DATA MINING FOR PREDICTING CUSTOMER SATISFACTION IN FAST-FOOD RESTAURANT

BAYU ADHI TAMA

Department of Information Systems, Faculty of Computer Science, Sriwijaya University

E-mail: bayu@unsri.ac.id

ABSTRACT

Managing customer satisfaction has become a crucial issue in fast-food industry. This study aims at identifying determinant factor related to customer satisfaction in fast-food restaurant. Customer data are analyzed by using data mining method with two classification techniques such as decision tree and neural network. Classification models are developed using decision tree and neural network to determine underlying attributes of customer satisfaction. Generated rules are beneficial for managerial and practical implementation in fast-food industry. Decision tree and neural network yield more than 80% of predictive accuracy.

Keywords: *customer satisfaction, classification model, data mining, fast-food restaurant, rules extraction*

1. INTRODUCTION

With a highly saturated and competitive market in fast-food industry, restaurant operators ought to be more selective in running their business strategies. Due to the rapid growth of above-mentioned industry recently in Indonesia, fast-food restaurant was chosen in this study. Various brand names expand their networks into cities and offer a wide range of menu, prices and services which could result to extreme competitive environment among the restaurants in winning and maintaining customers. Increasing competitive environment allows fast-food customers to choose among the available products and services with respect to personal preferences and brand names. Consequently, most of fast-food customers cannot be considered as potential loyal customers [1], [2].

A strong relationship between customer satisfaction and customer loyalty can be derived from managing customer experience [3], [4]. Some studies revealed findings that managing customer experience has potential association with food-related attributes. For example, some attributes such as taste, staff behavior, and food design have been identified as key factors in developing customer experience in restaurant [5]. Similarly, some attributes such as food presentation, staff competency, and food taste become the key clincher to customer satisfaction [6], [7].

In a row with above studies, firstly, the objective of this study is to identify determinant factor for customer satisfaction in fast-food restaurant. Although in other previous studies food quality, food presentation, and staff behavior have been recognized as determinant of customer satisfaction, few studies have related to other attributes such as menu pricing and discount, restaurant's cleanliness and neatness (sanitary condition), and waiting time [8]. Moreover, most of above studies used descriptive statistic and regression analysis method to analyze the sample, and fewer study focused on data mining, particularly the utilization of classification techniques for rule extraction. Rules are worthwhile pattern obtained by discovering the relationship among attributes in data. However, common methods such as descriptive statistic and regression analysis could not perform such task.

Meanwhile, machine learning as a data mining tool, has been broadly used to discover the correlation among attributes in data. Techniques such as decision tree (DT) [9] and neural network (NN) are prevalent due to their accuracy. Thus, secondly, the objective of this paper is to explore and to examine the utilization of information-based marketing strategy by extracting useful and interesting rules from fast-food customer data using DT and NN.

The rest of this study presents literature review related to customer satisfaction, data mining, and rule extraction techniques of neural network. Next, data set and experiment setup are discussed then we compare different rules obtained from DT and NN in terms of their accuracy along with potential

managerial implication of the findings for fast-food restaurant operators.

2. LITERATURE REVIEW

2.1 Customer Satisfaction in Fast-food Restaurant

Nowadays, marketing strategy has shifted from strategy that focuses on the brand (brand-centered) to customer-focused strategy (customer-centered). Many enterprises attempt to identify the factors that support customer satisfaction since it is a one of crucial business decision. As the heart of marketing, the ability to satisfy customers is great importance for many reasons [10]. Customer satisfaction is the proportion between what is customer's expectation and perceptions [11]. In business perspective, customer satisfaction is a measure of how products and services provided by an enterprise meet customer expectations [12].

In the context of fast-food industry, it is necessary to understand determinant factors driving customer satisfaction. For the restaurant managers, an understanding of underlying factors that influence customer satisfaction is a guidance to design services and offering [10]. Prior studies revealed the important attributes of customer satisfaction in restaurant. Since only few studies discussed customer satisfaction attributes in fast-food restaurant, thus we present and summarize attributes affect customer satisfaction in several types of restaurant as shown in Table 1.

Tangible (e.g. food quality and food presentation) and intangible (e.g. staff behavior and restaurant atmosphere) attributes were considered. For instance, Qu [13], Arora & Singer [14], Baek et al [15], and Gupta et al [16] considered attributes such as food quality (taste), food range, and food price are the most important attributes in restaurant. Several studies also indicate that nonfood-related attributes such as restaurant environment and staff friendliness were the determinant factors contributing to customer experience in restaurant [17], [18].

2.2 Data Mining and Rule Extraction Techniques

As stated by Hogan *et al* [19], the ability of an enterprise in acquiring and managing customer information is the key to maintain competitive advantage; but, managing customer relationships is hard to perform since customer, either as individual or in group, has different preferences and expectations. Hence, to overcome such problems, information-based marketing strategy is needed in

order to provide responsive decision making continuously.

Table 1: Attributes Related to Customer Satisfaction

Authors	Methodology	Result
Qu [13]	Regression analysis	Food and environment have the largest effect on intention of customer to come back.
Arora & Singer [14]	Regression analysis	Food quality and service have significant influence to customer satisfaction
Baek, *et al* [15]	Conjoint analysis	Menu price are the key attributes of choosing fast-food restaurant.
Gupta, *et al* [16]	Descriptive statistic	Food performance is considered as crucial attribute attributes of improvement for guests' come back
Park [17]	Correlation analysis	Food quality, staff kindness, and quick service are the most important attributes in fast-food restaurant
Jang, et al [18]	Factor analysis and Structural equation modeling	Service and food quality are key attributes influencing customer intention.

State of the art technological developments related to information management such as data warehouse and data mining enable enterprises to capture and analyze customer data up to the level of individual customers [20]. Confluence of data mining and machine learning in customer relationship management (CRM) domain have been widely employed as tools to discover the relationship among attributes in the data. In customer-related research, DT and NN have been considerably used in closed cycle CRM dimensions such as customer identification, customer attraction, customer retention, and customer development [12]. For instance, DT and rule extraction algorithm, called NeuroRule [21], had been applied to predict consumer preference of fast-food franchise. Classification rules that were generated from these techniques could achieve predictive accuracy of more than 80% [22]. An empirical experiment that was conducted by Hung, et al [23] shows these two techniques effectively enabled telecom service providers to make more accurate and effective churner prediction. Neural network also outperformed the Logit Regression in identifying potential churners [24].

Several methods of extracting rules from NN have been proposed since last decades. For instance, an earliest work called X2R has been introduced by Liu and Tan [25]. Through this feature selection were performed and concise rules from numeric dataset were generated even though generated rules are order-sensitive. Likewise, in 2002, Setiono *et al* [26] proposed a method that worked in a network with a single hidden layer and one linear output unit, called REFANN. Before applying REFANN, a pruning strategy, called NN pruning for function approximation (N2PFA), was employed in order to reduce the number of rules by removing redundant hidden units and irrelevant input attributes.

The latest work on extracting rules from NN called REANN was developed by Kamruzzaman and Islam [27]. REANN aims at searching of simple rules with acceptable predictive accuracy. Since REANN uses constructive pruning strategy, optimal network architecture can be specified automatically. Moreover, it can extract concise rules with acceptable accuracy.

3. RESEARCH PROCESS

In this section, we depict the research process of this study as in Figure 1. Details of each steps is described below.

3.1 Data Collection and Pre-Processing

According to Kotler and Armstrong [28], four tools are employed to measure customer satisfaction as well as providing complaint systems, satisfaction survey, 'ghost shopping,' and losing customer analysis. Hence, we collected data samples through a survey relating to customer satisfaction.

Empirical data were collected through a self-administered, survey of customers in big retail chain of fast-food restaurant in Palembang, Indonesia. The field work was conducted from January 2011 to February 2011. Some trained volunteers were involved in the survey and the author ensured and maintained the integrity of data quality. In total, 400 questionnaires with 10 input variables were distributed such as taste, serving, dessert, restaurant environment, washing basin, waiting time, staff behavior, menu pricing, discount, and store location. Final data contains 340 cases after we performed data pre-processing by eliminating duplicate instances or those with many inaccurate and missing values.
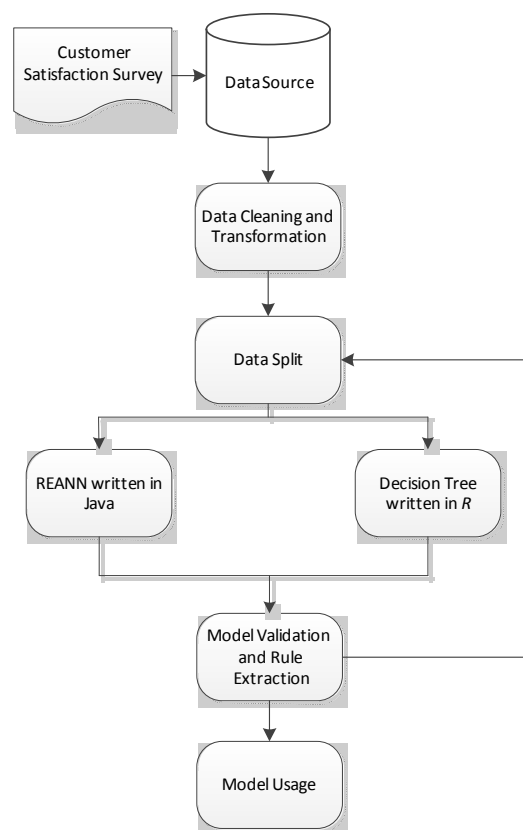

Figure 1: Diagram of Research Process

3.2 Data Mining and Model Evaluation

To extract the rules, we adopted well-known C4.5 program [9] was implemented as Rattle in R packages [29]. The interface (Figure 2) was developed based on REANN algorithm to provide researchers, restaurant manager, or decision makers a simple and easy tool in performing marketing research.

Rules are extracted from DT. For this purpose, of the 340 respondents who participated in survey, 133 observations (39.12%) and 207 (60.88%) observations were assigned to the Class 0 (unsatisfied) and Class 1 (satisfied), respectively. For NN training, Class 0 (unsatisfied) and Class 1 (satisfied) cases were assigned target values of (0, 1) and (1, 0), respectively.

To evaluate the model, datasets were partitioned into 2 parts. The first part was for training and rule extraction and the second part was applied for model validation. The first set was selected randomly and consists of two-third (66%) of the

dataset (226 cases) and the validation set consisted of the remaining 114 cases (33%).

Figure 1: A Java-based Interface for Rules Extraction with REANN Algorithm

4. RESULT AND DISCUSSION

4.1 Extracted Rules from C4.5 and REANN

Result from C4.5 algorithm that generates DT and classification rules is presented below. C4.5 is one of the most widely used tools for data mining because of its ease of use and high accuracy [30]. Results of the rules based on REANN algorithm are presented and a comparison between C4.5 and REANN are conducted in order to determine the accuracy.

C4.5 generated a total of 15 rules and one default rule. Seven rules classify samples as Class 1, the 7 rules and the default rule classify samples as Class 0. The accuracy of the rules extracted from C4.5 on the training and validation sets are 84.12 and 70.69%, respectively. Total of 2 extracted rules were obtained from REANN. The accuracy of the rules extracted from REANN on the training and validation sets are 84.78 and 73.40%, respectively. The result of those 2 algorithms can be summarized as shown in Table 2.

Table 2: Performance Evaluation

Algorithm	#Rules	Accuracy (%)	
		Training	Validation
Decision Tree (C4.5)	16	84.12	70.69
REANN	2	84.78	73.4

We merely choose interesting rules based on its meaning and usefulness for practical and implementation application. The most significant rules for Class 0 and Class 1 for C4.5 are presented as follows:

R1: IF staff behavior = good AND menu pricing = normal AND serving = less AND taste = good THEN Class 1

R2: ELSE IF staff behavior = good AND menu pricing = normal AND serving = normal AND restaurant environment = bad THEN Class 0

R3: ELSE IF staff behavior = good AND menu pricing = normal AND serving = normal AND restaurant environment = good THEN Class 1

R4: ELSE IF staff behavior = good AND menu pricing = normal AND serving = normal AND restaurant environment = normal THEN Class 1

R5: ELSE IF staff behavior = good AND menu pricing = expensive THEN Class 0

R6: ELSE IF staff behavior = good AND menu pricing = affordable AND discount = false AND dessert = good THEN Class 1

R7: ELSE IF staff behavior = bad Then Class 0

R8: ELSE IF staff behavior = bad AND discount = false THEN Class 0

While extracted rules generated from REANN can be shown as follows:

R1: IF staff behavior = good AND taste = good AND washing basin = clean AND waiting time = short THEN Class 1

R2: ELSE IF staff behavior = good AND washing basin = clean THEN Class 1

4.2 Discussion

From the experiment, it is shown that training set accuracy from the two methods are quite similar, the accuracy of the rules extracted from REANN on the validation set are higher by up to 2% than the accuracy of the rules from C4.5. This higher accuracy rates are also achieved by simpler rule sets consisting of fewer rule. Hence, in this case, by using REANN simple rules with high predictive accuracy were obtained as stated in [14].

However, of the 8 significant rules obtained from C4.5, Rule *R7* (dimension of service) could be considered to be significant because it can directly address the question "Why do customer tends to be unsatisfied in the fast-food restaurant?" It can be noted that staff behavior was the most determinant factor. In order to make customers satisfied, it is recommended that customer service be improved through self-motivation training.

In addition, rule *R2* can also answer the question "What factors in addition to price, service, and a serving that lead customers to be unsatisfied?" It can be concluded that inconvenient of location is significant attribute; therefore *R2* might be useful to restaurant operators in improving the convenient place. It might be of interest to seek further questions to the customer's opinion concerning the restaurant cleanliness and neatness. Through this, restaurant operators can then improve the environment of its outlets accordingly.

Moreover, there were no significant difference between the rules generated by C4.5 and REANN. Overall, a closer inspection of the rules generated by C4.5 and REANN could enable in identifying how the fast-food restaurant could strengthen their competitive advantage. Mostly, attributes dimension such as service and taste appeared in most rules, thus the findings suggested that the restaurant operator can enhance customer satisfaction by offering best service and tasty food. From business perspective, this experiment revealed essential recommendation to prioritize attributes by paying more attention to the best service attributes along with taste and price accordingly. However, the finding is contrary to [7] that stated food quality became the most determinant attribute in customer satisfaction.

4.3 Managerial and Practical Implications

This study provides some managerial implications which can be used by restaurant operators to better understand the factors related to customer satisfaction in fast-food restaurant. Fast-food restaurant have typical characteristic and become a trend in Indonesia. However, it is complicated to maintain fast-food customers become loyal customer. In strong competitive environment, creating a unique products or brands by potent positioning and differentiating is necessary to enhance competitive advantage.

This study revealed underlying factors of customer satisfaction in fast food restaurant such as staff behavior. It is recommended that restaurant operators understand required attitude when serving customers. Several crucial aspects regarding staff behavior such as staff's attitude, gesture, and emotional intelligence in facing and serving customers should be considered. Likewise, adequate staff training such as motivational training, basic customer service training aimed at enabling fast-food restaurant to offer excellent service to their customers should be considered. Lastly, continuous evaluation of the quality of customer service is needed as well to ensure efficient operation.

Apart from staff behavior, this study also uncovered convenient location as determinant factor. It is common that cleanliness is an important aspect of customer experience in the restaurant. There is correlation between customer expenditure, customer experience, and sanitary (cleanliness) condition.

Finally, food presentation and quality become essential factors in understanding customer satisfaction in fast-food restaurant. In this study it was also discovered that functional aspects of shaping customer experience i.e. taste [5] has an association with customer satisfaction. It is recommended that restaurant operators understand local taste of food in different areas. For instance, in the certain culture (area), customers prefer to have high level of spiciness, so that it is notable to make food products properly.

5. CONCLUSION

Data mining approaches based on decision tree and neural network to identify the most determinant predictor in influencing customer satisfaction in fast-food restaurant were proposed. Generated rules from C4.5 and REANN were quite similar, and REANN performed best accuracy in validation test compared to C4.5. The basic customer service attribute (i.e. staff behavior), restaurant's cleanliness, and food presentation was revealed to be significant predictor in determining customer satisfaction. This study provides some findings that are useful for the marketers, policy makers who have interest in customer satisfaction research. For practical implication, an understanding of underlying determinant factors of customer satisfaction could strengthen the competitive advantage of the fast-food restaurant.

In the future we intend to conduct a cross-sectional research by comparing the characteristics of fast-food customers in Indonesia and other countries. In addition, it also would be interesting by applying the model in other category of restaurants i.e. specialty and/or continental restaurant. Finally, it would be useful if more data samples could be collected in the future.

REFERENCES:

[1] C.Y. Wang, "Service Quality, Perceived Value, Corporate Image, and Customer Loyalty in the Context of Varying Levels of Switching Costs," *Psychology Marketing*, vol. 27, no. 3, pp. 252–262, 2010.

[2] M Haghighi, A. Dorosti, and A., Hoseinpour, A. Rahnama, "Evaluation of Factors Effacting Customer Loyalty in the Restaurant Industry," *African Journal of Business Management*, vol. 6, no. 14, pp. 5039-5046, 2012.

[3] R. Ladhari, I. Brun, and M. Morales, "Determinants of Dining Satisfaction and Post Dining Behavioral Intentions," *International Journal of Hospitality Management*, vol. 27, no. 4, pp. 563–573, 2008.

[4] H. Han and K. Ryu, "The Roles of the Physical Environment, Price Perception, and Customer Satisfaction in Determining Customer Loyalty in the Restaurant Industry," *Journal of Hospitality & Tourism Research*, vol. 33, no. 4, pp. 487–510, 2009.

[5] L.L. Berry, L.P. Carbone, and S.H. Haeckel, "Managing the Total Customer Experience," *MIT Sloan Management Review*, vol. 43, no. 3, pp. 85–89, 2002.

[6] Y., Jang, S. Namkung, "Does Food Quality Really Matter in Restaurants? Its Impact On Customer Satisfaction and Behavioral Intentions," *Journal of Hospitality & Tourism* , vol. 31, no. 3, pp. 387-409 , 2007.

[7] M. A. A Bashir, H. F. Ariffin, N Baba, and S. Mantihal, "The Impact of Food Quality and Its Attributes on Customer Satisfaction at Malay Restaurants," in *Hospitality and Tourism-Synergizing Creativity and Innovation in Research*, Sumarjan et al., Ed. London: Francais and Taylor, 2013, pp. 175-180.

[8] A Ponnama and M.S. Balaji, "Matching Visitation-Motives and Restaurant Attributes in Casual Dining Restaurants," *International Journal of Hospitality Management*, vol. 37, pp. 47-57, 2014.

[9] R. Quinlan, *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kauffman Publishers, 1993.

[10] Syed Saad Andaleeb and Carolyn Conway, "Customer satisfaction in the restaurant industry: an examination of the transaction-specific model," *Journal of Services Marketing*, vol. 20, no. 1, pp. 3-11, 2006.

[11] A. H. Kracklauer, D. Q. Mills, and D. Seifert, "Customer Management as the Origin of Collaborative Customer Relationship Management," *Collaborative Customer Relationship Management - taking CRM to the next level*, pp. 3-6, 2004.

[12] E.W.T. Ngai, Li Li Xiu, and D.C.K. Chau, "Applications of Data Mining Techniques in Customer Relationship Management: A Literature Review and Classifications," *Expert Systems with Applications*, vol. 36, pp. 2592-2602, 2009.

[13] H. Qu, "Determinant Factors and Choice Intention for Chinese Restaurant Dining," *Journal of Restaurant & Foodservice Marketing*, vol. 2, no. 2, pp. 35-49, 1997.

[14] R., Singer, J. Arora, "Customer Satisfaction and Value as Drivers of Business Success for Fine Dining Restaurants," *Services Marketing Quarterly*, vol. 28, no. 1, pp. 89–102, 2006.

[15] S., Ham, S., Yang, I. Baek, "A Cross-Cultural Comparison of Fast-food Restaurant Selection Criteria Between Korean and Filipino College Students," *International Journal of Hospitality Management*, vol. 25, no. 4, pp. 683–698, 2006.

[16] S., McLaughlin, E., Gomez, M. Gupta, "Guest Satisfaction and Restaurant Performance," *Cornell Hotel and Restaurant Administration Quarterly*, vol. 48, no. 3, pp. 284–298, 2007.

[17] C. Park, "Efficient or Enjoyable? Consumer Values of Eating-out and Fast-food Restaurant Consumption in Korea," *International Journal of Hospitality Management*, vol. 23, no. 1, pp. 87-94, 2004.

[18] S. S. Jang, J. Ha, and K. Park, "Effects of Ethnic Authenticity: Investigating Korean Restaurant Customers in the US," *International Journal of Hospitality Management*, vol. 31, no. 3, pp. 990–1003, 2012.

[19] Jhon E. Hogan et al., "Linking Customer Assets to Financial Performance," *Journal of Service Research*, vol. 5, no. 26, pp. 27-38, 2002.

[20] Markus Wubben, *Analytical CRM*. Wiesbaden: GWV Fachverlage GmbH, 2008.

[21] R. Setiono and H. Liu, "Symbolic Representation of Neural Network," *IEEE Comput*, vol. 29, no. 3, pp. 71-77, 1996.

[22] Y. Hayashi, M-H Hsieh, and R. Setiono,

"Predicting consumer preference for fast-food franchises:a data mining approach," *Journal of the Operational Research Society*, vol. 60, pp. 1221–1229, 2009.

[23] Shin-Yuan Hung, David C Yen, and Hsiu-Yu Wang, "Applying data mining to telecom churn management," *Expert Systems with Applications*, vol. 31, pp. 515–524, 2006.

[24] Michael C. Mozer and et al, "Predicting Subscriber Dissatisfaction and Improving Retention in the Wireless Telecommunications Industry," *IEEE Transactions on Neural Network*, vol. 11, no. 3, 2000.

[25] Huan Liu and Sun Teck Tan, "X2R: A Fast Rule Generator," in *IEEE International Conference on Systems, Man and Cybernetics*, vol. II, Vancouver, 1995, pp. 1631 - 1635.

[26] Rudy Setiono, Wee Kheng Leow, and Jacek M. Zurada, "Extraction of Rules From Artificial Neural Networks for Nonlinear Regression," *IEEE Transaction on Neural Networks*, vol. 13, no. 2, pp. 564-577, May 2002.

[27] S. M. Kamruzzaman and Monirul Md. Islam, "An Algorithm to Extract Rules from Artificial Neural Networks for Medical Diagnosis Problems," *International Journal of Information Technology*, vol. 12, no. 8, pp. 41-59, 2006.

[28] Philip Kotler and Gary Armstrong, *Principles of Marketing 14th Ed*. New Jersey: Prentice Hall, 2012.

[29] Graham Williams, *Data Mining with Rattle and R*. New York: Springer Science+Business Media, 2011.

[30] Tom M. Mitchell, *Machine Learning*. New York: McGraw-Hill Science, 1997.