American Society of Brewing Chemists

ASBC Annual Conference
June 4–7, 2017
17260 Harbour Pointe Dr
Fort Myers, FL 33908

Predicting Market Preference
From Reviews of Professional Tasting Panels on the Gastrograph Sensory System

D. Shah, R. Ahn and J. Cohen - Analytical Flavor Systems

Abstract

Most sensory panels at breweries do not contain a stratified sampling of the general or target population, so standard statistical methods cannot be employed in order to understand the perception and preferences of the average or target beer consumer. To project the preferences of professional panelists onto the preferences of the general population, reviews that were completed on the Gastrograph Sensory System were sampled in accordance with the tasting experience level distribution of the general population. The techniques LFDA (Local Fisher Discriminant Analysis) and PAM (Partitioning around Medoids) were used to maximize between-product similarity and minimize within-product similarity in flavor profile. The Random Forest method was then utilized to predict the distribution of perceived quality scores the general population would assign given any set of reviews, with built-in considerations for the style of beer and the tasting experience of the panelists.

Introduction and Data Collection

The standard sensory practices employed at breweries aim to predict whether a batch of beer is consistent enough to send out, to ascertain whether a new product will succeed in the general market, and to establish whether the brewery can improve their existing products. Unfortunately, it is difficult to predict whether the general population will taste an off-flavor in a deviant batch or whether they will enjoy the flavor profile of a new product. This is because the panels usually do not contain a robust and accurate representation of the general beer-drinking population at large. This study aims to more accurately determine the preferences of the beer-drinking population from reviews of professional panelists.

The Gastrograph Sensory System is a sensory platform that enables panelists to describe the product they are tasting across 24 broad-spectrum flavor attributes and somatosensations, supplemented by the specific reference flavors the panelists taste in any specific product. At the end of each review, the panelists are asked to assign a perceived quality score to that product.

The Gastrograph Sensory System was designed to be a comprehensive universal data collection interface where sensory panelists can rate the intensity of 24 broad-spectrum flavor attributes and somatosensations from 0 to 5: 0 being not present at all, 1 being at the edge of perception, and 5 being the highest possible intensity in a product. Reference flavors are marked within broad-spectrum flavor attributes so that their semantic meaning is properly categorized. At the end of each review, panelists assign a perceived quality score indicating their best assessment of the overall quality of the beer, regardless of personal preference, on a 7-point normally distributed scale.

Feature Engineering

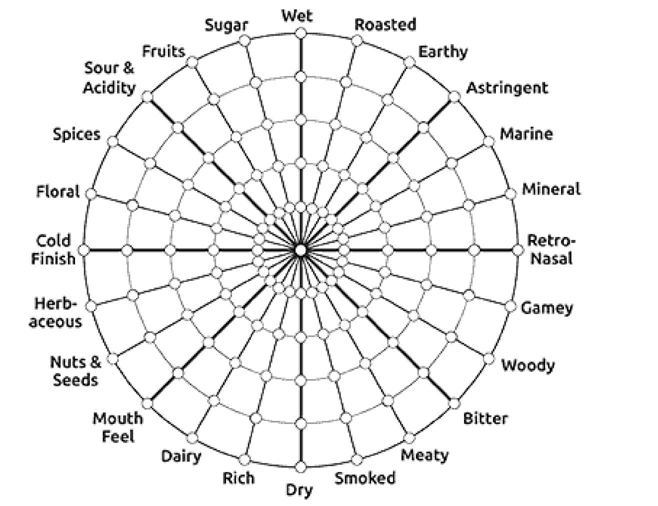

Figure 1: A visualization of the Gastrograph. The Gastrograph application by Analytical Flavor Systems measures 24 flavor variables as well as an overall perceived quality. All reviews in this study were done using this flavor wheel.

Each panelist has an account on the Gastrograph Sensory System allowing the tracking and controlling of their respective biases and flavor sensitivities. The users are assigned a per-review experience score indicating how likely it is that their review is an accurate representation of the product in question.

As cognizant tasting becomes more frequent, perception changes, and preferences evolve. The ability of the panelists to identify subtleties and nuances in flavor that one would not have identified in the past is quantifiable. The proficiency of any given panelist is predicted by their ability to identify the subtleties of the flavor profile in a product across multiple somatosensory attributes. Each panelist is assigned a dynamic experience score which is updated upon every new review submission.

The Algorithm and Models

Random Forest is a machine learning algorithm that uses hundreds of decision trees, each with a subset both of the variables and of the observations in the input data, to both predict the output perceived quality and learn the variables of most importance. Decision trees are a set of rules used to classify the data into categories. In this case, the categories are the different possible perceived quality scores on a scale of 1 to 7, 7 being the highest. The variables chosen to categorize each observation are the ones that maximize information gain when used as a splitter in the decision trees.

The random forest algorithm is run with the inputs being the flavor profiles, the reviewer's experience score, and the style classification of each review of beer, and the output being a probability distribution of possible perceived quality scores the average beer drinking market would give this beer.

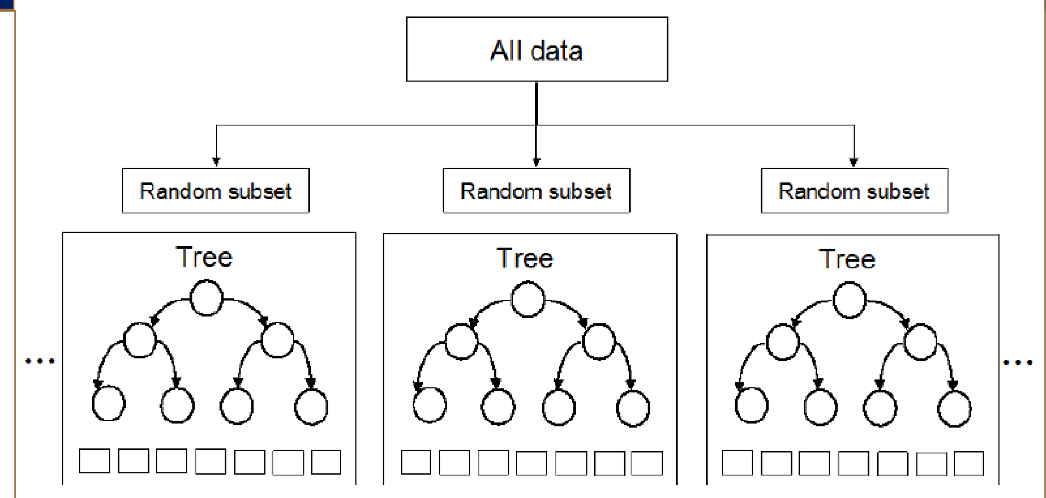

Figure 2: A visualization of the Random Forest model, which employs hundreds of decision trees to vote on the outcome of an observation. In this case, the decision trees decide what Perceived Quality score a review of a beer should be classified under, and the final result is calculated by a majority vote regarding the Perceived Quality score, with each decision tree getting one vote.

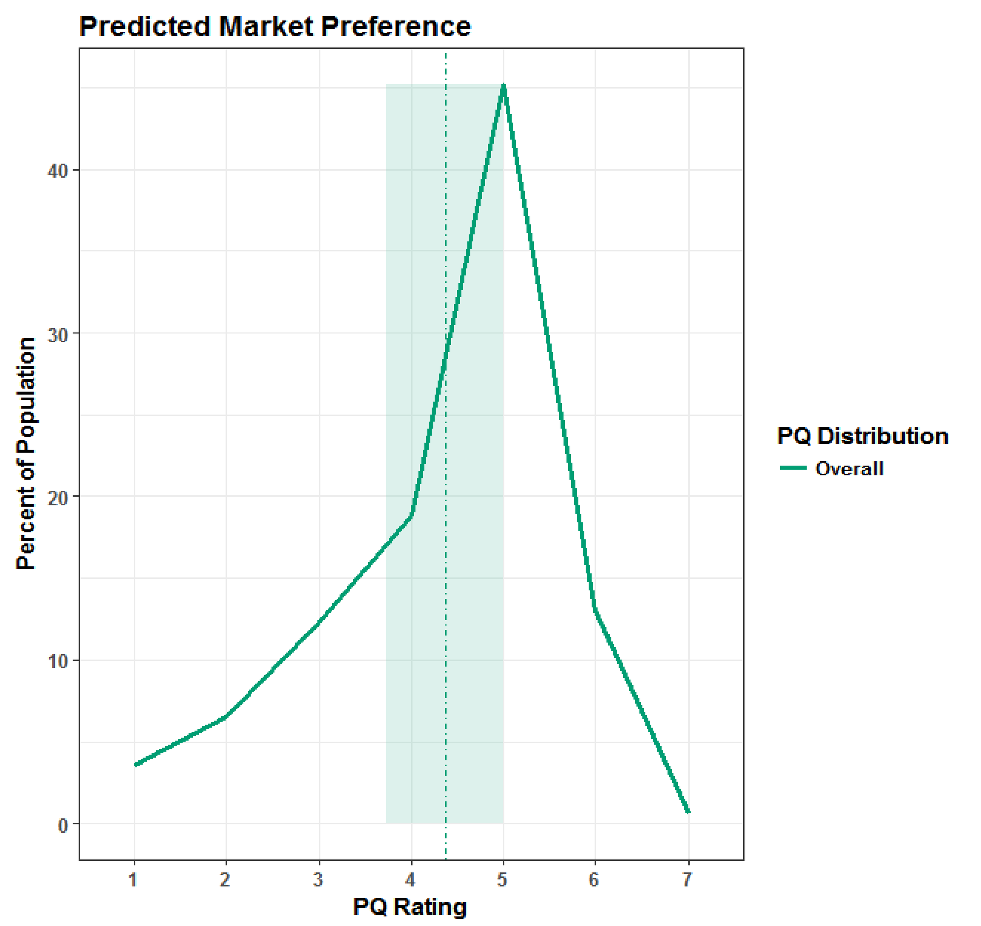

Figure 3: A visualization of the sample Predicted Market Preference distribution output by our model. Most people would give this sample beer a 5 out of 7 on the perceived quality score, and the average perceived quality score this beer would be assigned is a 4.4 out of 7.0.

Results and Conclusions

The output of this model is a distribution of perceived quality score probabilities for any input review(s). These probabilities are calculated from the proportion of decision trees that voted for this perceived quality score. These probabilities are interpreted as the predicted market preference of the input reviews. This is a valid interpretation because of the correct distribution of experience scores in the input dataset and the understanding that the decision trees model various segments of the population that do not care about each flavor attribute equally. Because the style classification of each product is included in the model, the model learns to discriminate between beers in a more robust way: the combinations of flavors that make a stout achieve a high perceived quality score are different than the ones that make an IPA achieve a high perceived quality score. The model achieves a mean absolute error of 0.98 perceived quality units on our set of test data, which is lower than the standard deviation of 1.3 perceived quality units in our dataset. This amount of error is expected to remain due to the inherent variability of perceived quality, especially when taking into account less experienced reviewers.

The result is a robust model of the distribution of perceived quality scores that the general population would taste, given any number of input reviews. Since reviews are classified by experience score, the model is unaffected by poor quality reviews on the Gastrograph Sensory System.

The percentage of the population that would prefer a specific batch of a product versus any other batch or versus the overall product is now able to be quantitatively measured. Other applications come from subsetting the population by gender and quantifying that difference in predicted market preference for any product, as well as formulating new products with a target flavor profile that results in successful market performance.

Resources and Acknowledgements

Research Funded by Analytical Flavor Systems, Inc.

"Random Forests - Classification Description". *Stat.berkeley.edu*. N. p., 2017. Web.

Leon, Arturo and Jason Cohen. "Classifying Reviewers By Experience: Modeling The Evolution Of Reviewer Experience". GastroNexus. N.p., 2015. Web. 27 Jan. 2017.