Sang Yun SEO, Sang Duck KIM, Seong Chan JO /Journal of Asian Finance, Economics and Business Vol 7 No 2 (2020) 221-228

221

Utilizing Case-based Reasoning for Consumer Choice Prediction based on the Similarity of Compared Alternative Sets[*]

Sang Yun SEO[**], Sang Duck KIM[***], Seong Chan JO[****]

Abstract

This study suggests an alternative to the conventional collaborative filtering method for predicting consumer choice, using case-based reasoning. The algorithm of case-based reasoning determines the similarity between the alternative sets that each subject chooses. Case-based reasoning uses the inverse of the normalized Euclidian distance as a similarity measurement. This normalized distance is calculated by the ratio of difference between each attribute level relative to the maximum range between the lowest and highest level. The alternative case-based reasoning based on similarity predicts a target subject's choice by applying the utility values of the subjects most similar to the target subject to calculate the utility of the profiles that the target subject chooses. This approach assumes that subjects who deliberate in a similar alternative set may have similar preferences for each attribute level in decision making. The result shows the similarity between comparable alternatives the consumers consider buying is a significant factor to predict the consumer choice. Also the interaction effect has a positive influence on the predictive accuracy. This implies the consumers who looked into the same alternatives can probably pick up the same product at the end. The suggested alternative requires fewer predictors than conjoint analysis for predicting customer choices.

Keywords: Similarity of choice alternative set, Case-based reasoning, Choice prediction, Product recommendation, Conjoint analysis

JEL Classification Code: C42, C44, D12, M31

1. Introduction

When consumers purchase a product, they usually compare product attributes such as price, size, brand, and so on. However, too much information can cause consumers to make less effective decisions due to information overload (Huang, 2000; Hwang & Lin, 1999; Jacoby, Speller, &

**First Author, Assistant Professor, Department of Business Administration, Kyungnam University, South Korea,
Email: syseo@kyungnam.ac.kr
***Second Author, Professor, Department of Business Administration, Kyungnam University, South Korea,
Email: sdkim@kyungnam.ac.kr
****Corresponding Author, Ph.D candidate, School of Management, Kyunghee University [Postal Address: 513-1, School of Management(Orbis Hall), Kyunghee University, 26 Kyunghee-daero, Dongdaemun-gu, Seoul,02447, South Korea].
Email: statistics@khu.ac.kr

Berning, 1974; Malhotra, 1982; Malhotra, Jain, & Lagakos, 1982). Today, a number of online stores feature a product recommendation system to assist customers with their shopping and decision making. The advantage of online store is to present various alternatives to consumers(Chung, 2017). This product recommendation system operates through collaborative filtering, which is a method of making automatic predictions about the interests of a user by collecting the preference information of many other users (Lekakos & Giaglis, 2006; Resnick, Iacovou, Suchak, Bergstrom, & Riedl, 1994). Therefore, the collaborative filtering system operates well when data are sufficient.

In this study, we propose a different method of predicting consumer product choice by utilizing the similarity in the characteristics of the choice alternatives. Our proposed method does not require a consumer's demographics, purchase history, or post-purchase satisfaction, which are data typically used in collaborative filtering recommendations.

Literature has long established that a consumer's decision-making process involves searching for information and evaluating alternatives before purchasing the product

(Engel, Blackwell, & Miniard, 1995). Investigating these alternatives can provide several pieces of information concerning the attributes consumers are considering. For example, if the consumer prefers a specific brand, his or her choice alternatives are more likely to include that brand. The stage of searching for information in the consumer decision-making process can yield meaningful clues to predict consumer preference.

A principle of market segmentation is splitting consumers or potential customers into several groups, with the customers in each group likely to have similar levels of interest, and to seek similar product characteristics or benefits. Usually, it is assumed that consumers within a unique segment have a similar preference structure (Desarbo, Ramaswamy, & Cohen, 1995; Kamakura & Russell, 1989). Indeed, respondents included in the same segment show identical responses to the marketing mix (Kamakura & Russell, 1989; Wedel & Kamakura, 2000). Therefore, it is believed that consumers within the same segment use identical decision heuristics, because consumers are inherently limited information processors and seek to conserve cognitive energy (Andrews & Currim, 2002).

Based on this assumption, this study suggests an alternative method for predicting consumer choice based on case-based reasoning, which makes use of the utilities of consumers with identical consideration sets at the product evaluation stage. This prediction method can be used to improve the effectiveness of a product recommendation system.

The rest of this paper is organized as follows: the next section reviews existing literature on case-based reasoning as well as the method to measure similarity between subjects. This is followed by a description of the research method; the simulation results; and the conclusion, including implications and future research.

2. Literature Review

2.1. Prediction of Consumer's Choice by Case-based Reasoning

Case-based reasoning is the process of solving new problems based on the solutions of similar past problems. The premise of case-based reasoning is that similar problems have similar solutions. Case-based reasoning comprises a four-step process. The first step is to retrieve from memory cases that are relevant to solving the target problem. The second step is to adapt the solution as needed to fit the target problem. The third step is to revise the previous solution to the target problem. The final step is to retain the resulting experience as a new case in memory

after the solution has been successfully adapted to the target problem (Aamodt & Plaza, 1994).

A number of online stores have improved their product recommendation systems by applying case-based reasoning for customer satisfaction, customer relationship management, and service customization. The basic role of a product recommendation system is to recommend products that match user requirements based on the retrieval process. The retrieval process is driven by the similarity of current user requirements or the target problem to past requirements or problems (Finnie & Sun, 2002). Therefore, the effectiveness of a product recommendation system depends on its ability to match and retrieve user requirements for the product, requiring a large variety of past cases to be stored in memory cases. Amazon.com and MovieLens are well-known sites using these recommendation systems.

A significant amount of research has been conducted on methodologies to upgrade the case-based reasoning for product recommendation. Chiu (2002) proposed a model to predict the customer's response using the genetic algorithm for insurance companies. Bousbahi and Chorfi (2015) proposed the MOOCs system, which matches learners' requests based on profile, needs, and knowledge, using case-based reasoning and special retrieval information techniques.

This paper also proposes an alternative algorithm that exploits the similarity between customers using case retrieval, and tests how this method works in terms of predicting consumers' choice. The similarity between subjects is calculated in terms of the unity between alternative sets that are compared in the product evaluation stage.

2.2. Similarity Evaluation

In this study, the similarity between subject i and j, $S(i, j)$, is measured in terms of the unity of compared alternatives (Xiong & Funk, 2006). In equation (1), k_i is the value of attribute k that the subject i prefers, and $f(k_i, k_j)$ is a similarity function of normalized distance between each attribute level. The weight W_k denotes the importance of the kth attribute. Thus, the similarity is the weighted sum of normalized distance between the attribute levels each subject prefers.

$$S(i,j) = \sum_{i,j}^{N} \sum_{k=1}^{K} f(k_i, k_j) \times W_k \qquad (1)$$

where i, j is the index of each subject; N is the total number of respondents; $f(k_i, k_j)$ is the similarity value between subject i and j for attribute k; and W_k is the

weighting parameter, which is limited to a value between zero and one.

The function applied to measure similarity between the preferred attribute levels depends on whether the attribute is categorical or continuous. For example, when the compared attribute is a continuous attribute such as price, equation (2) is applied as a similarity function. Equation (2) is close to one, as the preferred attribute level is the same.

Otherwise, when the attribute is categorical attribute, such as brand, then equation (3) is applied to measure similarity. In equation (3), the function is either zero or one, which corresponds to the case when the preferred attribute level is identical or not, respectively.

$$f(k_i, k_j) = 1 - \frac{|a_{k,i} - b_{k,j}|}{|max_{k,i} - min_{k,j}|} \qquad (2)$$

$$f(k_i, k_j) = \begin{cases} 1 \ (a_{ki} = b_{kj}) \\ 0 \ (a_{ki} \neq b_{kj}) \end{cases} \qquad (3)$$

where, a_{ki} is the attribute level at which the subject i prefers the attribute k; max_{ki} is the largest level of attribute k and min_{ki} is the lowest level.

When the consumer compares more than two alternatives, the similarity is averaged over every similarity of alternatives. Therefore, the similarity always ranges from zero to one.

3. Research Methods

3.1. Procedure

The empirical research is conducted using a self-reporting questionnaire. The product used in this empirical study is a laptop computer that is familiar to students participating in this research.

In the first step of the questionnaire, the subjects are exposed to sixteen profiles of a laptop computer and they are asked to choose a favorite profile that they are likely to purchase. Next, they are asked to choose three additional profiles, in their order of preference, that they would also consider buying. Next, the similarity of the alternatives is computed in terms of the number of alternatives chosen. In the second step, the participants are asked to evaluate each of the sixteen profiles, and the results are used to estimate the utilities of each attribute level by using the utility decomposition method of conjoint analysis. In the final step, the participants are asked to select a favorite profile among the four profiles for a validity test. This choice test for validity is conducted four times with a different validation set profile each time.

Therefore, the experiment factors hypothesized to affect predictive accuracy are the extent of similarity (S), and the number of alternatives used for similarity calculation (N). The predictive accuracy of case-based reasoning for the validity test is calculated by how many times the choice matches the choice of participants in the four times trials.

3.2. Stimuli for Experiment

The profile design used in the questionnaire focused on attributes of laptop computer. These attributes were selected by a pretest, in which the subjects were asked to select four important attributes among Brand, CPU performance, RAM capacity, HDD capacity, battery life, weight, and display size when buying a laptop computer.

The details of each attribute are provided in Table 1. The most preferred attributes are CPU, RAM, HDD, and weight, which 12.7% respondents of three hundred forty participants preferred.

Table 1 : Attributes and Level

Attribute	Attribute Levels			
Brnad	SEC	LG	HP	
CPU	Intel i7	Intel i5	Intel i3	
Memory	2G	4G	8G	16G
HDD	128G	256G	500G	1T
Battery Life	4 Hours	7 Hours	9 Hours	12 Hours
Weight	1~1.3kg	1.3~1.7kg	1.7~2.5kg	More than 2.5kg
Display Size	12 inch	13.3 inch	14 inch	15.6 inch

Thus, the product profiles are designed using these four attributes, and sixteen profiles are generated by orthogonal design. Table 2 shows each product profile and its attribute level. The participants were asked to choose considerable alternatives among sixteen profiles and to evaluate each product profile for purchase decision making.

Table 3 shows the profile set to be used for the validity test of similarity case-based reasoning to predict the subject's choice. The participants were asked to select a favorite profile in the set and repeat this choice test four times.

Table 2 : Laptop computer profiles and attribute level

No.	Brand	CPU	RAM	Weight	No.	Brand	CPU	RAM	Weight
1	HP	i7	2G	2.0kg	9	HP	i5	8G	1.3kg
2	SEC	i3	2G	1.3kg	10	HP	i3	4G	2.5kg
3	SEC	i3	8G	2.0kg	11	HP	i3	16G	1.7kg
4	SEC	i3	2G	2.5kg	12	LG	i5	2G	1.7kg
5	LG	i3	16G	2.0kg	13	LG	i7	8G	2.5kg
6	LG	i3	4G	1.3kg	14	SEC	i5	4G	2.0kg
7	SEC	i3	8G	1.7kg	15	SEC	i5	16G	2.5kg
8	SEC	i7	4G	1.7kg	16	SEC	i7	16G	1.3kg

The predictive accuracy is computed by how many times the subject's choice and the predicted choice using similarity case-based reasoning match. Therefore, the predictive accuracy for the validity test ranges from zero to one, at intervals of 0.25 for each subject.

Table 3 : Profile set for the validity test

Set	Profile	Brand	CPU	RAM	Weight	Baseline
Set 1	P1	SEC	i7	16G	2.5kg	26%
	P2	LG	i5	16G	1.3kg	69%
	P3	HP	i3	2G	1.3kg	5%
	P4	LG	i3	4G	2.5kg	1%
Set 2	P1	LG	i5	2G	2.5kg	2%
	P2	SEC	i5	16G	1.7kg	60%
	P3	HP	i7	16G	1.7kg	37%
	P4	SEC	i3	2G	2.0kg	1%
Set 3	P1	LG	i7	2G	2.5kg	2%
	P2	HP	i5	8G	2.5kg	8%
	P3	LG	i3	2G	2.0kg	2%
	P4	LG	i7	16G	1.7kg	88%
Set 4	P1	LG	i5	8G	2.0kg	55%
	P2	SEC	i3	4G	2.0kg	4%
	P3	SEC	i3	4G	1.7kg	30%
	P4	LG	i7	2G	2.0kg	11%

4. Empirical Study

4.1. Data Collection

For the empirical study, we employed a self-reporting questionnaire and three hundred twenty-two undergraduate and graduate school students participated in the survey. However, sixteen students were removed due to incomplete and unreliable responses, thus three hundred six subjects remained. The average participant age is approximately twenty-two years old and 56.2% are males.

4.2. Prediction Algorithm

We apply equation (1), (2), and (3) in order to calculate the similarity between each subject. The case-based reasoning predicts the target subject's choice from the following steps. First, the case-based reasoning searches the most similar subject case with target subject, based on the similarity between the alternative set. Second, the case-based reasoning calculates the utilities of each attribute level of the target subject by applying utility values derived from conjoint analysis of the most similar case subject. Third, the case-based reasoning calculates the sum of utilities for each profile to which the target subjects are exposed to in the validity test. Forth, the case-based reasoning recommends the highest utility profile for the target subject. If the target subject selects the recommended profile, then it is assumed that the case-based reasoning accurately predicts the target subject's favorite profile. This validity test of case-based reasoning occurs four times for each subject and the hit ratio is computed by the number of matches between both choices in the four trials.

The target subjects are randomly selected from all participants. We predicted the choice of one hundred target subjects using case-based reasoning. Therefore, the total amount of simulations is 100 (subjects) × 4 (number of chosen alternatives: 1, 2, 3, 4). In addition, we examine how the hit ratio changes according to similarity between cases. This repeated simulation is conducted by R software.

4.3. Results of Simulation

The simulation results are provided in Figure 1. The hit ratio of case-based reasoning is marked as the opened circle on the axis of similarity. There are four graphs according to the number of considered alternatives. The N=1 indicates that only one profile is considered for making a choice, and N=2 indicates that two alternatives are considered.
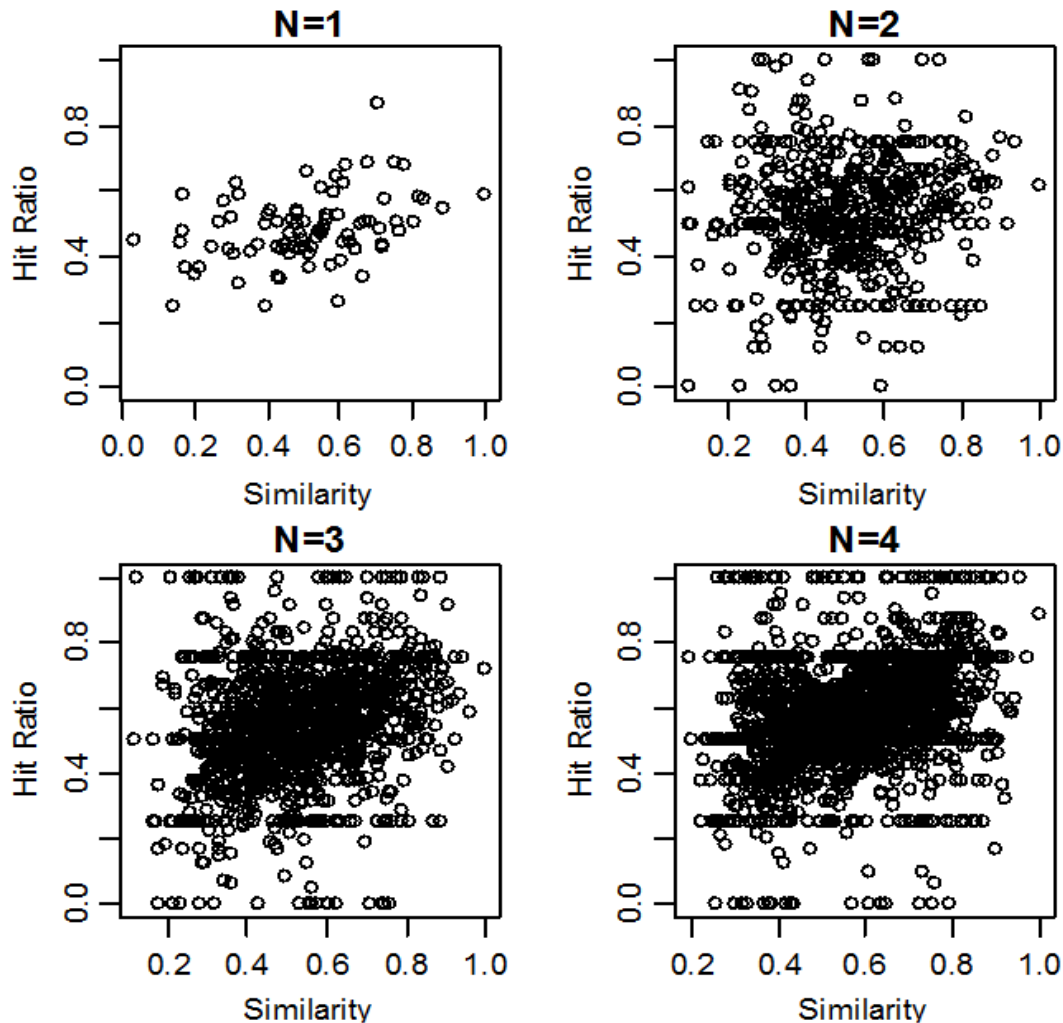

Figure 1 : Variations in predictive accuracy according to number of products and matches

The reason why there are less data points in the figure of N=1 and N=2 is that less attributes are compared for computing similarity between alternatives. In other words, when only one alternative is compared for similarity calculation, each of the four attributes is compared. However, when four alternatives are compared, then four alternatives × four attributes are compared for similarity calculation. Thus, the more alternatives compared, the more cases with different similarity may occur.

In the simulation result of N=1 and N=2, the hit ratio appears to ascend slightly as alternatives for making a choice become more similar. On the contrary, the ascending pattern of hit ratio appears more prominently in the simulation result of N=3 and N=4.

In order to compare the effect of similarity, we split the observed data of similarity into two groups, either high or low similarity based on the 0.5 similarity level. As shown Figure 2, the high similarity group is marked as a solid line, and the low similarity group is marked as a dotted line. Generally, the hit ratio of the high similar group appears higher than the low similar group, regardless of the number of alternatives. In addition, the hit ratio increases continuously when more alternatives are used for comparison for making a choice.

In order to test the effect of similarity and the number of alternatives on hit ratio, we use a 2 (similarity: high vs low) × 4 (N of alternatives: one, two, three, four) analysis of variance (ANOVA). The result of ANOVA is provided in Table 4. The result of ANOVA reveals a significant primary similarity effect and the number of alternatives.

The hit ratio average of a high similar group is significantly higher than the low similar group (M_{High} = 0.58 vs M_{Low} = 0.52, F= 105.585, p < .00). In addition, the average of hit ratio is significantly different from the number of alternatives (F= 31.417, p < .000). The average of hit ratio is ascending as more alternatives are used.
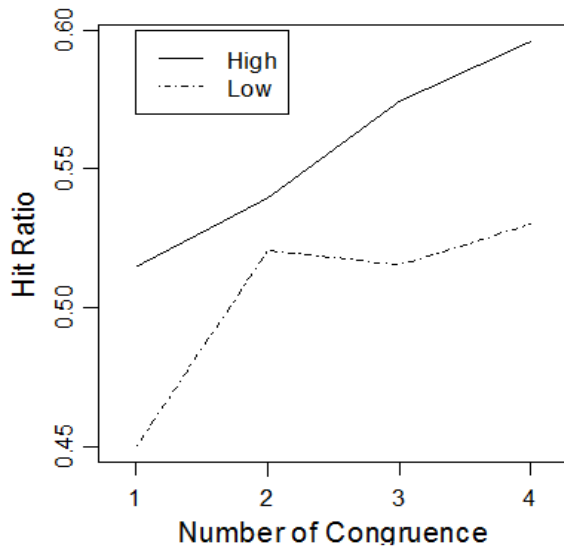

Figure 2 : Comparison of hit ratio by similarity and the number of alternatives
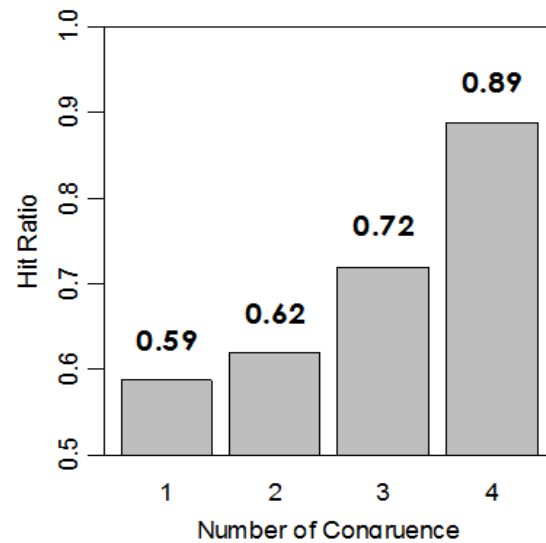

Figure 3 : Hit ratio by the number of perfectly matched alternatives

Figure 4 : Hit ratio by similarity and number of alternatives

Number of Alternatives Used in Reasoning	Similarity		Mean
	Low	High	
1	0.450	0.514	0.484
2	0.520	0.539	0.529
3	0.505	0.574	0.547
4	0.530	0.596	0.570
Total	M=0.520	M=0.580	0.556

Figure 5: ANOVA of hit ratio by the similarity and the number of alternatives

Factors	Sum Sq	Df	Mean Sq	F	Pr(>F)
High or Low in Similarity	3.06	1	3.062	105.585	2.00E-16
Number of Congruence	0.91	1	0.911	31.417	2.24E-08
High or Low in Similarity × Number of Congruence	0.13	1	0.128	4.441	0.0352
Residual	104.11	3590	0.029		

In addition, in order to investigate the effect of the number of alternatives, we draw a graph for the hit ratio of case-based reasoning taking advantage of only perfectly identical cases of alternatives. The case-based reasoning predicts a subject's choice based on the utilities derived from a subject whose considered alternatives are identical to each other. As shown Figure 3, the hit ratio is ascending, as more matched alternatives are used for case-based reasoning. The hit ratio is 59% with only one identical alternative, 62% with two alternatives, 72% with three alternatives, and 89% with four alternatives.

4.4. Discussion

In order to verify the effectiveness of case-based reasoning, we compare the hit ratio of case-based reasoning with the ratio of the most preferred choice in the validity test. In other words, in this study, an alternative method for recommending a product to subjects is to recommend the

product most subjects prefer. The number on the far right in Table 3 indicates the ratio of the subjects' favorite profiles. Profile 3 is the most preferred with 69% of set 1, profile 2 with 60% in set 2, profile 4 with 88% in set 3, profile 1 with 55% in set 4, and the average ratio of the most preferred profile is 67.7%.

We believe that case-based reasoning is valuable when the hit ratio of case-based reasoning outperforms the average ratio of the favorite profile. According to Figure 3, when more than three perfectly identical alternatives are used in case-based reasoning, the hit ratio exceeds the ratio of 72%. Therefore, an effective prediction by case-based reasoning using alternative similarity depends on how many identical alternatives are used.

5. Conclusion

5.1. Summary

In order to determine which product is suitable for marketing using the online store, it is necessary to consider the characteristics of the product line in the decision of consumer purchasing of the product line (Han & Kim, 2015). In the context of online stores, several product recommendation systems have traditionally relied on collaborative filtering, which is a method used to recommend items based on a user's past behavior. The main idea behind collaborative filtering is that similar users share the same interests and like similar items. Thus, the effectiveness of the collaborative filtering system depends on having access to enough information to solve problems. This study proposes an alternative method for recommendation systems that does not use past decision behavior; rather, it uses the characteristics of the alternatives considered by the consumer when making a decision.

After being briefed on the empirical study, the subjects were initially asked to select four alternatives among sixteen profiles, in the order of their preference, that they were willing to purchase. Then, they were asked to choose their favorite profile among the four profiles, in order to test the predictive accuracy of case-based reasoning based on the similarity of the identified alternatives. The case-based reasoning retrieves information on the subjects most similar to the target subjects and exploits the utility value derived by those subjects for calculating predictive validation profiles. The case-based reasoning system ranks the profiles according to the sum of the utilities for each attribute level and recommends or predicts the profile with the maximum utility among the alternative profiles as the choice of the target subject.

The simulation results revealed that the similarity of identified alternatives between the target subjects and retrieved subjects significantly contributes to the predictive accuracy of the case-based reasoning recommendation system. In addition, the more alternatives that are used in case-based reasoning, the higher is the hit ratio. The interaction between the number of alternatives and similarity is also significant.

This result implies that when the identified alternatives that a consumer is considering when making a choice are similar, the chosen alternative is likely to be the same. Thus, when online-shopping consumers postpone their decision making because they are searching for additional information, or have changed their purchase plan, and so on, case-based reasoning based on the similarity of identified alternatives can help recommend the right products for these consumers

In terms of theoretical contribution, the algorithm of case-based reasoning based on the similarity of identified alternatives can be applied by recovering sequential missing data. For example, traditional conjoint analysis requires ranking data for full profiles; however, there are frequent cases of incomplete data, such as repeated rankings or missing data, owing to participant oversight. The simulation result of this study shows that the missing values can be recovered by exploiting the data from similar cases in which the information is complete.

5.2. Limitations and Further Research

This work has several limitations. First, this study hypothesizes that the consumer decision process involves a comparison of alternatives. However, a consumer's decision process does not always go through this step. For example, consumers not highly involved in the product are unlikely to make the effort to compare products. Additionally, they often engage in routine purchase behavior based on information they have gathered during past purchase experiences. Consumers often make decisions based only on brand experience, store location, store loyalty, and so on (Cheng & Kim, 2019).

Particularly, it is well known that a consumer's decision process differs based on their product involvement. Therefore, it is worth investigating whether this case-based reasoning prediction is applicable to low-involvement products or situations.

Second, attribute importance was not considered in the calculation of similarity between alternatives. In other words, the importance of product attributes may differ for each customer; however, all attributes were assigned equal importance in this study. Unfortunately, we did not ask participants to rank their perceived importance of each attribute. However, several studies have attempted to apply

case-based reasoning after taking into account attribute importance by including weight parameters derived from the genetic algorithm (Chiu, Chang, & Chiu, 2003; Shin & Han, 1999). Additionally, Park and Han (2002) predicted a firm's bankruptcy with case-based reasoning, including weighting parameters derived from the analytic hierarchical process.

Finally, we did not consider the dynamics of the solution. In other words, the present solution may differ from a past solution to the same problem. We conjecture that the present solution is more likely to be relevant to the target problem, as the present solution is derived from updated circumstances regarding the problem. We need to consider the weight of the solution in each case, in terms of time series. It will be interesting to examine whether including the time series weighted value of each case into the model significantly upgrades its validity.

References

Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications, 7*(1), 39-59.

Andrews, R. L., & Currim, I. S. (2002). Identifying segments with identical choice behaviors across product categories: An intercategory logit mixture model. *International Journal of Research in Marketing, 19*(1), 65-79.

Bousbahi, F., & Chorfi, H. (2015). MOOC-rec: A case based recommender system for MOOCs. *Procedia-Social and Behavioral Sciences, 195*, 1813-1822.

Cheng, Z. F., & Kim, G. B. (2019). The Relationships among Brand Experience, Customer Perceived Value, and Brand Support Behavior in Service Industry. *Journal of Distribution Science, 17*(2), 91-100.

Chiu, C. (2002). A case-based customer classification approach for direct marketing. *Expert Systems with Applications, 22*(2), 163-168.

Chiu, C., Chang, P., & Chiu, N. (2003). A case-based expert support system for due-date assignment in a wafer fabrication factory. *Journal of Intelligent Manufacturing, 14*(3-4), 287-296.

Chung, J. B. (2017). Internet shopping optimization problem with delivery constraints. *Journal of Distribution Science, 15*(2), 15-20.

DeSarbo, W. S., Ramaswamy, V., & Cohen, S. H. (1995). Market segmentation with choice-based conjoint analysis. *Marketing Letters, 6*(2), 137-147.

Engel, J. F., Blackwell, R. D., & Miniard, P. W. (1995).

Consumer behavior. Orlando, FL: Dryden Press.

Finnie, G. & Sun, Z. (2002). Similarity and metrics in case-based reasoning. *International Journal of Intelligent Systems, 17*(3), 273-287.

Han, J. W., & Kim, W. Ki. (2015). The effect of product type and channel prioritization on effective digital marketing performance. *Journal of Distribution Science, 13*(5), 91-102.

Huang, M. (2000). Information load: Its relationship to online exploratory and shopping behavior. *International Journal of Information Management, 20*(5), 337-347.

Hwang, M. I., & Lin, J. W. (1999). Information dimension, information overload and decision quality. *Journal of Information Science, 25*(3), 213-218.

Jacoby, J., Speller, D. E., & Berning, C. K. (1974). Brand choice behavior as a function of information load: Replication and extension. *Journal of Consumer Research, 1*(1), 33-42.

Kamakura, W. A., & Russell, G. J. (1989). A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research, 26*(4), 379-390.

Lekakos, G., & Giaglis, G. M. (2006). Improving the prediction accuracy of recommendation algorithms: Approaches anchored on human factors. *Interacting with Computers, 18*(3), 410-431.

Malhotra, N. K. (1982). Multi-stage information processing behavior: An experimental investigation. *Journal of the Academy of Marketing Science, 10*(1), 54-71.

Malhotra, N. K., Jain, A. K., & Lagakos, S. W. (1982). The information overload controversy: An alternative viewpoint. *The Journal of Marketing, 46*(2), 27-37.

Park, C., & Han, I. (2002). A case-based reasoning with the feature weights derived by analytic hierarchy process for bankruptcy prediction. *Expert Systems with Applications, 23*(3), 255-264.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). *GroupLens: An open architecture for collaborative filtering of netnews.* Paper presented at the 1994 ACM Conference on Computer Supported Cooperative Work (pp.175-186). New York, NY: ACM.

Shin, K., & Han, I. (1999). Case-based reasoning supported by genetic algorithms for corporate bond rating. *Expert Systems with Applications, 16*(2), 85-95.

Wedel, M., & Kamakura, W. A. (2000). *Market segmentation: Conceptual and methodological foundations*. Boston, MA: Kulwer Academic Publisher.

Xiong, N., & Funk, P. (2006). Building similarity metrics reflecting utility in case-based reasoning. *Journal of Intelligent & Fuzzy Systems, 17*(4), 407-416.