

Modeling Marketing Response for a Communications Client

Cactus Strategies



By Carl Cozine

Response modeling for direct marketing involves determining what characteristics of a potential customer of some service make him/her a buyer. The cost of direct marketing campaign is directly proportional to the number of prospects being targeted: each contact, whether by phone, direct mailing, or a sales representative, increases costs. However, only those prospects that actually buy a product contribute to revenue. Reducing the number of contacted prospects, while increasing the response rate, saves on marketing costs and increases profit.

Cactus Strategies, a consultancy specializing in heuristic business tools and models, is developing a state-of-the-art customer care and acquisition system for a client, a communication services company. The client is focused on introducing a suite of new products to the market and is seeking to improve their direct marketing activities.

The created heuristic tool is capable of predicting which potential customers have actual demand for the introduced products and which customers are most likely to become buyers. The developed system assesses the probability that a potential customer needs the promoted products and might choose a particular class of service, based upon their business profile.

The key data mining elements of this project are the identification of a business rule that predicts demand for the client's services and the continuous improvement of this profile. The data mining algorithms of Megaputer Intelligence's PolyAnalyst Pro software are at the heart of the heuristic tool, providing insights beyond those discoverable by any other analytical package.

The data provided by the client included information about around 6,000 companies that expressed interest in the products similar to the newly promoted one. Since no information about nonbuyers of the new products is available, Cactus Strategies added to the data around 6,000 records of randomly selected companies, purchased from an independent vendor of business data. Several pieces of information were provided about each company: the size, location, business classification, operation type, time period the company was tracked, etc. – the total of 53 attributes. A preliminary analysis resulted in selecting 12 attributes thought to be most predictive.

In order to solve this problem, the data was analyzed with the help of Megaputer's PolyAnalyst Pro. This data mining suite features a broad selection of exploration engines for predicting values of continuous variables, explicitly modeling complex phenomena, determining the most influential independent variables, and solving classification and clustering tasks. A joint application of PolyAnalyst algorithms resulted in a model capable of accurately predicting purchase decisions.

The first step of the analysis was to transform the data into a form suitable for analysis by aggregating the values of some variables and substituting some original variables by their more predictive combinations. Then two rounds of the Find Dependencies algorithm facilitated quick identification of the set of variables that had the greatest influence on the purchase decision. The time taken by more elaborate machine learning algorithms of PolyAnalyst depends on the number of variables investigated, and thus eliminating redundant variables greatly speeds the analysis.

Then the Classify algorithm was run, analyzing only the variables found to be important. The Classify algorithm uses fuzzy logic to develop a continuous function modeling the probability that a record represents buyers/nonbuyer – and selects a threshold for minimizing the number of incorrect classifications. The PolyAnalyst Find Laws exploration engine, chosen to power the classification, allowed for the fast production of a viable and explicit model. After about eight hours of perfecting the model, the system found a classification rule predicting with 81% accuracy the probability that a potential customer will be a buyer based only on three their characteristics.

Later the Cluster algorithm helped refining the model even further. The resulting business rule had the advantages of being descriptive and easily applied to the bulk of the data, and of satisfying the client, particularly interested in predicting buyers using categorical variables. PolyAnalyst provides means to easily deploy discovered rules in external data storing applications ranging from Microsoft Excel to IBM Visual Warehouse. This has immediate business effect, as now the client can score all their potential customers by the likelihood of purchasing new products and direct market only to likely buyers.

The business value of the discovered model is readily visualized by PolyAnalyst Lift and Gain charts. The Lift chart evaluates the benefits of performing a model-based vs. random marketing campaign. The Gain chart illustrates the dependence of dollar-based profit on the number of model-suggested prospects contacted. It allows the company to optimize the number of prospects contacted to achieve a balance between the maximum profit and exposure. For a Gain chart, the cost per contact, profit per response, and maximum number of prospects for the marketing campaign are entered. For the selected parameters the profit peaks when 1,500 best prospects are targeted.

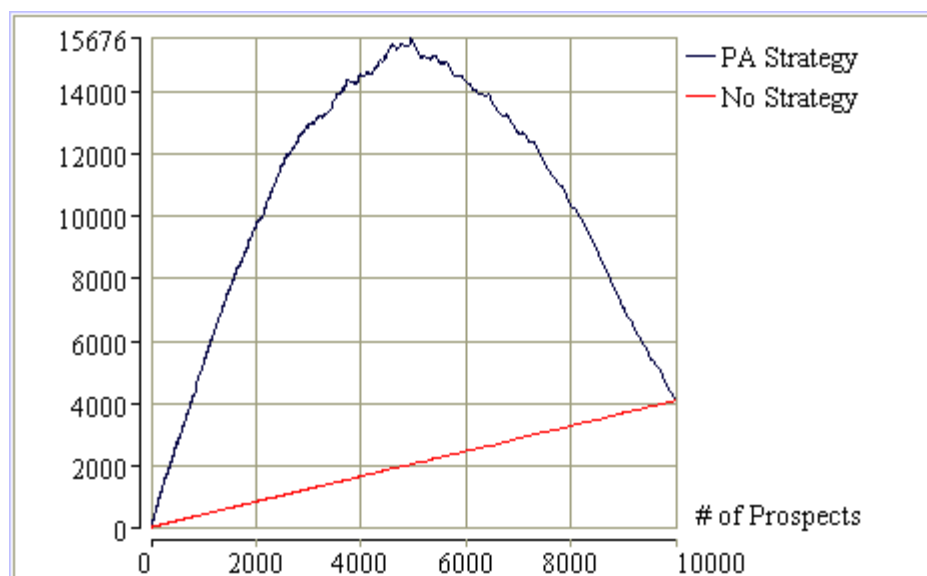


Fig 1. The PolyAnalyst gain chart helps identify the maximum profit location

With a response model at hand, the client can identify the most likely buyers prior to spending time and money communicating with prospects. Contacting only a fraction of prospects – those most likely to purchase – results in lower direct marketing expenses and better response rate, and therefore increases profit.

Carl Cozine is Principal of Cactus Strategies, LLC. He may be reached at (720) 283-7380.