World Academy of Science, Engineering and Technology
International Journal of Economics and Management Engineering
Vol:8, No:4, 2014

Data Mining to Capture User-Experience: A Case Study in Notebook Product Appearance Design

Rhoann Kerh, Chen-Fu Chien, Kuo-Yi Lin

***Abstract*—**In the era of rapidly increasing notebook market, consumer electronics manufacturers are facing a highly dynamic and competitive environment. In particular, the product appearance is the first part for user to distinguish the product from the product of other brands. Notebook product should differ in its appearance to engage users and contribute to the user experience (UX). The UX evaluates various product concepts to find the design for user needs; in addition, help the designer to further understand the product appearance preference of different market segment. However, few studies have been done for exploring the relationship between consumer background and the reaction of product appearance. This study aims to propose a data mining framework to capture the user's information and the important relation between product appearance factors. The proposed framework consists of problem definition and structuring, data preparation, rules generation, and results evaluation and interpretation. An empirical study has been done in Taiwan that recruited 168 subjects from different background to experience the appearance performance of 11 different portable computers. The results assist the designers to develop product strategies based on the characteristics of consumers and the product concept that related to the UX, which help to launch the products to the right customers and increase the market shares. The results have shown the practical feasibility of the proposed framework.

***Keywords*—**Consumers Decision Making,Product Design,Rough Set Theory,User Experience.

I. Introduction

As the rapidly development of the electronic technology and its extensive application, a variety of notebook products are launched promptly with different segment of users and the swift change of consumption cycle. In general, users get the first and the most important impression of products from the product appearance [1], which highly relevant to users' judgment and the satisfaction of the product. Notebook product should differ in its appearance to engage users and contribute to the user UX. It is crucial to refer the UX during the process of a new product design, for the purpose of launching the product that in line with the expectation of customers. UX is the feeling of the users regarding to use the product in several situations and screenplay [2]. Owing to the different factors of users, the reaction and the acknowledgment of the users to the productsare highly affected [3], [4].

Focus on the realistic need for capture UX, it's crucial to evaluate the factors that affect the customers' product appearance selection. By clarifying the corresponding factors, it is possible that we can find the target customers to launch the products. However, when collecting customer voice by questionnaire, it is difficult to extract the UX and generalize the relationship between the users' reaction and the experimental factors by the questionnaire results. Thus, it requires a systematic method to analyze the data and to conclude the important rules. Data mining is the useful tool to find out the relationship between users reaction and the factors by generating simple rules and detect the useful information hidden in the data. Under the pressure of time to market, the derived rules can assist the designers to communicate with the users, and help the manufacturers to lunch the right product efficiently [5].

There are several methods can be used for data mining, in particular, rough set theory (RST) [6] is a data mining technique to explore the relationship between users reaction and the factors by generating simple, useful rules. This approach can be effectively used especially when the information characteristic is imprecision, uncertainty, and vagueness [7], [8]. The assumption of RST such as the distribution and the independency among the variables is needless, In addition the come out "IF-THEN" rule is easy to explain and understand in practice. In recent years, RST approach has been well developed and applied in many areas, such as yield enhancement in TFT-LCD [9], personnel selection [10], fault diagnosis on distribution feeder [11], and medical decision making [12]. However, little research has been done for product design and consumer decision making. Therefore, it still opens various tracks for doing research works in this field.

This study proposed a data mining framework to capture UX based on rough set theory to extract useful information from consumer product appearance preferences and response. The proposed framework helps designers to evaluate various products in concept or prototype stage so as to identify the ideal product. The result of this framework can help to launch the new products that match customer needs and increase the market sales. The rest of this paper is organized as follows. Section II describes the fundamental of this approach and

C.F.Chien, corresponding author, is with the Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu 30013, Taiwan, ROC (phone: +886-3-5742648; e-mail: cfchien@mx.nthu.edu.tw).

R. Kerh is with the Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu 30013, Taiwan, ROC phone: +886-928-962958; e-mail: s102034607@m102.nthu.edu.tw).

K.Y. Lin is with the Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu 30013, Taiwan, ROC phone: +886-919-821186; e-mail: s9834810@m98.nthu.edu.tw).

reviews related studies. Section III addresses the proposed approach. Section IV evaluates this approach with the empirical data for validation. Section V concludes with discussion of contributions and future research directions.

II. Fundamentals of Applicable Methods

Data mining is the computational process of discovering potentially useful patterns from large amount of data sets [13]. This approach can help the decision maker to find out the valuable knowledge hidden in the raw data and the key factors influence the result of experiment. Owing to the progress of information technology, the technique has been widely used as a tool for decision making and analysis.

Extracted information and knowledge from data mining procedure is a useful basis for decision maker to observe the relevance of data information and the factors which affect product acknowledgement of users. Data mining has been widely applied in numerous area [14]-[17], which basically including four types of process. A briefing of this process is described as follows:

A. Problem Definition

The process of data mining starts from identifying the right problem to solve and constructing the objectives and the related attributes. Set up goals by understanding the background information and discussing with the experts, further, consider the expected results and benefits before mining process commence and at last to sort out the mining outcome systematically.

B. Data Preparation

The collected data often may contain noise value, missing, and inconsistent information. Data preparation is a procedure to enhance the quality of the information and make the mining process more efficient. In this procedure, it eliminates the noise value and identify the outliers also fill up the missing value and correct the inconsistent terms, and more, it sometimes converts the form of the data so as to facilitate the integration of the mining process.

C. Rules Generation

In this procedure we employed the method and the theory according to the mining problem type. For this study we constructed the model based on the rough set theory and used the software tool: Rough set exploration system (RSES) to extract the rules from the data.

D. Results Evaluation and Interpretation

By means of the mining process prototype and the data analysis, the useful information can be extracted as the standards of future notebook appearance design. In addition, we can interpret the meaningful rules and the sample types hidden in the data by discussing with the experts.

III. Framework of Proposed Approach

As shown in Fig. 1 is the data mining framework to capture UX. Basically, the proposed framework consists of four stages including problem definition and structuring, data preparation, rules generation, and results evaluation and interpretation. Further details are described as follows:

A. Problem Definition

Users responded their preference of products by completing the questionnaires under the design of the UX experiment. The scores of questionnaires reflected the level of satisfaction of users after experienced the product. However, the sample amounts were limited considering the factors of experimental expenditure and time. Besides, the factors influenced user reaction can be more to dozens, thus to generate the limited samples needs to premeditate numerous factors. Furthermore, the reasons of products reaction might came from the relationship of different combination of factors. Thus, it's important to clarify factors influenced user reaction rules which might be further estimate the type of market potential consumers, thereby assisting the corporate product strategy to increase market share.

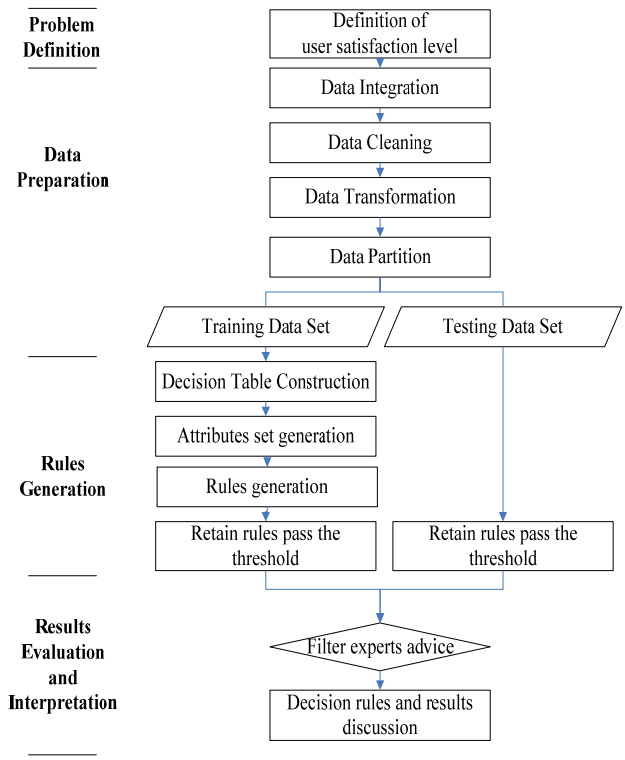

Fig. 1 Data mining framework to capture UX

B. Data Preparation

From the reaction of users after experienced the products, the data collected including two parts, the users' background and the product experienced information. The data of user background consists of user groups (such as gender, age, profession and major), past experience (such as related products using experience, related products possession experience and the frequency of using), current states (purchased motivation, purchased requirement) and the acknowledgement of products (perception, value and the attractiveness of products). The product experienced

information contains the data of the experimental process of all products (such as functional performance, experience time and the ease of using), the experimental reaction data (the aesthetic, functional satisfaction, the timing of using product and the product recommendations) etc. In order to complete the experiment, users need to experience all the different types of products.

The information collected must be sufficiently analyzed so as to solve the research problem. The process of data preparation can not only enhance the data quality but make the mining procedure and result efficiency. The data collected need to be integrated to consolidate from multiple sources into consistent data store. Furthermore, data cleaning approach was proposed for removing noise data and outliers, and correcting inconsistencies. Data transformation is proposed to transform the data to mining purpose. Data reduction is conducted to diminish the data dimension and enhance the efficiency of analysis. Finally, classified the data as training data set and testing data set so as to train and to verify the rules, in which generates rules by training data set, and verifies the effect of rules by testing data set.

C. Rules Generation

Rough set theory is conducted to generate the possible reducts and derive candidate rules with the training data. The procedures are shown as follows:

1. Construct the decision table: Sort out the data into the table of factors corresponded to the reaction.
2. Generate attributes set: Generate all possible reducts with removing unnecessary attributes and retained attributes which have obvious relationship with reaction variables.
3. Generate rules: Generate candidate rules based on the selected reducts from attribute set.
4. Filter the decision results: Consider various reactions of different users and delete the decision results which are not the focus class.
5. Calculate the support value of the rules: Calculate the support value of each rule, the formula is shown as bellows.

$$\text{Support}(\text{Rule}i)=P(\text{subset data selected by Rule}i) \quad (1)$$

6. Calculate the confidence level and the lift of the data, the formula is as bellows.

$$\text{Confidence}(\text{Rule } i)=\text{P}(data \mid \text{subset data selected by Rule}i) \quad (2)$$

$$\text{Lift}(\text{Rule } i)=P(data \mid \text{subset}i) \,/\, P(data \mid \text{population}) \quad (3)$$

7. Filter the rules: Retained the rules under the condition of the confidence level and the lift value higher than the setting threshold, and vice versa.

D. Results Evaluation and Interpretation

The rules obtained from the user reaction results corresponded to user background and experiment process, in which can interpret user decision and speculate the behavior of users. According to the discussion of experts for the interpretation of results, we are able to transform the UX results into useful customer information, which can not only provide key customer information to the enterprise but also assist the enterprise to predominate customer characteristics, so as to understand the objects and estimate the demand of the product sales and furthermore making marketing decision.

IV. Empirical Study

In path with the continued innovation of style and brand of the computer products, the product style, the exterior, the new functions, and the feature of product, have been in the fierce competition situation. Therefore, if we can understand the requirement of customers in the concept and prototype stage, we can develop a product which will not only be loved by customers but also contend important factors that customers considered.

The empirical study was conducted for Electronic Manufacturing Service (EMS) company in Taiwan and the experiment was executed during 20XX/1/1~20XX/3/31. The experiment recruited 168 subjects withdifferent background.

Experience the appearance performance, including weight, thickness, and size for the design of 11 different notebook prototypes. To comprehend the UX and sort out the consumer behavior, it is crucial that each user should experience the entire testing product. The data of user corresponding to product appearance reaction collected are 1848 in this study.

A. Problem Definition

This study probed the reaction of users corresponding to their background and the perception of product exterior. For the purpose of considering factors homogenization and avoiding the single level of factors which influence severely to the verified results, this study considered gender, major and the perspective of aesthetic of users as the basis of filtration. In particular, the experiment arranged users to actual experience the product, besides setting experimental standard operating procedure which provided users basic knowledge of experiencing the product. In order to perform user sense of the exterior and the feeling for the product, users should fill out the questionnaire of each product after completing the experiment. This study exploited the scale of Centrality of Visual Product Aesthetics (CVPA) [18], which to verify users feeling for the appearance of the product, and if the score of the scale were in high condition, it will show that users are more satisfied with the product aesthetic appearance.

B. Data Preparation

The data of this study contained the user background, the characteristic of product appearance, and the reaction information of users. The factors which were studied for user background and the characteristic of product appearance including several items, such as gender, major, personal aesthetic, arc aesthetic, ratio of thickness, proportion of product and the screen size, and the material feeling. The description and the level of factors are classified in Table I. In addition, by using the CVPA scale, we will know how important the users

view the aesthetic, which is the level of their personal aesthetic. The study divided users into high, medium, and low aesthetic group by the user response of questionnaire.

According to the score and the testing result of the user reaction data, we can divided the decision-making results into three categories, which were high level aesthetic reaction (425, 18%), medium level aesthetic reaction (1807, 77%) and low level aesthetic reaction (120, 5%). The results of three categories showed that there is a significant difference performance in analysis of variance (ANOVA) test, which the p-value is equal to zero.

For the purpose of generating rules and further validating the data, the study considered the order of the experiment time, thus the original data were split into training data, which has the earlier experiment time (1881, 80%) for deriving rules, and the testing data which has later testing time (471, 20%) for estimating the validity of derived rules. Five folds cross-validation were completed for the data set; moreover, the average value of support, confidence level and the lift will be tested to check whether it passes the threshold or not?

TABLE I
DIVIDED LEVEL OF FACTORS

Factor	Level
Gender	2
Major	6
Personal Aesthetic	3
Arc Aesthetic	5
Thickness	5
Product size	5
Screen	5
Material	5

C. Rules Generation

The analytic decision table inclusive of 9 factors and 1 decision attribute, and we can know the result of factors corresponding to decision-making attributes through the attribute sets. In this study, we investigated the rules and the reasons of high level aesthetic reaction and the low level aesthetic reaction, but excluded the medium aesthetic reaction. In addition, the rules were generated by eliminating the permutations situation and deleting 10 rules which is less than support threshold.

The description of the rules were explained as the following: by taking the first rule for instance, if the subject was a high aesthetic user and he or she satisfied the arc shape of the product, we can conclude that the overall appearance reaction of the subject to this product is high. It can be seen that there are 72 subjects from the data agree with this rule. As for the second rule, it indicated that if female users satisfied with the arc shape and the size ratio of the product, then the overall appearance reaction of the subject to this product is high. There are 49 subjects support this result. The following third and the forth rule can be explained similar to the previous two rules as the result shown in Table II.

After setting the threshold value, which we set the confidence level was 70%, and the lift was 1. It can be seen that the remained rules on Table III all pass the threshold and can be saved in the rule set for verification and for expert explanation. As for the confidence level threshold for testing data were set as 70% and the lift is 1. It can also be seen that the rules on Table IV all pass the threshold, which might indicate that the rules generated from the data were verified.

D. Results Evaluation and Interpretation

From the extracted rules and the interview with the subjects, we can conclude that there are four users' characteristics. The first one is that if the subject was a high aesthetic user and he or she satisfied the arc shape of the product, then the overall reaction of the product appearance is good. We found out that the impression of notebooks with high standard design and famous brand will deeply rooted in the high aesthetic customer's mind, while the leading fashion brand notebooks often have the arc appearance design, thus impacts the product selection of the high aesthetic users.

The second result we got is that if the female users satisfied with the arc shape and the size ratio of the product, then the overall reaction of product appearance is good. Through observation, high aesthetic female users will also be impacted by the notebooks with fashion brand and the arc appearance design, and thus will lead to choose the product with arc design. Whereas, the physiological structure of female users is different from the male user. For example, female's strength ability is smaller than male's, so the weight load, or the size of products will affect the overall comfort level. While the comfort level of the product is important to the female users; thence, for high aesthetic female users, they tend to choose the products with arc appearance design and the appropriate size ratio.

The third feature is that if the subject majored in engineering and he or she has a high satisfaction of the product arc appearance and the screen ratio, then the overall reaction of product appearance is good. By practical understanding of user habits, the users majored in engineering have to use the software program to accomplish their work for a long term; hence, they will in general emphasize the notebook screen feeling. These may conclude that if the screen ratio can meet the engineering majored users' requirement, then the overall reaction of the product will be good.

The last rules we obtained is that if the subject had low satisfaction of the arc appearance and the material feeling of the product, then the overall reaction of product appearance ispoor. Users take the arc design and the material tactile as the important factors they considered for the exterior of product. They thought that the unique material and the arc design will become the feature of the notebook brand, so as to impact their product decision.

V. CONCLUSION

This study developed a data mining framework to extract useful rules from the relationship between users' background information and their reaction to the product appearance design for a notebook computer. Based on the results of derived rules, we developed useful strategies for product designer to design the product that in line with customer's expectation. Furthermore, the result can assist to clarify the customer

segmentation and help the manufacture company to allocate the marketing resource to the right customer.

With an effective and systematic procedure to analyze questionnaire data and clarify the experimental factors and the response variable, the UX extracted from the empirical study provided useful rules hidden in the data. The studied result can assist product managers to understand the properties of the target customer, and to manage the product features and the brand images. In addition, the result may helpto introduce the product which match the requirement of the customers, and that may raise the market shares and increase the profits. For future study, it's feasible to develop a data mining structure for different types of product feature, such as new function and operation system, in order to increase the related application, and to gain a maximum profit as possible.

TABLE II
DECISION TABLE OF PRODUCT APPEARANCE REACTION

	IF								THEN	
Item	Gender	Major	Personal aesthetic feeling	Arc	Thickness	Size ratio	Screen ratio	Material	Product reaction	Support
1	x	x	3	5	x	x	x	x	Good	72
2	2	x	x	5	x	5	x	x	Good	49
3	x	2	x	5	x	x	5	x	Good	26
4	x	x	x	1	x	x	x	1	Poor	10

TABLE III
RESULT FROM TRAINING DATA RULES

Rule	Candidate Rule	Sample numbers[1]	Sample numbers[2]	Confidence level	Lift	Acceptable
1	If the subject was a high aesthetic user and he or she satisfied the arc shape of the product. Then, the overall reaction of the product appearance is good.	72	62	86.10%	4.38	Yes
2	If the female users satisfied with the arc shape and the size ratio of the product. Then, the overall reaction of product appearance is good.	49	43	87.80%	4.46	Yes
3	If the subject majored in engineering and he or she has a high satisfaction of the product arc appearance and the screen ratio. Then, the overall reaction of product appearance is good.	26	25	96%	4.88	Yes
4	If the subject had low satisfaction of the arc appearance and the material feeling of the product. Then, the overall reaction of product appearance is poor.	10	10	100%	5.08	Yes

[1]Sample numbers that match the rules condition [2]Sample numbers that match the rules condition and result

TABLE IV
RESULT FROM TESTING DATA RULES

Rule	Candidate Rule	Sample numbers[1]	Sample numbers[2]	Confidence level	Lift	Acceptable
1	If the subject was a high aesthetic user and he or she satisfied the arc shape of the product. Then, the overall reaction of the product appearance is good.	18	15	83.38%	4.23	Yes
2	If the female users satisfied with the arc shape and the size ratio of the product. Then, the overall reaction of product appearance is good.	15	13	86.70%	4.39	Yes
3	If the subject majored in engineering and he or she has a high satisfaction of the product arc appearance and the screen ratio. Then, the overall reaction of product appearance is good.	6	6	100%	5.08	Yes
4	If the subject had low satisfaction of the arc appearance and the material feeling of the product. Then, the overall reaction of product appearance is poor.	3	3	100%	5.08	Yes

[1]Sample numbers that match the rules condition [2]Sample numbers that match the rules condition and result

ACKNOWLEDGMENT

This research is supported by The Toward World Class University Project from Ministry of Education (102N2075E1), Hsinchu Science Park (102A26), National Science Council, Taiwan(NSC100-2628-E-007-017-MY3, NSC102- 2622-E-007- 013), and Quanta Computer, Inc.

REFERENCES

[1] Hoegg, J. A., and J. W. Alba, "Seeing is believing (too much): The influence of product form on perceptions of functional performance," Journal of Product Innovation Management, 28 (3), 346–359 (2011).

[2] Creusen, M. E. H. "Research Opportunities Related to Consumer Response to Product Design," Journal of product innovation management, 28(3), 405–408 (2011).

[3] Hassenzahl, M. and N. Tractinsky, "User Experience – a Research Agenda," Behaviour and Information Technology, 25(2), 91-97 (2006).
[4] Veryzer, R. W. and de B. B. Mozota, "The Impact of User-Oriented Design on New Product Development: An Examination of Fundamental Relationships," Journal of Product Innovation Management, 22(2), 128–143 (2005).
[5] Berry M. J.and G.S. Linoff, Data Mining Techniques: For Marketing, Sales, and Customer Support, John Wiley & Sons, Inc, (1997).
[6] Pawlak, Z., "Rough Sets," International Journal of Computer and Information Sciences, 11(5), 341-356 (1982).
[7] Greco, S., B. Matarazzo and R. Słowiński, "Rough sets theory for multicriteria decision analysis," European Journal of Operational Research, 129 (1), 1–47 (2001).
[8] Yang, X., J. Yang, C. Wu, and D. Yu, "Dominance-based rough set approach and knowledge reductions in incomplete ordered information system," Information Sciences, 178(4), 1219-1234 (2008).
[9] Hsu C.Y., C.F. Chien, K.Y. Lin and C.Y. Chien "Data mining for yield enhancement in TFT-LCD manufacturing: an empirical study," Journal of the Chinese Institute of Industrial Engineers, 27(2), 140-156 (2010).
[10] Chien, C. F. and L. Chen, "Using Rough Set Theory to Recruit and Retain High-Potential Talents for Semiconductor Manufacturing," IEEE Transactions on Semiconductor Manufacturing, 20(4), 528-541 (2007).
[11] Peng, J., C.-F. Chien and B. Tseng, "Rough set theory for data mining for fault diagnosis on distribution feeder," IEE Proceedings- Generation, Transmission, and Distributions, 151(6), 689-697 (2004).
[12] Pawlak, Z., "Rough set approach to knowledge-based decision support," European Journal of Operational Research, 99(1), 48–57 (1997).
[13] Hang, J. and M. Kamber, Data Mining Concepts and Techniques, Elsevier, (2006).
[14] Harding, J. A., M. Shahbaz, S. Srinivas and A. Kusiak, "Data Mining in Manufacturing:A Review," Journal of Manufacturing Science and Engineering, 128, 969-976 (2006).
[15] Chien, C. F., W. C. Wang and J. C. Cheng, "Data mining for yield enhancement in semiconductor manufacturing and an empirical study," Expert Systems with Applications, 33(1), 192-198 (2007).
[16] Ngai, E.W.T., L. Xiu and D. C. K. Chau, "Application of data mining techniques in customer relationship management: A literature review and classification," Expert Systems with Applications, 36(2-2), 2592-2602 (2009).
[17] Braha, D., Y. Elovici, M. Last, "Theory of actionable data mining with application to semiconductor manufacturing control," International Journal of Production Research, 45(13), 3059–3084 (2007).
[18] Bloch, P. H., F. F. Brunel and T. J. Arnold, "Individual differences in the centrality of visual product aesthetics: Concept and measurement," Journal of Consumer Research, 29(4), 551–565 (2003).