Customer Classification

A marketer often wants to know the segment to which a customer belongs, but lacks an individual identifier needed to do so. This dilemma limits the marketer’s ability to treat the customer differently with tailored messages and offers. Classification analysis addresses this by creating a model to assign customers to groups, based on a set of known characteristics of that individual. The model is built from data with both group membership and the set of individual characteristics, then applied to new, unclassified customers. This process yields a predicted group of segment membership. Offers can be customized to take into account segment priorities and responsiveness.

Classification analysis starts with a predefined set of customer groups which can be based on a single variable such as frequency of product use or interest in a new product concept. Alternatively, the customer groups may be derived from complex combinations of variables, such as “women shoppers who have certain product needs, channel requirements, and levels of dissatisfaction with available offers.” They may have been derived through analysis of frequency distributions, crosstabs, or multivariate procedures such as cluster analysis. The groups are given for this purpose, although some classification methods can be inverted to develop some insights into group structure itself.

Each customer is characterized by a set of variables that can be used to predict membership in the groups. The predictors could be demographic characteristics that indicate product use, attitudinal measures related to needs-based segments, or customer transaction histories, such as account information for various banking services. Classification analysis answers the question: For customers with unknown group membership, which groups does the individual resemble most, based on these measured characteristics?

The analysis first must determine whether customer group membership can be predicted at all, and second, which are the best predictors. If the predictors are integral to the group definitions, the rate of correct classification is likely to be high. For example, if the groups were derived from a cluster analysis of attitudinal statements, then the individual attitudinal statements should be good “predictors” of group membership, perhaps yielding a classification rate close to 100%. Alternatively, if the predictors are drawn from transaction histories that only weakly reflect attitudinal differences, then the rate of classification is likely to be low, perhaps only minimally better than random assignments.

The output of the analysis identifies the variables that significantly predict group membership and the nature and strength of the relationship. The analytical approaches include discriminant analysis, multinomial logit, Classification and Regression Trees (CART) and Chi-square Automatic Interaction Detection (CHAID). These methods vary in terms of the underlying model for linking predictors and groups such as linear, logistic, and binary splitting. They make different assumptions about interaction and nonlinear effects and have different capabilities to identify and estimate them.

Each method characterizes the statistical relationships differently, in coefficients or tree diagrams. This leads to different methods for applying the estimated model to new, as yet unclassified customers. Scoring a new set of customers could be done with a series of equations or by applying a logic tree.

All the methods provide a way to assess the fit of the model, including an analysis of the relationship between predicted and actual group membership from which the percent of customers that were correctly classified can be calculated. In some applications, there’s a compelling requirement to identify the smallest set of predictors that achieve a satisfactory level of prediction accuracy. Often, a cost is associated with additional predictors, such as data collection or analytical processing. Such a cost can be traded off against the prediction accuracy, seeing as more predictors generally yield a higher hit rate.

Estimated prediction accuracy from a test sample usually overstates the actual hit rate that would be experienced in a newly selected sample of customers. Holdout samples for internal cross validation and in-market testing can be used to estimate more realistically.

In some cases, no information is available about the size of the groups before the analysis. Suppose that clusters have been developed from a convenience sample known to be nonrandom. In other instances, however, there is prior knowledge of group sizes that can be taken into account. If three groups are known to be distributed 10%, 30%, and 60% in the total population, this information can guide the analysis itself. It generally will lead to different and presumably better prediction equations. The classification algorithm takes into account both the similarity of a new customer to each group and the overall likelihood of being in each group.

Applications

In practice, applications often are divided into two types: a) those in which the information on the predictor variables is known in advance, and b) those in which the information is elicited during an interaction with the customer.

Suppose a bank classifies its customers according to channel-use behavior groups, such as branch, ATM, and online. A clas-
sification model could be developed based on past transaction behavior with the bank, predicted from a set of variables such as the mix of branch compared to other types of transactions, tenure with the bank, and asset and loan balances. This model can be applied to the entire database of customers to identify channel affinity. Outbound customer contacts (by mail or telephone) can be targeted according to the predicted group membership of the customers. Offers can be tailored to these different customer groups based on revealed channel groupings. (Classification analysis also can be used as a research tool to establish segment identifiers for further sampling and analysis.)

On the other hand, suppose a computer company is developing a telemarketing capability to offer a range of hardware, software, and training products to inbound calling customers. To quickly classify customers into segments, a small number of questions can be asked during the call. Ideally, these questions would be drawn from a classification model that has significantly better than chance level of accuracy. The customer’s experience over the phone then can be tailored to his or her predicted needs. Different scripts would be used for different customer types, based on the customer’s specific priorities.

These applications of classification analysis have the potential to be leveraged through both the Internet and traditional channels. Outbound e-mail contacts could be informed by previous information about the customer, and segmenting customer lists can lead to more tailored messages and customer responses.

Inbound contacts to a Web site can be structured to reflect the type of customer. Customer behavior at the site can be used to classify the customer into groups that are treated in a particular way. Some Web design personalization software attempts to customize the experience through the sequence of ads, information, and offers presented to the customer.

One issue likely to increase in importance is the effect of privacy concerns. Customers may not want their personal information used for classification purposes and as a result, withhold information. Such a move may severely limit Web marketers’ ability to use individual customer information for tailoring and targeting.

Consequently, databases of customers may become more sparse; and some people will give information freely while others will not. Credit card companies and other intermediaries may hold the information, but never release it to other marketers. A given customer profile may be spotty, with information in some product categories (e.g., business travel-related), but not others (e.g., such personal items as apparel).

An implication for classification analysis is that the pattern of missing data could become an important variable. Customers may be differentiated along a dimension relating to their willingness to release information, and there even may be a price sensitivity associated with information release. Because customers are providing something of value, they may demand and receive compensation (money, information, or some other kind). A component of offer design could be the use of information gathered in the buying process.

**Exceptions**

While some situations call for customer classification, others may not. A neighborhood retail bookseller should have some idea about the different customers who visit the store frequently and have appropriate assortments for them, based on their interests. Alternatively, an online bookseller may not need to segment if each customer can be treated as an individual and find any book he or she might ever want. After all, any grouping scheme gives up some information about individuals as it aggregates them into segments. More importantly, the use of grouping could result in revenue and profit loss to competitors who better anticipate and satisfy customers’ individual priorities.

If there is extremely good information about each customer and/or if customers can design and select customized products, a need for grouping probably does not exist. While this may be a desirable goal, this situation doesn’t exist today for many providers. Even in information-intensive industries such as financial services, many firms don’t have the capabilities for one-to-one marketing. They lack the information technology to allow custom designed offers and the knowledge to design the best offers for each individual.

Still, technology is changing so rapidly that these constraints soon may be eliminated. In the near future, it may be unnecessary to know much at all about the customer in advance. The provider’s key capability will be to sense changing customer priorities from their in-market behavior and provide a flexible organization that can respond appropriately. Collaborative filtering engines which elicit customer preferences and/or behavior represent a step in this direction. In a sense, they push classification analysis down closer to the individual customer level.

Until these visions are more fully realized, a role for segmentation and classification will exist. Identification of customer segments will help marketing researchers understand the structure of the market and develop strategies for tailored value propositions. It might even become more important as marketers struggle to understand and treat customers who are reluctant to part with valuable information about themselves.