Approach_\(12-19-11\).pdf](http://marketinnovationsinc.com/White_Papers/MCIS_Approach_(12-19-11).pdf)

Market Innovations Inc., *Optimizing New Telcom Stand-Alone and Bundled Offers Using Predictive Technologies*
[http://marketinnovationsinc.com/White_Papers/Optimizing_Service_Provider_Bundles_\(12-20-11\).pdf](http://marketinnovationsinc.com/White_Papers/Optimizing_Service_Provider_Bundles_(12-20-11).pdf)

Modeling Technologies, May 2011

Parr-Rud, Olivia, *Business Intelligence Success Factors* (SAS/Wiley 2009)

SASCOM, Third Quarter, 2010

“Why Predictive Analytics is a Game-Changer,” Dave Rich and Jeanne G. Harris, *Forbes*

<http://www.forbes.com/2010/04/01/analytics-best-buy-technology-data-companies-10-accenture.html>

SASCOM, Fourth Quarter, 2010

“Survey Puts Fraud Loss at \$2.9 Trillion Globally in 2009,” [Tammy Whitehouse](#)

Compliance Week, <http://www.complianceweek.com/survey-puts-fraud-loss-at-29-trillion-globally-in-2009/article/189557/>

SASCOM, First Quarter, 2011

“Combining Structured, Semistructured and Unstructured Data in Business Applications,” Jacob Ukelson (InfoManagement Direct, December 2006)

<http://www.information-management.com/infodirect/20061201/1069202-1.html>

SASCOM, Fourth Quarter, 2011 (Reprinted from WRAL Tech Wire, www.wraltechwire.com)

The Use of Social Network Analysis in Mobile Operator Marketing, iDiRO Technologies,
<http://www.idiro.com/>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries.
 ® indicates USA registration. 105620_S90792.0412