



"turning data into dollars"

Tom's Ten Data Tips – Oktober 2007

Affinity Analysis

Affinity analysis is an association technique based on the premise that the products consumers purchase, and the preferences they express, are indicative of their future behavior. By identifying product affinity patterns, one can predict future behavior to enhance service and promote cross-selling.

The links between customers and the products they choose can be modeled both uni-dimensionally, and one-dimensionally. In the former case, the order in which products were chosen (and preferences expressed) is disregarded for the purpose of making recommendations. Using one-dimensional affinity analysis, one also derives information from the *order* in which products are chosen. This is also referred to as sequence analysis.

1. Scaling And Sequence Analysis Are matching Sides Of The Same Coin

The cross-over from scaling to sequence analysis algorithms is not well documented in the (research) literature. Traditionally, algorithms for sequence analysis have been the domain of KDD, Machine Learning, and computational science communities. Scaling algorithms (Guttman, Rasch, Mokken, etc.) have been the realm of the social sciences. There is considerable synergy possible by integrating and blending these domains.

2. Collaborative Filtering Is A Special Flavor Of Affinity Analysis

Collaborative filtering is a set of techniques used to make predictions about interests and preferences. Technically, this is based on distance calculations between a set of ratings for an individual whose pattern is as close as possible to other people in hyperspace. Then preferences of similar users drive the ensuing recommendations.

Collaborative filtering has to do with explicit preferences or tastes that consumers have expressed, whereas affinity analysis is more generic. In affinity analysis individuals need not have overtly or voluntarily expressed *any* preference. The "collaborative" part stems from

consumers “voting” together on a preference, and thereby “collaborating” to achieve consensus.

3. Product Sequences Are Useful For *Prediction As Well As Insight*

The characteristic sequences of products that are discovered with one-directional affinity analysis can be used in two ways. As a *prediction* they indicate what is most likely to happen in the near future. That way, a likely “next step” can be recommended, in line with past behavior. Another application altogether is when the typical order in which purchases are made is used to *explain* what is happening in the market. The various ways that customer careers can traverse through the product assortment is a powerful means of segmenting customers.

4. Voting = Distance

The voting scheme that is employed in collaborative filtering, is mathematically represented by the distance function that is employed. Each customer is represented as a vector, and the vector that is closest in high-dimensional space will be the customer one is expected to resemble most. Although the act of voting appears rather different from calculating spatial distances, in collaborative filtering these two are conceptually equivalent.

5. There Are Many Distance Functions Possible

Intuitively, most people think of distance in terms of Euclidian distance, the “traditional” way in which we calculate distance in geographic space. But even there, we need not confine ourselves to the shortest distance between two points: one could think in terms of the actual path traversed rather than the shortest line connecting two points, or the *time* it takes to get from A to B.

Mathematically, there are many other distance calculations possible. Some examples are: city-block distance, Minkowski distance, geometric distance, Jaccard distance, Cosine distance, edit distance, and of course the most well known, the Euclidian distance. These variations correspond with the myriad possible voting schemes.

6. Sequence Analysis Provides Longitudinal Analysis On Cross-Sectional Data

An interesting feature of some scaling procedures (like Mokken scale analysis), is that the *sequence* of activities *within* the customer

lifecycle is inferred from making comparisons *across* customers. In this way, even if there is *no* historical data present, it is nonetheless possible to model longitudinal patterns.

7. Sequence Variables Are Powerful Predictors

Sequence variables that are the outcome of turning product- or service sequences into scales (using Mokken-scaling, for instance), can be used as a prediction in and of themselves, or as input variables in some other prediction. These scaling variables can be input variables in a regression, or some other algorithm that is used for predicting.

Because scaling variables are a compound of meaningful customer behavior patterns, they correlate with crucial variables. Therefore their "information density" tends to be very high (they summarize information across multiple variables). It is for this reason that they tend to show up with high weights in subsequent predictions being made.

8. Distance Calculation Is Computationally Efficient

One of the reasons why algorithms for affinity analysis are used so often as prediction engines for a website is that the recommendations after clicking through a link need to be made "on the fly". Therefore, the response time needs to be near immediate.

Due to the nature of the calculation, such millisecond response times are possible. Another advantage they have is that tweaking the calculation by altering the weights of the distance coordinates is easy to test. This allows for easy testing which voting scheme or distance function works best.

9. Compound Scale Scores Relate To LTV

The set of products a customer holds as a subset of the products that comprise a scale, determines his position on the scale. This is simply modeled as a Poisson distribution. The real number bears a clear relation with the likelihood of churn.

Customers hold a position on multiple scales, and each scale's subsequent transition can be related to a compound attrition probability (as a Markov chain). When the transition probabilities and product profitability are factored in, together with acquisition cost(s), a "winning" Life Time Value strategy can be devised.

10. Narrow Assortments Merit 'Unfolding' Of Services

The techniques of affinity analysis, and particularly collaborative filtering, were developed in settings where the number of possible recommendations is overwhelming: FMCG, book sales, etc. Such companies hold very large numbers of SKU's, and the techniques described in this newsletter worked particularly well because they kept the analysis tractable.

There are many possible trajectories that customers can traverse through an assortment. Besides product uptake, the way in which products are used can also be modeled. In that way, groups of customers who hold just one single product, can still be projected on multiple variables. For example: have they activated the product, how often do they use it, are they using collateral services, what is the balance, etc. This "unfolding" of services allows for sophisticated behavioral analysis.