IDENTIFYING HIGH POTENTIAL BIOMASS-HOME-HEATING CUSTOMERS: A BAYESIAN CLASSIFICATION[1]

Adee ATHIYAMAN
Illinois Institute for Rural Affairs,
Western Illinois University, Macomb, USA

E-mail: A-Athiyaman@wiu.edu

Abstract:
This paper presents a Bayes' classification computer application that would help biomass managers optimize their marketing decision making. Programmed in Mathematica, this decision tool would help managers understand the size of the high potential market at the US county level: number of households that would be receptive to a telemarketing / direct marketing campaign about pellet heating appliances, for example.

Key words: Market Segmentation, Biomass, Residential Heating, Interactive App, Decision Theory

Introduction

A recent McKinsey Global Survey of marketing executives suggests that the most pressing competitive challenge for the executives is gaining customer insights to drive sales (Davis and Freundt, 2011). This is not surprising given that the present-day business environment consists of:

- ✓ multiple customer segments – our multicultural society and the more divided income groups have created multiple customer segments;
- ✓ multitude sub-brands and line extensions that target these segments;
- ✓ multiple media (for example, web banners, magazine ads, and Facebook pages), and
- ✓ multiple distribution touch points (Internet, product re-sellers, big-box retailers, third-party tele-sales providers, etc.).

This complex environment calls for new marketing capabilities such as data-management and analytics - skills that marketing executives say that their companies do not have (Breuer et al, 2013). While it is not uncommon for companies to outsource analysis, it is now well known that managers seldom act on numbers that they don't fully understand (Wierenga, 2002).

In this paper, we outline a relatively straight-forward computer program that would help biomass businesses make better decisions about customers. We believe that the biomass industry needs such "technical" assistance since only a mere 10% of the industry's market potential has been tapped as at date (Athiyaman, 2014).

Classification Algorithms

Linear discriminant function (LDA)

LDA is often employed to discriminate among groups of objects such as customers and products (Sharma, 1996). It is similar to a regression function with a nominal dependent variable:

$y = \lambda_1 x_1 + \cdots + \lambda_n x_n$, where, the λs are weights to be applied to the *x* properties of the objects.

For two groups, the λs are found by maximizing an objective function such as the following:

$$MaxG = \frac{(\overline{y_1} - \overline{y_2})^2}{\sum_{\substack{i=1,2 \\ j=1,..,n}} \left(y_{ij} - \overline{y_i}\right)^2}$$

The index number *y* is used to allocate a new object to be classified to the class whose value of the function is the closer.

Problems with the use of LDA stem from its stringent measurement requirements - continuous or ratio scale property of the object being measured is presupposed (Lilian and Rangaswamy, 2004). In marketing, information on customer characteristics (for example, home ownership, gender, etc.) is essentially nominal; discrete measures in statistical terminology (Maddala, 1986). Hence, it is our contention that the Bayesian classification procedure should be employed for classification tasks in marketing.

Classification based on Probability Axioms

We illustrate the procedure using data derived from a survey of 25,000 households conducted during late 2014 (see Athiyaman (2015b) for details about the survey). Our interest is in building a customer segmentation scheme for the biomass residential heating industry. Specifically, we want to segment households in the US into "high" versus "low" potential customers.

The survey data suggests the following attributes as important determinants of purchase:

Dichotomous Attribute	**Correlation with Purchase Intention**
x_1 –Attitude	0.66 (Negative = 0; Positive = 1)
X_2 – Biomass heating is low cost	0.44 (No = 0; Yes = 1)
X_3 – Education of the head of household	0.41 (High School = 0; College = 1)
X_4 – Year Home was Built	-0.38 (1979 or older = 0; 1980 and newer = 1)

For a dichotomous purchase intention measure (high likelihood of purchase (H) versus low likelihood (L)), the probabilities of the two groups having or not having each of the above attributes are shown in Table 1. To elaborate, in Table 1 the attribute pattern *X* = 1 includes all four dichotomous attributes each with level = 1. The reverse is true for attribute pattern *X* = 16: each of the four attributes has level = 0.

Table 1. Attribute Possession and Conditional Probabilities: Estimates from Survey Responses

	Attribute Patterns (levels)				Conditional Probability	
X	x_1	x_2	x_3	x_4	***P*(*X*|High)**	***P*(*X*|Low)**
1	1	1	1	1	0.28	0.15
2	1	1	1	0	0.42	0.33
3	1	1	0	1	0.06	0.01
4	1	1	0	0	0.11	0.05
5	1	0	1	1	0.04	0.06
6	1	0	1	0	0.05	0.05
7	1	0	0	1	0.01	0.01
8	1	0	0	0	0.02	0.02
9	0	1	1	1	0.00	0.07
10	0	1	1	0	0.01	0.04
11	0	1	0	1	0.00	0.02
12	0	1	0	0	0.00	0.01
13	0	0	1	1	0.00	0.06
14	0	0	1	0	0.00	0.11
15	0	0	0	1	0.00	0.01
16	0	0	0	0	0.00	0.01

Note: Total number of responses = 2717

Table 1 suggests that 99% of all customers classified as "high potential" possess positive attitude towards biomass heating. Furthermore, 42% of all "high potential" customers are educated beyond high school, live in older homes, and believe that biomass heating is low cost. In fact, the age of home matters as much as 14 points in predicting customer's purchase likelihood; older the home, higher is the purchase probability. How could a biomass company make use of this information to fine-tune its marketing?

Decision Aid

Assume that a biomass company is planning to advertise nationally. To minimize opportunity loss, the firm employs decision theory. Specifically, based on US Census data (American Community Survey (ACS) Selected Housing Characteristics) the firm estimates the prior probability of a randomly chosen household being classified as high sales potential, and applies this number to opportunity loss estimates to gain insights into expected payoffs (Table 2).

Table 2. Opportunity Losses: Household Level Illustration

Decision	State of Household		Opportunity Loss
	High Potential	**Low Potential**	
D_1 – Advertise	\$0	\$42	$E(Loss) = \$0 \times .1 + \$42 \times .9 = \$37.8$
D_2 – Do not advertise	\$1500	\$0	$E(Loss) = \$1500 \times .1 + \$0 \times .9 = \$150$

Note:

(i) ACS estimates of priors: proportion of households using biomass for heating = 0.1;

(ii) Advertising to a low potential household will cost the company \$42. This estimate was derived from averaging national telemarketing costs and national direct marketing costs per household (source: WebpageFX);

(iii) Not advertising to high potential is assumed to cost the company, actually the industry, sale of a pellet stove which is estimated at \$1500; "top-line" or revenue loss for the company from lost sale for a stove is \$1500 (US Environmental Protection Agency).

If the company were to choose an act or decision on the basis of information given in Table 2, then action "D_1 – Advertise" will be chosen (the company is minimizing the maximum loss). However, if households could be screened for possession of attribute patterns given in Table 1, then a much more precise decision could be made about advertisements.

To elaborate, the joint probabilities of each of the 16 attribute patterns given in Table 1 can be computed by weighting the prior probabilities of **H**igh/**L**ow potential (p=0.1 / p=0.9) with the conditional probabilities given in Table 1. Since "**H**" and "**L**" constitute a partition over the set of households, the marginal probabilities represent the sums of the joint probabilities (Green, 1964). Finally, the posterior probabilities are the results of the application of Bayes' theorem:

$$p(H|X) = \frac{p(H)\times p(X|H)}{p(H)\times p(X|H)+p(L)\times p(X|L)}$$

Based on the posterior analysis, the company can determine which segments of households should be targeted by advertisements. The critical posterior probability that informs us of the relevant attribute pattern(s) to target is:

$$C^* = \frac{\text{Opportunity Loss } (\mathrm{D}_1)}{Opportunity\ Loss\ (D_1) + Opportunity\ Loss\ (D_2)}$$

In other words, the decision rule for advertising is:

$If\ p(H|X)\ > C^*, take\ action\ D_{1;}$
$If\ p(H|X) < C^*\ , take\ action\ D_{2;}$
$If\ p(H|X) = C^*, take\ either\ action\ D_1\ or\ D_2$

Interactive Application

Figure 1 shows the interactive web application that implements the marketing decision model. It is available online at www.instituteintelliegence.com. To illustrate, suppose that a biomass company wants to see the number of households in McDonough County, Illinois, that would be receptive to a telemarketing / direct marketing campaign about pellet heating appliances. In the starting page, the analyst / decision maker will specify the county that is of interest – McDonough in this instance. This will result in a listing of number of owner-occupied households in the county. Then, the analyst will specify the attribute pattern that is of interest: assume that it is one of the first six rows of Table 1: "1, 0, 1, 0", for example. This specification would result in a listing of the proportion / number of households in the county that are associated with the given pattern (the home owner has positive attitude towards biomass heating, believes that it is inexpensive, has a high school education, and lives in a home that is classified as "new"). In addition, the app will also specify the number of households that satisfy the decision criterion,

$If\ p(H|X)\ > C^*, take\ action\ D_1\ (advertise)$

Appendix 1 shows that that this classification procedure or decision model yields 93% savings to the company than the "generic" decision theory framework given in Table 2. Although privacy concerns prevent us from specifying the addresses of these households, the app is being developed to “map” the geographical location of majority of these households (for example, one or more of the 10 census tracts in McDonough County).

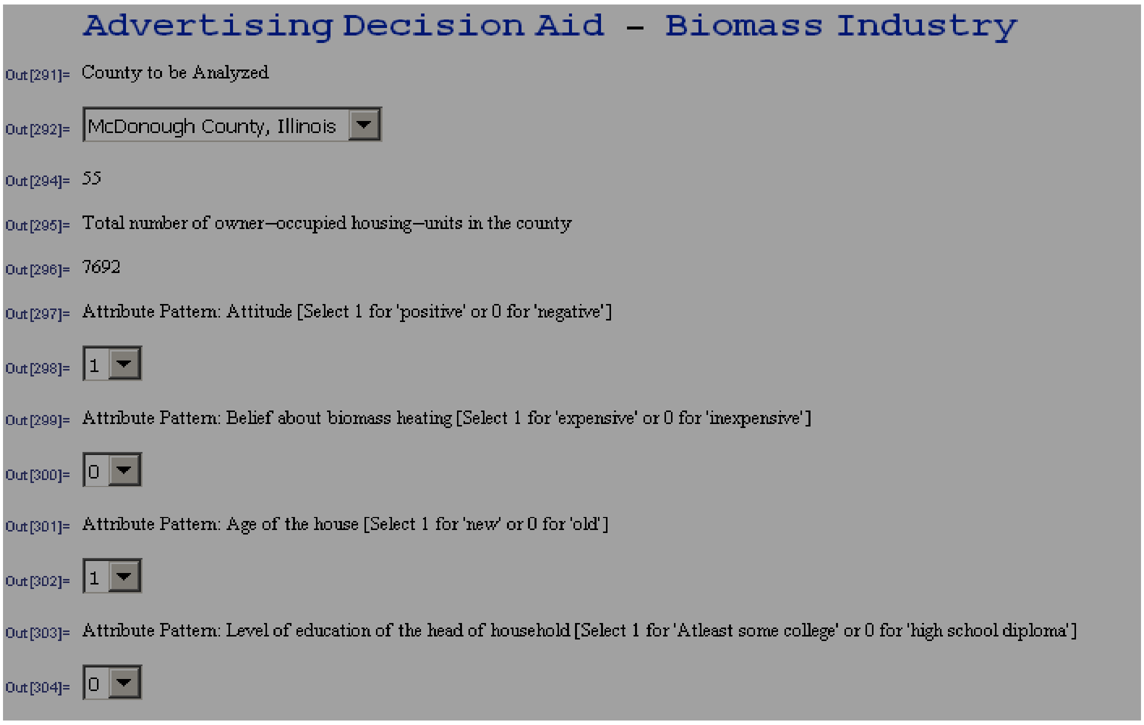

Figure 1. The Classification Application

Summary and Conclusion

Marketing in this age of micro segmentation requires skills such as knowledge of distributed computing and machine-learning techniques. Most companies do not possess these skills hence the demand for “canned” software that would aid marketing decision making. This paper presents a Bayes' classification computer application that would help biomass managers optimize their marketing decision making. This decision tool would help managers understand the size of the high potential market at the county level: number of households that would be receptive to a telemarketing / direct marketing campaign about pellet heating appliances, for example.

Earlier, Athiyaman (2015a) highlighted that company induced “push” marketing influences product “purchase” or “closure”, albeit at the final stages of customer decision sequence. Given this finding, it is essential that the biomass industry utilizes analytical tools such as the one presented in this paper for marketing-program optimization. As aptly observed by Aron and van den Direst (2014), use of data-driven decision tools is a necessary condition for being successful in todays' marketplace.

References

1. Arons, M.D.S., van den Driest, F. and Weed, K. **The ultimate marketing machine,** Harvard Business Review, Vol. 92, July/August 2014, pp. 54-63
2. Athiyaman, A. **Tapping Unstructured Data to Gain Insights into Consumer Purchases of Biomass Residential Heating Appliances: An Exercise in**

Big Data Analytics, Biomass Research Report 3, Working Paper, Macomb, IL: Institute Intelligence, 2015a
3. Athiyaman, A. **Analysis of Consumer Actions: A Descriptive Study of the Biomass Residential Heating Industry,** Biomass Research Report 5, Working Paper, Macomb, IL: Institute Intelligence, 2015b
4. Athiyaman, A. **Market Potential for Residential Biomass Heating Equipment: Stochastic and Econometric Assessments,** Biomass Research Report 1, Working Paper, Macomb, IL: Institute Intelligence, 2014
5. Breuer, P., Forina, L. and Moulton, J. **Beyond the Hype: Capturing Value from Big Data and Advanced Analytics,** McKinsey & Company, 2013
6. Bogdanović, J., Vladana Lilić, A. and Gijić, N. **Implementation Of Advanced Analytics In Consumer-facing Companies**, International Journal of Economics & Law, Vol. 10, 2014, pp. 35-42
7. Davis, C. and Freundt, T. **What Marketers Say About Working Online, McKinsey Global Survey Results,** McKinsey Quarterly, November, 2011
8. Green, P. E. **Bayesian Classification Procedures in Analyzing Customer Charcateristics,** Journal of Marketing Research, May 1964, pp. 44-50
9. Lilien, G.L. and Rangaswamy, A. **Marketing Engineering: Computer-assisted Marketing Analysis and Planning,** DecisionPro, 2004
10. Lilien, G.L., Rangaswamy, A., Van Bruggen, G.H. and Wierenga, B. **Bridging the Marketing Theory–Practice Gap with Marketing Engineering,** Journal of Business Research, Vol. 55, 2002, pp. 111-121
11. Maddala, G.S. **Limited-dependent and Qualitative Variables in Econometrics,** No. 3, Oxford: Cambridge University Press, 1986
12. Sharma, S. **Applied Multivariate Techniques,** New York: John Wiley & Sons, 1996
13. Wierenga, B. **On Academic Marketing Knowledge and Marketing Knowledge that Marketing Managers Use for Decision-making,** Marketing Theory, Vol. 2, No. 4, 2002, pp. 355-362
14. * * * **DP04: Selected Housing Characteristics"American Community Survey,** Accessed February 6, 2015. http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_13_5YR_DP04&prodType=table
15. * * * **Pellet Stove Fact Sheet,** US Environmental Protection Agency, Accessed February 6, 2015, http://www.epa.gov/burnwise/pdfs/PelletStoveFS08-04-11.pdf
16. * * * **The Cost of Advertising Nationally Broken Down by Medium,** WebPageFX, Accessed February 6, 2015, http://www.webpagefx.com/blog/business-advice/the-cost-of-advertising-nationally-broken-down-by-medium/

Appendix 1.

The Rationale for Using Customer Attribute Patterns in Marketing Decisions

In this section, we empirically demonstrate that taking action based on a posterior analysis for each subgroup of sales prospects is optimal than deciding on a prior analysis alone (Table A1).

Table A1. Utility or Benefits of Analyzing Customer Attribute Patterns

Attribute Pattern "X" (See Table 1 in Text)	Posterior: P(H|X)	P(L|X)	Marginal P(X)	Expected Opportunity Loss (D2: Do not advertise)	Expected Opportunity Loss (D1: Advertise)	Optimal Decision
1	0.173484	0.826516	0.16	41.86046512	5.584091	D1
2	0.125185	0.874815	0.34	63.6627907	12.45682	D1
3	0.362438	0.637562	0.02	8.720930233	0.429545	D1
4	0.193563	0.806437	0.06	16.56976744	1.932955	D1
5	0.06747	0.93253	0.06	6.104651163	2.3625	D1
6	0.102088	0.897912	0.05	7.848837209	1.932955	D1
7	0.102088	0.897912	0.01	0.872093023	0.214773	D1
8	0.131638	0.868362	0.02	3.488372093	0.644318	D1
9	0.00	1	0.06	0	2.577273	D2
10	0.015983	0.984017	0.04	0.872093023	1.503409	D2
11	0	1	0.02	0	0.644318	D2
12	0	1	0.01	0	0.214773	D2
13	0	1	0.05	0	2.147727	D2
14	0	1	0.10	0	4.295455	D2
15	0	1	0.01	0	0.429545	D2
16	0	1	0.01	0	0.429545	D2

The opportunity loss of $339 is 45% less than the decision based on a prior analysis alone (expected opportunity loss for a prior-only decision is $618).

[1] **Acknowledgement**
State funds for this project were matched with Federal funds under the Federal-State Marketing Improvement Program of the Agricultural Marketing Service, U.S. Department of Agriculture.