The Art and Science of
Predictive Analytics

Predictive

Driven by the explosive growth of technology and ongoing successes coming from the development of its applied sciences, "Predictive Analytics" has rightfully earned a critical role within the vanguard of industry, government, and science.

Yet, there are many myths and misperceptions related to what Predictive Analytics actually is, how it works, and what it can or cannot do. In reality, its tools and methods are based in proven science. At the same time, its successful application is knowledge and skills-based. Accordingly, at RPM² Predictive Analytics is both an art and a science. One cannot exist without the other. Both are critical factors for success.

DATA MINING vs. PREDICTIVE ANALYTICS

There are many terms for what we do. You may hear it referred to as "Data Science", "Machine Learning", or "Database Marketing". And while one could argue finer points of difference, each essentially involves the same processes. At RPM², we prefer two terms, "Data Mining" and "Predictive Analytics", because this distinction is important.

Data Mining is the discovery of previously unknown and useful *patterns* of information from databases or streams of data. Because these sources are large, Data Mining employs the skillful use of computer-assisted algorithms to identify these patterns.

However, Data Mining derives these useful patterns from historic data. What if we want to project these patterns into the future or onto another set of data? That's the role of **Predictive Analytics**.

To illustrate a classic application, Data Mining can be utilized to identify patterns that differentiate historic marketing responders from nonresponders. Predictive Analytics takes the process one step further and projects which individuals will respond to future marketing efforts. This distinct, future focus is what makes Predictive Analytics a compelling, strategic force for industry leaders.
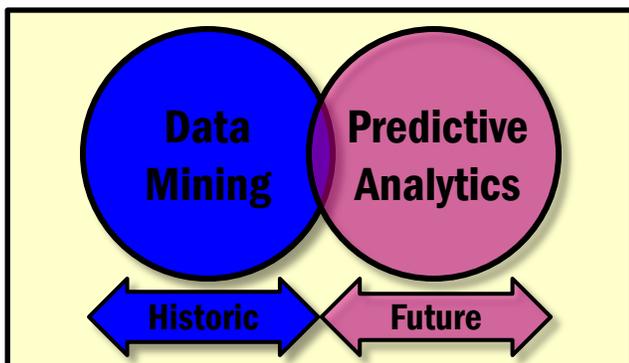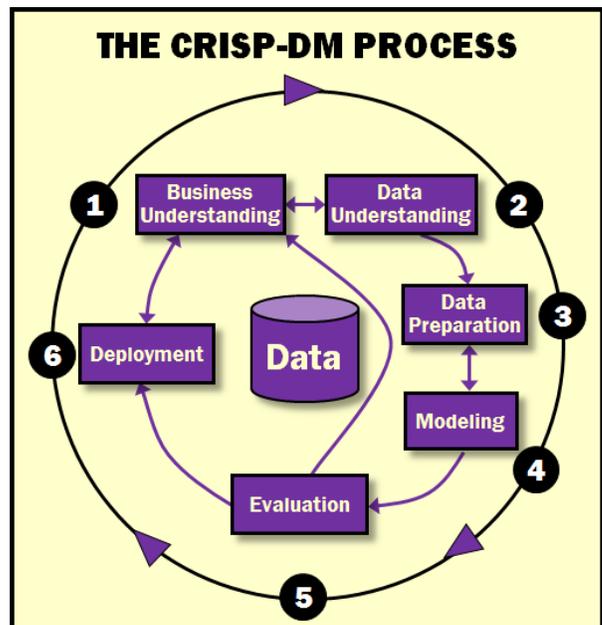

PREDICTIVE ANALYTICS IS A PROCESS

Perhaps the two most common misperceptions about Predictive Analytics are that: 1) It's mainly about the use of sophisticated algorithms and 2) It's automated software. Sadly, these misperceptions are often driven by our own industry. So if these are misperceptions, what are the facts?

First, Predictive Analytics involves much more than an exotically-named algorithm. It actually is a **diagnostic and prescriptive process** where the use of modeling algorithms comprises only one part of the whole. Second, while total automation is a noble goal that we hope to achieve in time, current "automated modeling" approaches either completely miss or minimize key parts of the process.

At RPM², we employ the "Cross-Industry Standard Process for Data Mining" or "CRISP-DM" as our template for developing Predictive Analytics projects. It is the leading process in the industry:

THE CRISP-DM PROCESS

CRISP-DM is a six-step process that is centered on data. We see that its six steps are major elements of an interconnected loop that has an iterative nature. To illustrate, when developing a Predictive Analytics project, a modeler will typically go through several cycles of steps two and three before proceeding to step four. And it's not until step four that we see the algorithms employed. Even so, the most important step is step one. If step one isn't done or if you get it wrong , you risk total project failure:

"Predictive Analytics Is An Objective-Driven Process"

WHAT'S THE BEST ALGORITHM TO USE?

This is where the art of Predictive Analytics comes into play. Though a wide and increasing range of algorithms are available, in short, no single "best algorithm" exits for every application (another common misperception).

The decision regarding which algorithm to employ results from the diagnostic part of the Predictive Analytics process. For example, every situation is unique and an operating environment may favor one algorithmic approach over another. So, choosing the right algorithm or algorithms is prescriptive in nature.

Even so, there are some algorithms that, because of their maturity and relative ease of implementation, are more commonly-employed. They include:

• Classification Trees, and
• Logistic or Linear Regression.

They are widely-used simply because they are well-known and effective. Yet the palette of available algorithmic tools has expanded. So a Predictive Analytics solution may include other, lesser-known but equally-powerful algorithmic approaches such as:

• Bayesian Models,
• Neural Networks,
• Support Vector Machines,
• Genetic Algorithms, and
• Multivariate Adaptive Regression Splines.

But perhaps most important, as a direct result of the advancement of computing technologies, more powerful sampling techniques are now practical. These same advances have also made it practical to combine these cutting-edge sampling techniques with multiple algorithms into a potent approach that is known as "**ensemble modeling**".

THE GOAL OF PREDICTIVE ANALYTICS

When it comes to business, the ultimate objective of Predictive Analytics is profit maximization. And it achieves this objective through the targeted realization of opportunities and financial efficiencies.

At the same time, Predictive Analytics does not exist in isolation. It usually must be coordinated with several functions within the enterprise and often with external vendors to be successful. Stated differently, **a Predictive Analytics program is only as successful as its weakest link.**

At RPM², we realize this fact and we leverage our extensive experience and resources to go beyond typical model development and help ensure your program's success.

COMMON APPLICATIONS

The critical questions every successful business needs to address are:

❶ Who are my best customers?
❷ How do I keep them?
❸ How do I find more like them?
❹ What do my customers want?
❺ How can I grow their involvement?

Predictive Analytics has solutions for each, including:

• **LOOK ALIKE** models identify highest propensity consumers within your current customer base or in the world of prospects. Most often, these are used to identify best customer look-alikes though they could be look-alikes of customers who have other characteristics.

• **SEGMENTATION** models that identify different groups of consumers by their purchase behaviors or characteristics. These are used to target offers or messages that motivate by appealing to their unique preferences.

• **RETENTION/ATTRITION** models that are designed to identify customers that you are at risk of losing so you can reach out and save the relationship.

• **UPSELL/CROSS-SELL** models that are quite effective at increasing your total and average sales.

• **RECOMMENDER/MARKET BASKET** to help you identify and group products that are most often purchased together.

• **UPLIFT** models maximize sales promotion returns. Which customers do you subsidize because they will purchase without a promotional offer? And which customers are truly generating incremental sales … your promotional efforts are motivating their purchase behaviors?

• **FRAUD DETECTION** models to help identify individuals or entities who are taking advantage of your business .

WHAT'S NEXT?

Successful development of a Predictive Analytics program is usually not an end, but is a first step. More often than not, initial success leads to a "Can we do more like this?" question with a "Yes!" answer. In turn, this leads to additional, more highly-targeted models that collectively produce even greater levels of profit maximization.

And, of course, there's more we can add to this conversation. But we've run out of space. So feel free to give us a call. It's toll-free and we'll be delighted to continue the discussion and answer your questions!

866.267.5369

Rapid Progress Marketing and Modeling, LLC

1760 Delaware Ave NE
St. Petersburg, FL 33703

Phone: 727.528.8578
Sales: 866.267.5369
Fax: 866.274.3030

www.RPMSquared.com