DATA MINING CASE STUDY: RETAIL

Consumer Packaged Goods Company
Multi-Model Study

Exclusive Ore Inc.[1]

Exclusive Ore executed a multi-model data mining study for a leading packaged food product manufacturer. In its category this company is the leader, by far, in terms of both brand name recognition and market share.

There were several categories of information available for the study:

- Consumer survey and tastes. The CPG company did a nationwide taste test/survey, in which consumers were asked to taste various samples. The samples included competitor's products, company's products, and experimental recipes. Consumers rated each product on several scales including the presence of certain tastes as well as like and dislike. The consumers were also surveyed with respect to lifestyle, product recognition, product purchasing, etc.
- Expert tastings for flavors. The same products tasted by the consumers were also tasted by experts and ranked with respect to various flavors and effects on the palate.
- Chemistry data. The same products that had been tasted by consumers and experts were also subjected to a chemical analysis that measured each product in various chemical dimensions.

Exclusive Ore built several models using these data.

The first was to cluster the customer data according to tastes and survey responses. While no remarkable pattern emerged from the survey responses, the clusters that were based solely on taste test responses showed a strong geographical correlation. That is, although geographic location was not used to create the clusters (only taste test responses were used), the resulting clusters had a strong geographic bias. From a marketing perspective this suggests that tastes for this product differ regionally, and that there may be ways to do regional targeting based on taste. These models were built with a K-means clustering algorithm in Silicon Graphics MineSet.

The second model attempted to predict consumer cluster preferences based on expert taste ratings. That is, each customer was assigned to a taste cluster, and then that information, along with other demographic information was used to build a predictive model. The resulting models were fairly accurate, and in many cases the cluster assignment was an important attribute. The models were built using a Naive-Bayes classification algorithm from Red Brick Data Mine.

The third model attempted to establish a relationship between product chemistry and flavors. The chemistry data was merged with the expert taste data. Predictive models were built for selected flavors and sensory characteristics that were very accurate at predicting the presence of a

[1] To find out more about Exclusive Ore Inc., and its custom solutions to data warehouse, database and data mining problems, please go to www.exclusiveore.com or click here.

particular flavor or attribute based on chemical factors. These models were built using a neural network algorithm in ISL Clementine.

SUMMARY

Clustering and predictive modeling techniques produced models that, in the end, showed how chemistry can be used to predict flavors, that consumer flavor preference varies regionally, and that consumer preferences can be predicted based on flavors.

The combinations of these models enabled the CPG company to start thinking about a strategy that targets products regionally by taste, and that permits them to shape product tastes by controlling the chemistry.

To find out more about Exclusive Ore Inc. and its custom solutions to data warehouse, database and data mining problems go to www.exclusiveore.com or click here.

Exclusive Ore Inc.
PO Box 1024
Blue Bell, PA 19422
(215) 643-3110