Reprinted from IFPress (2020) 23(1) 3,4

Predicting New Segment Opportunities

Daniel M. Ennis, William J. Russ, and Benoît Rousseau

Background: Computer-aided design of new products is made possible by the confluence of a number of factors in advanced analytics and computational speed. Recent interest in AI and its sub-branch, Machine Learning, has drawn attention to this area. Although the interest in this area and the goals have remained relatively constant over several decades, we are much closer to achieving these goals today than ever before. In a recent technical report[1], we discussed how Linear Programming can now be used in combination with Graph Theory and TURF to create optimal product bundles from unimaginably large numbers of possibilities. Another example of interest, in this report, is *Unfolding* which is a computationally intensive analytic technique based on a great idea proposed by Clyde Coombs in the 1950s. Unfolding of liking or purchase interest involves the determination of ideal and product coordinates in a low-dimensional drivers of liking space. This method is a key to understanding what drives consumer liking and hardly anyone would disagree with the process model that it specifies or the objective it attempts to accomplish.

After about a half-century of frustrating failure to achieve nondegenerate solutions to the unfolding problem (degenerate solutions are ones with little or no interpretability) the issue was finally resolved. One of the successful methods for unfolding, developed in 2001, is **Landscape Segmentation Analysis® (LSA)** which is based on a probabilistic similarity model involving latent product and individual ideal point coordinates[2,3,4]. Other methods have also been proposed since then[5]. Once an LSA map is generated, it can be used as a computer-aided design tool to predict the success of new products based on available descriptive sensory data, or other descriptive data, without the expense of conducting a consumer test. Since the method also provides individual ideal point locations, LSA facilitates segment identification and can be synergistically linked to newer methods of analysis, such as Machine Learning, to characterize segments[6,7]. This facility for synergistic analytics makes LSA an attractive candidate to link to newer and older methods such as Conjoint Analysis, MaxDiff, and Decision Trees or Random Forests. In this technical report, we will explore the capability of using LSA as a computer-aided design tool and show how it can be used to predict the performance of new prototypes or marketing concepts and to associate them with emerging segments.

Scenario: You work for a company with historical roots in alcoholic beverages and there are emerging trends that indicate that there is consumer interest in new beverage categories involving sensory, health and social benefits which derive from the original category. Your competitors and new players are very active in this area and you see declining revenues from your existing product lines. For some time, you have been creating an extensive database of internet-based characterizations of fruit-flavored beverages, some traditional and some more exotic. These characterizations include perceived imagery, health, and sensory variables and, using Graph Theory, you previously explored bundles of flavors, concepts, and benefits of fruit-flavored beverages[1]. Recently a colleague evaluated liking for twelve of the more commonly identified fruit-flavored product names from your database among 1000 consumers and she provided you with this data. The number of potential prototype alternatives in your database is extensive and you are looking for a method to accurately explore their predicted acceptability broadly and to certain consumer segments in particular. Using LSA, you would also like to see opportunities suggested by the structure of the LSA space. Your objective is to make a recommendation to management regarding product opportunities and segment characterization[6,7] to develop a marketing plan for alternative new fruit-flavored beverages that the company is not currently marketing.

Unfolding as a Hedonic Process Model: The idea behind LSA is that there is a latent drivers of liking space and that item points and individual ideal points can be located in this space. This process involves fitting a probabilistic similarity model to the data which solves the degeneracy problem identified earlier. Once this has been accomplished, variables that describe the space can be accommodated easily by regressing these variables into the space. The model includes parameters for item variance and individual subject biases, such as a tendency to rate high or low. A great advantage of this approach is that unfolding does not depend on the explanatory variables, as it does with external preference mapping, and identifies ideal points rather than ideal directions, as occurs with internal preference mapping.

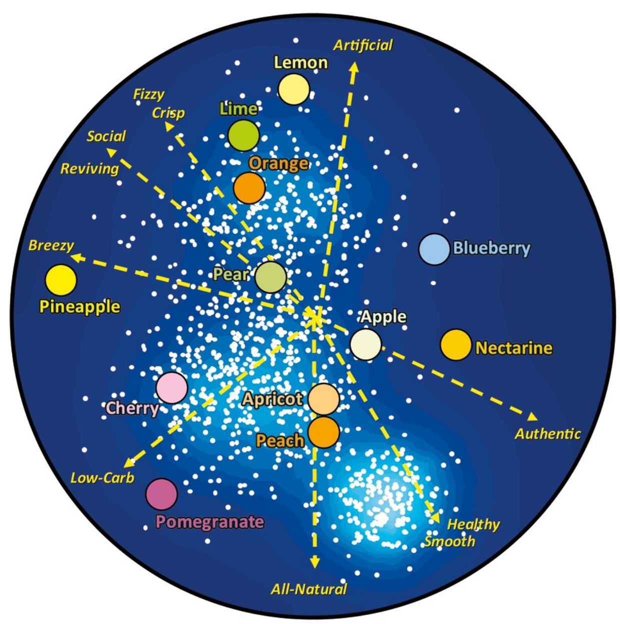

Figure 1. LSA solution to liking ratings for twelve fruit-flavored alcoholic beverages. Individual ideal points of consumers are shown as white dots on the map, which show three distinct segments.

Reprinted from IFPress (2020) 23(1) 3,4

Once the LSA space has been discovered and described, liking for new items that were not in the original dataset, and to which consumers they appeal, can be predicted. In other words, a computer-aided design tool has been created and the results can be linked to other methods of analysis synergistically.

Application to the Beverage Set: The LSA map will have about a quarter of the number of observations in parameters, but it is still necessary to check for stability. Cross-validation can be used both in creating the map and in testing the new product placement predictive ability. This involves dividing the data set into groups called folds; it is typical to use 5 or 10 folds. A map is then created for each combination of groups with one group left out. For example, with 5 folds, there would be five maps, each using four of the groups. You visually inspect these maps and if they are all very similar, it indicates stability. The maps can also be compared analytically using inter-point distances. You assess the predictive ability by placing existing products using only part of the data. Finding the placed products to be near their actual locations, the map can be considered predictive. Additionally, the variance in the predicted locations was used to calculate regions of placement to quantify the uncertainty in future predictions.

Figure 1 shows the results of the LSA analysis on the twelve beverage items. Most of these beverages appeal to two large segments apparently separated by citrus and non-citrus flavors. However, the analysis also reveals a new, minor segment that has not been fully satisfied by the existing flavors.

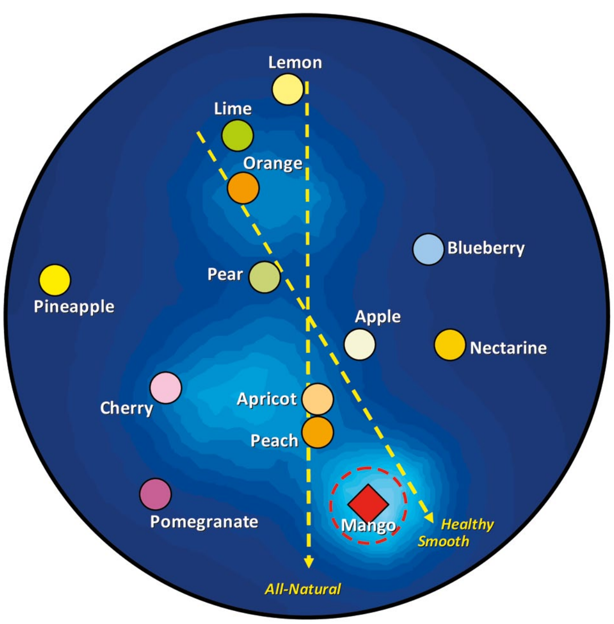

Figure 2. The LSA map with the location of the mango fruit flavored beverage predicted to fall in a area of uncertainty within the region located in the southern-most segment on the LSA map.

The closest possibilities are peach and apricot, but these flavors are not optimal. The new segment appears to be associated with imagery from health and natural descriptors. A search through your database locates the optimum fit as shown in Figure 2. The fruit flavor most closely linked to the new segment's interest is mango, which was not present in the original set of items tested by your colleague. Segmentation characterization[6,7] using Machine Learning is one method you plan to use to describe the new segment in term of psychographic and demographic variables.

Conclusion: A great deal of progress has been made in developing and using new analytic tools that were simply not available even a decade ago. Examples include the use of Graph Theory and Linear Programming for TURF-type problems, but it also includes tools from AI, such as Machine Learning, and solutions to the degeneracy problem in multidimensional unfolding. Some of this progress has been due to improved computational speed and greater memory capacity, but most of the computational improvement comes from superior and more insightful algorithms.

Insights from using analytic tools synergistically has made rapid progress possible in a wide number of applications, sometimes based on 70 year-old ideas that are still relevant today, and this situation will continue into the future. In this technical report, we showed how a computer-aided design tool can be used to establish the existence of a new segment and also to determine the item offering that would appeal to that segment. We also showed how advances in computational methods has made it possible to solve a problem that was of interest a generation ago, but still captures our imagination today.

References

1. Ennis, J. M. and Russ, W. J. (2017). Large TURF problems: Finding custom solutions. In Ennis, D. M., Rousseau, B., and Ennis, J. M. (Eds.) (2018). *Tools and Applications of Sensory and Consumer Science*. (pp. 132-133), Richmond, VA: The Institute for Perception.
2. Ennis, D. M., Palen, J., and Mullen, K. (1988). A multidimensional stochastic theory of similarity. *Journal of Mathematical Psychology,* **32**(4), 449-465.
3. Ennis, D. M., and Johnson, N. L. (1993). Thurstone-Shepard similaity models as special cases of moment generating functions. *Journal of Mathematical Psychology,* **37**(1), 104-110.
4. Ennis, D.M., and Rousseau, B. (2004). Motivations for product consumption: Application of a probabilistic model to adolescent smoking. *Journal of Sensory Studies,* **19**(2), 107-117.
5. Busing, F., Groenen, P., and Heiser, W. (2005). Avoiding degeneracy in multidimensional unfolding by penalizing on the coefficient of variation. *Psychometrika,* **70**(1), 71–98.
6. Russ, W. J. and Ennis, J. M. (2018). Characterizing sensory segmentation using machine learning. *IFPress,* **21**(4) 3-4.
7. Russ, W. J. and Ennis, J. M. (2019). A three-step approach to characterizing consumer segmentation via machine learning. Pangborn Conference paper, Edinburgh, Scotland. Available at https://www.ifpress.com/presentations.