USING INTELLIGENT PROFILING TO GENERATE SMART LISTS: AN EMPIRICAL TEST

Don Kridel, University of Missouri-St. Louis, Department of Economics, 8001 Natural Bridge Rd, St. Louis, MO 63121 USA, (314)516-5553, kridel@umsl.edu

Dan Dolk (IS/Dk), Naval Postgraduate School, Information Systems, Monterey, CA 93943, USA (831)656-2260, drdolk@nps.navy.mil

ABSTRACT

Kridel and Dolk [1] describe a web-based service that uses advanced data mining and dynamic model building techniques to conduct intelligent profiling. This *smart list* approach is designed to assist small-to-medium businesses (SMBs) in leveraging their analytical marketing capabilities. We begin validation of this approach by comparing the performance of the *smart list* for one target company versus the service provided by one of the industry leaders in customer list generation. Initial results indicate that the *smart list* approach outperforms the traditional approach by a substantial margin for both response rate and ROI.

INTRODUCTION

[1] describes data mining techniques derived from research in the areas of active decision support and automatic model generation ([2], [3], [4], [5]) to generate context-specific customer lists for SMBs. In this approach, the general objective is to automate parts of the demand forecasting econometric modeling process, specifically, determining appropriate dependent and independent variables and their associated functional form. Heuristics are used to generate equations to evaluate candidate prospect lists, and then iteratively refine these equations until they converge to a final prospect list. In this paper we seek to perform an analytical comparison of these *smart lists*, i.e. the predictive-modeling-based intelligent profiles, versus the industry standard simple lists. For a selected target company we ran two campaigns, one using *smart lists*, and one using a purchased list from an industry leader.

CURRENT PRACTICE

The industry standard simple lists require the firm to “guess” at what their prospects “look like”. For example, the typical list vendor offers the following selection criteria:

- Distance-based: households or businesses within a given zip code or pre-specified distance from your business
- Demographic-based: households or businesses with specific demographic characteristics. For example, households with income over $100,000 and whose age is less than 45; or business with over $1M in sales and over 100 employees.
- Combined demographic and distance-based selections.

Not surprisingly, the resulting prospect list is only as good as the firm’s guess. Further, there is no way to “weigh”, or trade off, appropriate characteristics. For example, is a more-wealthy household 4 miles away a better (or worse) prospect than a less-wealthy household that lives 2 miles away?

INTELLIGENT PROFILING AND SMART LIST (IP-SL) APPROACH

The *smart list* process works as follows:

- The customer uploads a customer list.
- IP-SL matches the uploaded customer list to a national business (B2B) or household (B2C) database.

- A custom predictive model is built that generates an optimal customer profile by identifying a "dependent variable" (matched customer list and sampled non-customers) and iteratively determining independent variables and their respective functional form (based on customer profile reports and active econometric methods [2]) for a series of logit regression models.
- IP-SL generates a "smart list" from model scored list [1], and this profile is used to score (rank) every household in the market area.
- The best (highest ranking) households are selected for the customer prospect list.

By scoring each household using the regression parameters and household characteristics, the modeling process specifically solves the trade-off problem, e.g., whether distance is relatively more (or less) important than income.

AN EMPIRICAL TEST

Sugar Creek Gardens (a small high-end perennial garden store in suburban St Louis, MO) was utilized as a target company for comparing IP-SL with the simple list approach. Sugar Creek conducted a campaign utilizing 2 lists, one from a competitor and one from IP-SL. Prospects were mailed postcards that had to be returned to receive a discount on a single purchase. (The postcards and spending were collected to carefully measure the response and effectiveness of the campaign.)

Sugar Creek's provided customer mailing list contained approximately 5,000 names; after cleansing and matching, approximately 3,300 names were utilized in the IP-SL. For the IP-SL part of the empirical test, we used BizFusion™, a service based on this process available through www.copperkey.com. The 1,000 highest-ranked prospects were mailed the discount coupon. For the competitor list, a hybrid distance and demographic based sample was selected. (The list was selected on the basis of income, age, home ownership and distance from the firm.)

Table 1 summarizes the results of the two campaigns. The competitor list generated a response of 1.4% (within the "so-called" expected or normal return of 1-2%); the IP-SL list generated a response of 7.8%. It is worth noting that the intersection set between the two lists was less than 1%! With respect to response rate, the IP-SL list outperformed the competitor list by a factor of over 557%. Perhaps, more importantly, since revenues were collected as well, return-on-investment (ROI) could be calculated. Two different measures are provided:

- Direct ROI: compares only the cost of the campaign to the revenues generated—this measure is often used to evaluate DM campaigns;
- Full ROI: compares total incremental costs (campaign, cost-of-goods sold, and discount) to incremental revenues; for this measure, immediate, one-year and three-tear results are provided (retention rates and average spend for subsequent purchases come from the business owner)

By either measure, IP-SL significantly outperforms the traditional simple list approach. Indeed, the simple approach does not immediately cover its costs, though is expected to generate a positive ROI over the first year (and beyond).

CONCLUSIONS

The results show a dramatic increase in response rate and ROI performance for the IP-SL method compared to the simple list approach for this case. It suggests that our data mining technique may be a significant marketing multiplier for generating customer prospects. Unsurprisingly, we also found that the analytical approach embraced by IP-SL worked far better than human intuition, since the two lists were effectively mutually exclusive. Our intention is to extend the study to more test cases (for both B2B and B2C).

Table 1: Comparison of alternative list performance

	IP-SL	Competitor Targeted
List Size:	1000	1000
One-time charge for IP-SL service	$ 495	$ -
List charge	$ 200	$ 210
Mail charges	$ 200	$ 200
Postcards (creative, stock, printing)	$ 500	$ 500
Total Campaign Costs	1395	910
Direct Mail response rate	0.078	0.014
New Customers generated	78	14
Average spending per New Customer	$ 51.91	$ 51.91
Gross $ generated	$ 4,049	$ 727
Net $ generated (net of campaign costs)	$ 2,654	$ (183)
"Direct ROI"	**90.2%**	**-120.1%**
Average material cost (CGS)	$ 2,173	$ 467
Generated revenue--net of all costs	$ 481	$ (650)
Immediate ROI	**34.5%**	**-71.5%**
First year:		
Gross $ generated	$ 5,596	$ 1,222
Average material cost (CGS)	$ 3,350	$ 39
Net $ generated	$ 2,246	$ 1,182
One year ROI	**61.0%**	**29.9%**
Three year horizon:		
Gross $ generated	$ 9,172	$ 2,706
Average material cost (CGS)	$ 4,959	$ 707
Net $ generated	$ 4,212	$ 1,998
Three year ROI	**202.0%**	**119.6%**

REFERENCES

[1] D.J. Kridel and D.R. Dolk, "An On-line Marketing Consultant for Small and Medium Businesses", *Proceedings of the 32nd WDSI Conference*, Lihue, HI, April 2003.

[2] Castillo, D.G., D.R. Dolk, and D.J. Kridel, "GOST: An Active Modeling System for Costing and Planning NASA Space Programs", *Journal of Management Information Systems*, Winter 1991-92, Vol. 8, No. 3, pp. 151-169.

[3] Dolk, D.R., and D.J. Kridel, "An Active Decision Support System for Econometrics", *Decision Support Systems*, 7, 1991, pp. 315-328.

[4] Dolk, D.R., and D.J. Kridel, "Modeling Telecommunications Demand Analysis", *Interfaces*, March-April 1993, Vol. 23, No. 2, pp. 3-13.

[5] Dolk, D.R., and D.J. Kridel, "Towards a Symbiotic Expert System for Econometric Modeling", *Current Research in Decision Support Technology*, edited by R.W. Blanning and D.R. King, IEEE Computer Society Press, 1993, Chapter 7.