When You Think Analytics, Think DOMINOES

Greg Hart, Principal

Executive Summary

Predictive analytics and data mining continue to garner more and more attention. As the business landscape becomes more competitive, it is increasingly important to leverage existing data assets for gaining deeper insights, or perhaps to even survive. Predicting sales, warranty claims, and personnel attrition becomes a competitive advantage. Getting started with analytics is easier than you might think. Let's explore some of the key concepts.

Introduction

Perhaps when you were growing up, you were exposed to the game of dominoes. You may have been drawn to the game, due to its perceived simplicity or maybe you were a numbers person. Before long, you figured out that the game became more competitive, required strategy (with a bit of luck) and was often frustrating. Ultimately, the pieces fit together, regularly quite elegantly. There was great satisfaction in seeing the game progress, yielding results that conceivably were not anticipated. Maybe you just marveled at how careful vision, planning and execution was required to line up dominoes, unfolding to reveal unique and innovative patterns, That's very similar to the journey one experiences when embracing predictive analytics.

Definition of the Business Problem/Opportunity

As with all successful projects, it's imperative to identify the problem to be solved or the opportunity to be exploited. Say your initial focus is in the customer area. What characteristics in your business truly determine whether you have a high-margin/low-margin customer, what drives a customer to buy from you again, and what indicators might be present to highlight a slow-paying customer? Knowing what questions to answer drives how you organize your resources to tackle the effort.

Organized Approach

To undertake an analytics initiatve, you'll need to conduct a readiness assessment. Diving in without proper preparation will result in disappointment from your project sponsors, and frustration with the analytics team. At a minimum, you'll need to concentrate on the following three areas:

People Elements

- Data Extraction/Manipulation Skills
- Data Analysis Skills
- Statistical Methods Skills

Process Elements

- Data Collection/Processing Steps
- Survey Process (if intended to supplement data)
- Statistical Analysis

Technology Elements

- Hardware – do you have adequate horsepower to crunch the numbers?
- Software – do you have the appropriate toolkit to do the job?
- Data – do you have the right data to analyze?

Modeling

Analysis of Variance (ANOVA), linear regression, and correlation. Does the mere sight of these items send shivers down your spine, reflecting on your college statistics course? Yes, they are components of an analytics solution. And, properly managed, can be most valuable in your overall modeling process. Crafting basic models, testing various scenarios and iterative model refinement are part of modeling the solution.

Information

Information is the lifeblood of any analytics project. At a minimum, you will need to isolate high-quality data with the following characteristics.

- Valid – Is the data trustworthy?
- Complete – Is there a preponderance of missing data?
- Assessable – Can you get to the information you need?
- Integrated – Is the information related in a proper fashion?
- Representative – Does the dataset sufficiently represent the problem domain?

There is oftentimes significant manual data creation to sufficiently round out a model. Collecting information from external sources can be both challenging to acquire as well as to integrate. Plan accordingly.

Normalization

Once you have a handle on the information scope, you'll need to take some preparatory steps in order to move forward. Initially, you'll need to clarify the structure of the data. This will likely involve joining information from multiple datasets, in order to achieve a comprehensive view of the information. This will be followed by interrogating the data and classifying it for proper processing and analysis. Typically, you will have a combination of *qualitative* (categorical, e.g., gender) and *quantitative* (measurement, e.g., salary) information. To aid in summarization and more sophisticated analysis, you will likely need to create logical groupings of your information (e.g., age group ranges.)

Ownership

You will likely identify a key project sponsor (or two.) Now that you have collected, interrogated and evaluated your information, along with candidate models/scenarios, you'll want to do a checkpoint with your sponsors to confirm that project ownership is still solid. Given that the exploration portion of the analytics project is sometimes perceived as trial-and-error and yielding contradictory results, you will need to take particular care to reassure your leadership that that is the nature of such projects. Also, since answers to questions beget new questions, scope can often creep. If you find yourself inadvertently drifting into business areas not stipulated in the *Organized Approach* section, stay in control. Predictive analytic projects often gain the attention of other teams both curious and envious of the tools and techniques being employed. Maintain focus, and log the additional candidate initiatives for future pursuit.

Exploration

Here's where all of your hard work pays off. You will identify outliers in your data analysis. Patterns will begin to emerge, as well as data anomalies. These will dictate iterative model refinement. You'll find yourself circling back for additional datasets to augment your findings. All along the way, you will be validating expected behaviors, proving out hypotheses, and confirming gut instincts. You'll be comparing variables (because you can) that prior to this initiative would have been prohibitively difficult or expensive. There's a sense of excitement in looking prospectively at information. Here's where you'll let the tools do the heavy lifting, and you'll do the heavy thinking.

Software

It should come as no surprise that software plays a critical component in analytics. With rich visualization graphics and depth of capabilities, software in this space has become more sophisticated. On one end of the continuum, some organizations have started small and constructed rudimentary models with spreadsheet applications. Database vendors have also extended their product offerings to include varying levels of analytic and data mining capabilities. On the other end of the continuum are those vendors that offer specific niche products devoted to statistical analysis. Like any software selection, your choice will be predicated on features, business requirements, business maturity, and costs. Start with a pilot project. Use external resources that have experience in this domain to shorten the discovery cycle, as well as reduce the overall risk. You'll learn a great deal, while keeping the investment manageable.

Conclusion

Predictive analytics has become more affordable, accessible, applicable and advantageous than ever before. It takes a special balance of people, process and technology to achieve success. It taps into the intuition and creativity of your business leadership. It's not an easy journey, but one that can pay substantial dividends. The old adage, "… if you want something you've never had, you must do something you've never done …" couldn't be more pertinent. Just think DOMINOES.

****

About Pariveda Solutions, Inc.

Pariveda Solutions works with organizations to improve their profitability through the deployment of process and technology. Pariveda delivers solutions in the areas of IT Strategy, IT Executive Advisory services, Program and Project Management, Application Development, System Integration, CRM and Business Intelligence. Pariveda's goal is to establish relationships with clients on a local level and deliver high value solutions.

Pariveda Solutions was ranked the 16th fastest growing company in the Dallas Business Journal's 2007 edition and the 34th fastest in the 2008 edition of the Dallas One Hundred, comprised of the 100 fastest-growing private companies in the DFW Metroplex. Pariveda Solutions was also recently named to the Dallas Business Journal's "Best Places to Work" for 2008 as well as one of Consulting Magazine's 7 "Small Jewels" for 2008. Launched and headquartered in Dallas, Texas, Pariveda Solutions has grown to over 150 employees since 2003. The company has additional offices located in Chicago, Denver, Houston and Seattle.

Learn more at www.parivedasolutions.com