



STATISTICS SOLUTIONS

INTELLIGENCE IN DATA

Customer Acquisition and Management Through Lead Propensity Scoring

www.StatisticsSolutions.com

877-437-8622

Key Findings

- In our data-driven world, potentially profitable patterns are buried among the volumes of in-house data.
- Customer acquisition algorithms determine what type of marketing campaign should be performed.
- Data-driven marketing saves money and increases the return on investment by allowing marketing campaigns to specifically target individuals who are most likely to purchase a company's products and services.

Introduction

Most businesses realize growth through the acquisition of new customers. This typically occurs when new customers become aware of your product, become a new prospect in the market for your product, or were previously consuming the product from a competitor. For each of these cases, data mining techniques are a proven way to segment these new customers and increase the response rates of your marketing campaign.

The traditional approach to this type of marketing is some combination of mass marketing advertisements, including newspaper advertisement, billboards, radio, television, etc., and direct marketing, such as mailing or telemarketing campaigns targeting specific potential customers. For example, a diaper manufacturer is likely to perform a mass marketing campaign targeting venues frequently viewed by current and future parents, such as parenting magazines, or on billboards adjacent to baby stores. The advertisements may also appear in mass market publications that are read by customers of the appropriate demographic.

Traditional approaches to direct marketing campaigns are quite similar to mass marketing campaigns. The marketing campaign manager typically selects the demographics that they are interested in, and contracts with data providers to obtain lists of customers who meet the necessary criteria. Unfortunately, in many cases the demographic information could be as wide as, or even wider than, the mass marketing audience. The data providers collect information on customers by gleaning addresses, names and demographics from rebate processing centers, warranty registration centers, or by simply purchasing that information from businesses or data vendors. For a diaper marketing campaign, the target demographic would be all persons 18 to 30 years old; perhaps a marketer on a budget could limit the demographic to only women between 18 and 30 years old. The data vendor would take that demographic request and return a list of names and addresses so that the diaper company could contact the customers directly. A perhaps more sophisticated marketing director may ask for names and phone numbers of all women between the ages of 18 and 30 who have registered a warranty on a recently purchased stroller.

Of course the demographics selected here are mostly based on hunches and best guesses. Contacting women between 18 and 30 years old is hardly an

effective way to carry out a direct marketing campaign. The more sophisticated marketer looking for stroller purchasers is relying on their intuition that people who registered stroller warranties are probably pregnant or have recently had a baby. **Unfortunately, this intuition does not account for the pregnant women who do not register their stroller warranties, nor does it account for women who are pregnant with their second child, who, in all likelihood, would just use the stroller from their first child.**

We find that these types of direct marketing campaigns are more of an art than a science. The intuitions about what to market to whom is a valuable skill set derived from years of experience. More often, direct marketing campaigns yield poor results and stuff our mailboxes with junk mail.

Data Driven Customer Acquisition

The recent boom in data collection and storage has given way to extremely large and complex data sets. The intricacies of newly available data are not nearly as simple as the diapers example from above. In the new data-driven world, useful patterns are buried amongst the vast volumes of complex data. This reality has driven the development of adaptation of new data mining tools that are able to find interesting and actionable items within the data.

These “data-gems” are not always obvious or initially intuitive, even to the seasoned marketing campaign manager. Fortunately, data mining tools are able to explain ‘why’ a particular data-gem is interesting. For example, our intuition may indicate that women who registered warranties for strollers may be interested in buying diapers because they are pregnant with their first child. However, data mining tools can provide a much richer set of information, such as women who purchased strollers priced within \$25-50 buy store brand diapers, while women who purchased more expensive strollers along with a name-brand diaper bag are 55% more likely to purchase name-brand diapers. Clearly the data mining results may show a much richer set of results as well as potential explanations for the results. Formally, this is known as market segmentation, wherein potential customers are grouped into segments. These market segments can be specifically addressed in a marketing campaign in order to

provide customized advertisements leading to an increase in response rates.

Market segmentation is only half of the story: so far, we've only discussed who to advertise to, not what to advertise. Returning to our previous example, given the two market segments: 1) women who buy cheap strollers buy store-brand diapers, and 2) women who buy expensive strollers and diaper bags buy named-brand diapers, what type of marketing campaign should be performed? Should the diaper company offer coupons to the first segment, and testimonials to the second segment? Coupons to both segments? Or testimonials to both segments? Data mining can help make these decisions.

With market segments in hand, the diaper company begins the process of customer acquisition. The goal of customer acquisition is to turn potential customers into actual customers of your product or service. Customer acquisition can be difficult to measure because there are different types of customer responses to consider.

Responses that occur as a result of a marketing campaign are called response behaviors. The most simple response behavior measurement is a yes/no response. In this type of response we measure 'yes' if the customer bought the product, and 'no' if the customer did not buy the product. This type of response is typically the one that matters most to a company's bottom line, but is the least informative to the marketing manager. Customer behaviors frequently contain subtle, complex, and measurable information that can be used to deeply understand the effectiveness of the marketing campaign.

Beyond yes/no response behaviors, categorical responses provide a rich set of information that can be used to tweak the marketing campaign. Returning to the diaper example, a categorical response may be the number of customers who received coupons that purchased size 1 diapers, size 2 diapers, etc. as well as the number of customers who purchased 25 count, 50 count or 100 count diaper bags, as well as the number of customers who spent between \$0-10, \$11-25, \$26-50, etc. in diapers. Combining this response data with the particular store, day of the week, and time, etc. allows the campaign manager to further customize the marketing campaign.

Categorical responses may also contain information about the type of customer that responds to the campaign. These are typically segmented into new and

existing customers. New customers can be new to the market, *i.e.*, new mothers, or not new to the market but new to the brand (*i.e.*, they switched brands). Existing customers can continue to purchase your product or switch to a competitor's product.

Begin With A Test Campaign

Test campaigns operate by performing an initial marketing campaign on only a small subset of the prospective customers, as well as sets of random customers. As the response rates begin to roll in from the test market the data mining tools begin to create statistical models on the results which provide actionable information to the campaign manager. In this process, the use of random customers is needed in order to provide a baseline that the prospective customer segments can be compared against. More importantly, responses from random prospects can give insight into market segments for future consideration. For example, a test campaign for diapers on the random population may find that women aged 45-55 frequently responded to the campaign. With this information the diaper company creates a new market segment consisting of mothers of pregnant women, that is, grandmothers. Furthermore, the test campaign found that only 2% of women aged 28-30 years old switched brands, indicating that the diaper marketing strategy did not convince respondents to change their diaper brand.

The results of a small test campaign can inform or reinforce the statistical model within the data mining tool, which in turn, influences the campaign manager's strategy. As a result of the test campaign, informed by the statistical model, the campaign manager decides to narrow the market segment for new mothers from women aged 18-30 years to women aged 18-27 years, as well as create a new market segment strategy targeting new grandmothers.

With the statistical model containing preliminary results in hand, the marketing campaign can begin to expand in size and scope. As new response behaviors roll in the model is refined, the marketing strategy is tweaked and the campaign is expanded even further.

Unlike the traditional approach, which is informed by intuition and best guesses, the data-driven marketing approach saves money and increases the return on

investment by allowing marketing campaigns to specifically target individuals who are most likely to purchase your company's products and services.

ase Study: HSBC Credit Card Direct Mail Campaign

In 2005 HSBC reported their results for a recent direct mail campaign. Prior to their direct mail campaign they understood that credit cards are issued based on the prospective customers' FICO score, which is based on the customers' financial history. However, the FICO score was a poor indicator of the usage patterns of credit cards. HSBC wished create a new indicator that modeled prospective customers' usage rate in order to perform a direct mail campaign aimed at prospective customers with a high predicted usage rate.

The first step of this campaign was to collect data on the prospective customers including previous account information, transaction information, demographics from a third party data vendor and cluster information from a third party data vendor.

Based on this information, a statistical model was created accounting for seasonal trends and by ignoring old data and giving more weight to more recent data. The data cleaning and data aggregation steps resulted in thousands of variables potentially useful in creating predictive models.

Using existing usage rate information as validation, HSBC discovered the specific variables that are key deterministic factors (e.g., age, number of credit cards, current debt load, etc.) in high usage rates. The final model was able to determine which prospective customers are likely to use a credit card in the next 12 months.

From this information HSBC created specific marketing strategies: 1) recouping inactive customers, 2) finding customers who are likely to be customers during special events (e.g., Christmas, back-to-school, etc.), 3) determining which customers are likely to *not* respond to any marketing campaign, 4) detecting inactive customers likely to respond to direct mail campaign, and 5) preventing attrition by detecting customers with declining usage rates.

HSBC reported results on three direct mail campaigns:

Seasonal Sales Event:

- Mailings to 300,000 random customers showed a return on investment of 7%
- Mailings to top 300,000 model-selected customers showed ROI of 59%
- Mailings to top 100,000 model-selected customers showed ROI of 168%

Reactivity Mailing:

- Mailings to inactive customers (more than 8 months) had response rate of 1% showing ROI of -59%
- Mailings to model-selected customers that are inactive (more than 8 months) had response rate of 2.3% showing ROI of 182%

Early Rewards Enrollment:

- Mailings to model-selected customers showed ROI of 965%

Statistics Solutions Company Biography

Statistics Solutions has been in the statistics consulting business for 18 years. For those interested in data mining solutions, we offer a wide range of consulting services and technology to assist you in optimizing your marketing campaigns, including the building of acquisition models, cross-sell and up-sell models, and churn models. For organizations without an analytic capability in-house, data mining activities will show an immediate ROI to your marketing campaigns.

For those firms with existing analytic capabilities in-house, we offer an outside opinion. For example, we specialize in trying different models (in addition to the typical SVM, gradient boost machine, partial least square, glmnet, generalized additive model, neural networks, etc.) and assumptions on the data, in order to get better lift. We also do meta-modeling (bagging & boosting of an existing

models, and ensemble of a collection of different models), where we have seen improvements of 5%-10%!

For more information, you can reach us at 877-437-8622,
Info@StatisticsSolutions.com, or visit our website at www.StatisticsSolutions.com.

Statistics Solutions, June 2012. James Lani, Ph.D. & Tim Weninger.

References

Berson, Smith, and Thearling. Building Data Mining Applications for CRM. McGraw-Hill Companies, 1999.
Hofmann. Marketing Strategies for Retail Customers based on Predictive Behavior Models. Salford Systems Data Mining. New York, 2005