ABSTRACT

Companies in the early stages of adopting analytics are testing and exploring strategies that produce ROI to justify the capital investment of implementing analytics. Other companies farther down the adoption path have made that investment and found the required ROI early on. By doing so, the early investors created a repertoire of analytical business strategies and their tactical results. It is these companies that are working on the next level of analytics – global optimization at the macro and micro level. Examples are: yearly marketing spend optimized across all channels down to the most logical entity, campaign spend of a charity to reduce funding gaps or increase donations for topics of interest, reducing a bank’s risk exposure of each investment instrument on a position basis, or assisting a global company in ensuring that local product development can be adopted world wide. This last topic is the focus of this paper and is addressed using affinity models.

INTRODUCTION

Consider a global bank that has developed and tested a new product in one of its hundreds of local markets. The bank wants to leverage the results of the local product history in one market into a global launch to maximize its capital investment returns yield. As an early adopter of proven ROI analytics, the bank has a repertoire of tactical results over a two year period. Instead of using traditional account-level data such as purchase history or D&B data, the bank used the propensity-to-buy (PTB) models from 10 core product groups in North America.

Utilizing the asset of the scores history, the bank can implement a sales and marketing strategy for locally tested products that can be developed and deployed for the other regions without previous sales histories of the local product. The PTB scores history covers most of the core products across all divisions of the bank. Models are built and refreshed on a monthly basis and their results and scores stored. In combination with the current sales history of the local new product and the local PTB scores history, the organization is in a position to correlate purchases during the past year against the local core product scores history from the past two years to predict which global entity is likely to purchase the new product developed locally.

SOLUTION

The analytical benefit of having a reliable, trusted and implemented analytical strategy over at least a two year period allows for progression to the next level of analytics with measurable ROI. The inference engine leverages the scores history for independent variables and purchases of the new product over the past year for the definition of the dependent variable. The two-year time requirement is for proper statistical sampling: one year of data for model building and one year of data for model validation.

An affinity model is a statistical model that can predict a customer’s likelihood of buying a particular product or service from a particular product category. An affinity model analyzes a customer’s past purchasing patterns for a product or product category and predicts the customer’s affinity for a different product or product category. The affinity model correlates local sales utilizing a locally small, but globally vast number of logistic regression based PTB scores.

The scores history consists of thousands of models created at the local level for core products. An analytical dataset can be created by combining the leveled PTB scores from the 10 core product models. The analytical dataset consists of primary keys (for matching) and core product PTB scores from the local market where the new product was developed and tested. The file consists of buyers (events) of the new local product and non-buyers (non-events), also from the local market. Events are indicated by a value of one while the non-events are indicated by a value of zero, thus creating a binary variable which is the dependent variable that will be used to fit a logistic regression.

The analytical dataset is then divided into training and a validation file, in order to fit and validate the logistic regression. The logistic regression is utilizing the two-year scores history represented by the independent variables and actual purchases of the dependent variable. It is assumed that the underlying PTB models are built on consistent specifications exposed to identical business rules – without analyst bias. The goal of the logistic
regression is to create an affinity relationship between the adoption of a new product in a local market and then to deploy where the product has never been sold before.

To deploy the affinity model, a similar data set has to be assembled for scoring purposes. The main difference between the deployment dataset and the analytical dataset from which the logistic regression parameters are estimated is that the dependent variable does not exist in the scoring dataset since there is no purchase history for the locally developed product in the rest of the global market. Once the logistic regression is fit and the scoring code created, the affinity model can be applied to the rest of the global market, providing PTB scores globally. It is implicitly assumed that the characteristics of the local market are mirrored in global markets to which the affinity model is deployed.

Validation of the affinity model approach is done by using another product that has been sold in all markets. The analytical datasets are created in the exact same way as detailed above – with one major difference. While the scoring data set does not have a dependent variable, the validation dataset has a binary dependent variable that can be used to create lift charts, ROC curves and other performance measures.

Usually, when the results of the validation model are compared to the original PTB scores at the various local levels, the results may be slightly below the optimal propensity to buy model’s predictive ability. Yet when comparing the results to the next best solution – third party data – the results are typically better than third party-based PTB models since the corporation’s own customer base was used to build the affinity model. Hence, the need to purchase 3rd party data may be eliminated.

CONCLUSION

The innovation lies in leveraging the library of analytical models to build predictive insights EVEN when there is no purchase history.

Applications of the affinity model approach are far reaching. Corporations which acquire companies without a global presence can ensure that their acquisitions can be successful and provide maximum monetary return in the shortest timeframe possible.

Other opportunities are the reduction in reliance on third party data or, in combination with 3rd party data, create better predictions than with 3rd party data alone.

ACKNOWLEDGMENTS

We would like to thank all of our clients who let us help them push the analytic envelop, especially the support of the IBM corporation, a true visionary in the world of applied predictive analytics.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Matthias Kehder, CEO
Enterprise: Modern Analytics
Address: 1010 Turquoise Street, Suite 250-255
City, State ZIP: San Diego, CA 92109
Work Phone: 858-488-0771 x114
Fax: 858-488-0775
E-mail: mkehder@modernanalytics.com
Web: www.modernanalytics.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.