Available online at www.sciencedirect.com
SCIENCE DIRECT®
Decision Support Systems 37 (2004) 215–228
Decision Support Systems
www.elsevier.com/locate/dsw

An intelligent system for customer targeting: a data mining approach

YongSeog Kim[a,*], W. Nick Street[b]

[a] *Business Information Systems, Utah State University, Logan, UT 84322, USA*
[b] *Management Sciences, University of Iowa, Iowa City, IA 52242, USA*

Received 1 August 2002; accepted 1 December 2002

Available online 15 February 2003

Abstract

We propose a data mining approach for market managers that uses artificial neural networks (ANNs) guided by genetic algorithms (GAs). Our predictive model allows the selection of an optimal target point where expected profit from direct mailing is maximized. Our approach also produces models that are easier to interpret by using a smaller number of predictive features. Through sensitivity analysis, we also show that our chosen model significantly outperforms the baseline algorithms in terms of hit rate and expected net profit on key target points.

Keywords: Customer targeting; Data mining; Feature selection; Genetic algorithms; Neural networks; Ensemble

1. Introduction

The ultimate goal of decision support systems is to provide managers with information that is useful for understanding various managerial aspects of a problem and to choose a best solution among many alternatives. In this paper, we focus on a very specific decision support system on behalf of market managers who want to develop and implement efficient marketing programs by fully utilizing a customer database. This is important because, due to the growing interest in micro-marketing, many firms devote considerable resources to identifying households that may be open to targeted marketing messages. This becomes more critical through the easy availability of data warehouses combining demographic, psychographic and behavioral information.

Both the marketing [8,19,33] and data-mining communities [4,32,27,13] have presented various database-based approaches for direct marketing. A good review of how data mining can be integrated into a knowledge-based marketing can be found in [41]. Traditionally, the optimal selection of mailing targets has been considered one of the most important factors for direct marketing to be successful. Thus, many models aim to identify as many customers as possible who will respond to a specific solicitation campaign letter, based on the customer's estimated probability of responding to marketing program.

* Corresponding author. Tel.: +1-435-797-2271; fax: 1-435-797-2351.
E-mail addresses: ykim@b202.usu.edu (Y.S. Kim), nick-street@uiowa.edu (W.N. Street).

0167-9236/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0167-9236(03)00008-3

This problem becomes more complicated when the interpretability of the model is important. For example, in database marketing applications, it is critical for managers to understand the key drivers of consumer response. A predictive model that is essentially a "black box" is not useful for developing comprehensive marketing strategies. At the same time, a rule-based system that consists of too many if-then statements can make it difficult for users to identify the key drivers. Note that two principal goals, model interpretability and predictive accuracy, can be in conflict.

Another important but often neglected aspect of models is the decision support function that helps market managers make strategic marketing plans. For example, market managers want to know how many customers should be targeted to maximize the expected net profit or increase market share while at least recovering the operational costs of a specific campaign. In order to attain this goal, market managers need a sensitivity analysis that shows how the value of the objective function (e.g., the expected net profit from the campaign) changes as campaign parameters vary (e.g., the campaign scope measured by the number of customers targeted).

In this paper, we propose a data-mining approach to building predictive models that satisfies these requirements efficiently and effectively. First, we show how to build predictive models that combine artificial neural networks (ANNs) [37] with genetic algorithms (GAs) [18] to help market managers identify prospective households. ANNs have been used in other marketing applications such as customer clustering [1,16] and market segmentation [2,21]. We use ANNs to identify optimal campaign targets based on each individual's likelihood of responding to campaign message positively. This can be done by learning linear or possibly nonlinear relationships between given input variables and the response indicator. We go one step further from this traditional approach. Because we are also interested in isolating key determinants of customer response, we select different subsets of variables using GAs and use only those selected variables to train different ANNs.

GAs have become a very powerful tool in finance, economics, accounting, operations research, and other fields as an alternative to hill-climbing search algorithms. This is mainly because those heuristic algorithms might lead to a local optimum, while GAs are more likely to avoid local optima by evaluating multiple solutions simultaneously and adjusting their search bias toward more promising areas. Further, GAs have been known to have superior performance to other search algorithms for data sets with high dimensionality [28].

Second, we demonstrate through a sensitivity analysis that our approach can be used to determine the scope of marketing campaign given marginal revenue per customer and marginal cost per campaign mail. This can be a very useful tool for market managers who want to assess the impacts of various factors such as mailing cost and limited campaign budget on the outcomes of marketing campaign.

Finally, we enhance the interpretability of our model by reducing the dimensionality of data sets. Traditionally, feature extraction algorithms including principal component analysis (PCA) have been often used for this purpose. However, PCA is not appropriate when the ultimate goal is not only to reduce the dimensionality, but also to obtain highly accurate predictive models. This is because PCA does not take into account the relationship between dependant and other input variables in the process of data reduction. Further, the resulting principal components from PCA can be difficult to interpret when the space of input variables is huge.

Data reduction is performed via feature selection in our approach. Feature selection is defined as the process of choosing a subset of the original predictive variables by eliminating features that are either redundant or possess little predictive information. If we extract as much information as possible from a given data set while using the smallest number of features, we cannot only save a great amount of computing time and cost, but also build a model that generalizes better to households not in the test mailing. Feature selection can also significantly improve the comprehensibility of the resulting classifier models. Even a complicated model—such as a neural network—can be more easily understood if constructed from only a few variables.

Our methodology exploits the desirable characteristics of GAs and ANNs to achieve two principal goals of household targeting at a specific target point: model interpretability and predictive accuracy. A standard GA is used to search through the possible combinations of features. The input features selected