# Prediction of Purchase Behaviors Based on Customer Demand Value Using Factorization Machines

Yuya Miyamoto and Michiko Tsubaki
The University of Electro-Communications, Department of Informatics, Chofu, Japan
Email: m1630121@edu.cc.uec.ac.jp, tsubaki@se.uec.ac.jp

***Abstract*—This study focuses on Integrated Marketing Communication which is an important factor of Customer Relationship Management. The purpose of this study is to compare the customer sales data of two different types of stores in the confectionery industry, using the framework which is "Sources of Communication Messages in a Relationship". In addition, this study proposes the method to categorize the customer's pursuit values for each store based on sales data. And also this paper proposes the new approach that can promote sales while increasing the value of usage in customers' living, based on devising a method to improve the efficiency of customer direct mail delivery, using Factorization Machines. The result shows the value that customers desire in each two shops is different from not only on-site investigation but also customer sales data. Furthermore, a customer demand value is defined in this study, the customer direct mail delivery is established to combine customer demand value with the result of purchase behavior prediction using Factorization Machines, makes that the value of usage in customers' living increase and sales promote.**

***Index Terms*—factorization machines, customer demand value, purchase behavior, direct marketing, customer relationship management, prediction**

## I. Introduction

Customer Relationship Management (i.e., CRM) is known to be a important factor in improving the relationship between customers and the company, thus, it is important for companies to provide the demanded value to customers (Grönroos(2007)[1]). Don (1992)[2] has proposed Integrated Marketing Communication (i.e., IMC) which indicates the need to integrate all forms of communications including media marketing, direct marketing and advertising, to the provision and consumption of products and services and the communications between customers as an essential factor to CRM. In addition, Gronroos (2000)[3] proposed "Sources of Communication Messages in a Relationship" (Figure1.) which extended four sources of communication messages proposed by Duncan and Moriarty (1997)[4]. This includes "Planned Messages", "Product Messages", "Service Messages", "Unplanned Messages" and "Absence of Communication". "Planned Message" is the result of the planned marketing communication campaign and it is often used in Direct Mail (i.e., DM), televisions, and publications. It is generally an one-sided message from companies, but as for DM, companies could send messages according to individual customer's needs and it is relatively easy to for companies improve it. Therefore, this paper focuses on the DM included in the "Planned Messages". Although electronic devices are widely used today, DM continues to play an important role in many corporate communication strategies (Simon and Lynda, 2015)[5]. At Canada Post Corporation (2015)[6], DM is said to be more persuasive than digital media, and it is much easier for customers to understand the advertising contents. Also in previous studies on DM distribution, it has been verified that the different types of DM (e.g., the purposes of promotion, the purposes of building customer relationships) lead to different results of purchasing behaviours (Gazquez-Abad, Canniere and Martinez-Lopez, 2011)[7]. Migueis, Camanho and Borges (2017)[8] compared direct marketing reactions at banks using various methods of machine learning. Chen et al. (2016)[9] also described the validity of conducting purchasing behavior predictions using Factorization Machines in e-commerce. However, in CRM in terms of the delivery of DM, none of the previous researches took the value desired by customers into considerations, which is considered important.

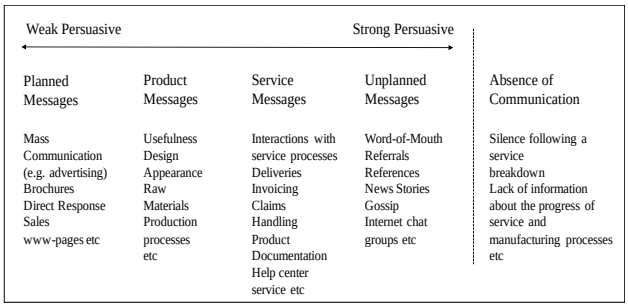



Figure 1. Sources of communications messages in a relationship.

Therefore, the purpose of this research is to compare the customer sales data of two different types of stores in the confectionery industry, using the framework proposed by Gronroos (2000)[3], and to categorize the customer's pursuit values for each store. Therefore, this paper

Manuscript received July 6, 2018; revised October 31, 2018.


doi: 10.18178/joams.6.4.190-196

proposes the new approach that can promote sales while increasing the value of usage in customers' living, based on devising a method to improve the efficiency of customer direct mail delivery, using Factorization Machines.

## II. Analysis Data

The data analyzed in this study is data of two different type of confectionary stores in Osaka, Japan. Followed by the analysis, the raw data of each of the two stores are scrutinized to comply with the research purpose. The summary of the data after the examination is shown in Table I.

TABLE I. Summary of the Data After Examination

| Variables | Shop Y | Shop M |
|---|---|---|
| Number of Records | 341,972 | 2,104,036 |
| Term | 1, April, 2014 to 31, May, 2017 | 11, October, 2006 to 7, June, 2017 |
| Number of Customers | 17,410 | 79,236 |
| Number of Products | 1,306 | 7,612 |
| Number of Purchasing Purposes | 6 | 9 |

## III. Comparison of Two Shops from the Viewpoint of a Service Science

### A. *Comparison and Examination Among Shops Based on Sources of Communication Messages in a Relationship*

Table I shows the comparison between two stores based on the source of communication messages in customer relationships proposed by Gronroos (2000)[3].

TABLE II. Comparison of Two Shops

| | Planned Messages | Product Messages | Service Messages | Unplanned Messages | Absence of Commu-nication |
|---|---|---|---|---|---|
| Shop Y | △ | ◎ | △ | ○ | △ |
| Shop M | ○ | △ | ◎ | △ | ○ |

The evaluation of two stores are shown in Table II. Confectionary shop Y focuses more on the development of products, by having a famous pastry chef to direct the process of development. It has a strong product message, and it seemed to have provide an unplanned message within the geographical area. On the other hand, confectionary shop M does not conduct any product developments as Shop Y does, however, it actively collects customer requests and feedbacks to improve the operating system of the store. It was also discovered that Shop M continuously distributes DM to customers, which adds another strength to Shop M's service messages. A more detailed analysis of customer needs will be conducted in the next section, based on the premises shown in Table II.

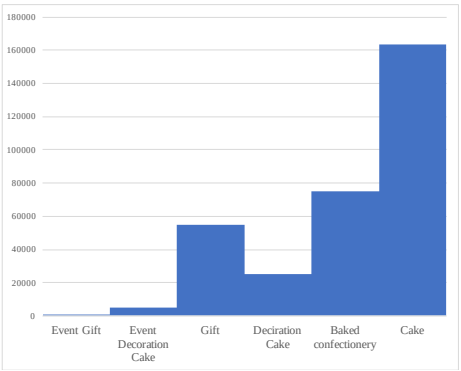
### B. *Understanding Customer Needs and Comparing Customer Segmentation Among Shops*

Each confectionary store has labeled data on its products, which indicates customer's purchasing purpose for the product (Table III). This study will first use this data in attempt to understand customer needs.

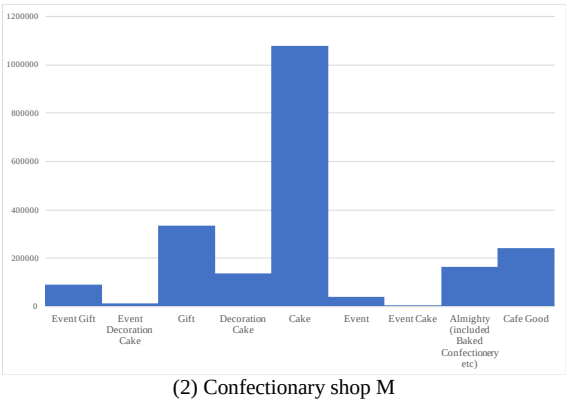
TABLE III. Purchasing Purpose at Each Confectionary Shop

| Shop Y | Shop M |
|---|---|
| Event Gift | Event Gift |
| Event Decoration Cake | Event Decoration Cake |
| Gift | Gift |
| Decoration Cake | Decoration Cake |
| Baked confectionery | Cake |
| Cake | Event |
| | Event Cake |
| | Almighty (included Baked Confectionery etc) |
| | Caf é Good |

The distribution of customer's purchasing purpose at each confectionary shop are shown next.




(1) Confectionary shop Y




(2) Confectionary shop M

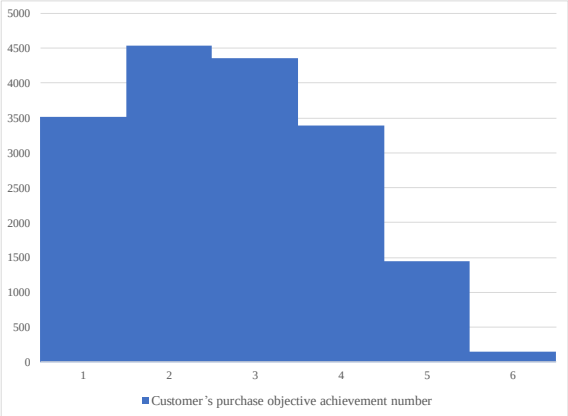
Figure 2. The distribution of customer's purchasing purpose at each confectionary shop.

Fig. 2 show that Customers demand 1st) Cake, 2nd) Baked Confectionery, 3rd) Gift at the confectionary shop Y, and do 1st) Cake, 2nd) Gift, 3rd) Cafe Good at the confectionary shop M.

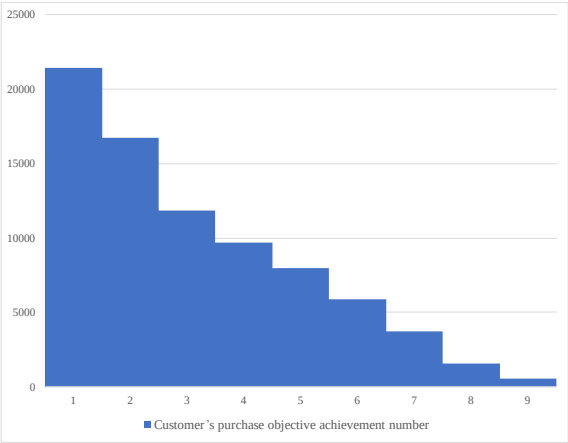
Customer's purchase objective achievement number is shown in Figure 3. Purchase objective achievement number represents the number of purchasing objectives that customers have had when entering the stores.

As shown in Fig. 3, purchasing objective achievement number 3 of shop Y is the peak and the value desired by

customers are diversified. On the other hand, shop M has 1 as its peak, which indicates that each customer has uniquely determined value when entering the store. Therefore, it is possible that the result is directly affected by the use of DM, which ultimately lead to an increase in peak of the number of achievement for purchasing purpose. As result, customer's utility value in life and store's sales is further increased.




(1) Confectionary shop Y




(2) Confectionary shop M

Figure 3. The distribution of customer's purchase objective achievement number at each confectionary shop.

## IV. Customer Demand Value and Expansion of Customer's Comsumption Behavior Hypothesis

Customer demand value is the value that are desired by customers in practice. In other words, each customer aims to achieve the pursuit value by means of goods and services. The difference between value and demand value is that value is a principle or standards that individuals already hold, and the demand value is a value that can be sought to realize the already existing value. One may not be able to realize its value because of economical or environmental constraints. It is also extremely difficult to measure those constraints individually.

In addition, past researches have attempted to analyze the effects of value on customer's purchasing behavior based on the assumption that there is a strong correlation between customer's value and their purchasing behavior (Piyathasanan et al., 2014[10]; Kim et al., 2012[11]; Miyamoto and Tsubaki, 2016[12]). Questionnaires are often used as a common method when questioning the participants regarding their values, however, it is extremely difficult and expensive to conduct surveys among all customers. There is also a problem of accuracy of the data. Since the surveys are mostly self-assessed, the data could be subjective to participants' own assessment of selves. Therefore, this paper seek to estimate the customer demand value from purchasing behaviors, based on the hypothesis that customers make purchasing decisions based on their pre-existing values.

Watabe and Tsubaki (2016)[13] proposed a consumption and usage behavioral model by conducting on-site investigation, hearing survey and literature survey based on important factors of consumer purchasing behavior. Figure 4 shows the proposed model confined and extended to the confectionary industry.

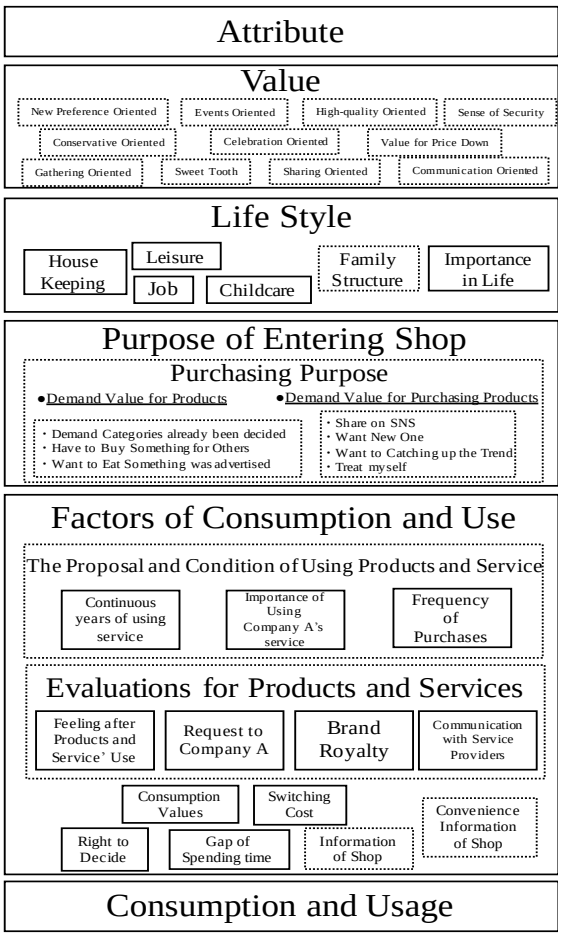



Figure 4. Consumption and usage behavioral model extended to the confectionary industry.

This study focuses and extends the aspects of Values. Based on the past experiences of pastry chefs and managers of the pastry shops, it is predicted that consumers' desired values include the following: 1) new preference oriented, 2) conservative oriented, 3) high-quality oriented, 4) "value for price down" oriented (e.g., feeling of getting a good deal), 5) celebration oriented, 6) events oriented, 7) communication oriented,

8) gathering oriented, 9) sweet tooth, 10) sense of security, 11) sharing oriented. The value thought to be derived from purchasing behaviors are shown in Figure5, and the method of calculation for identifying these values from data are shown in Table IV.

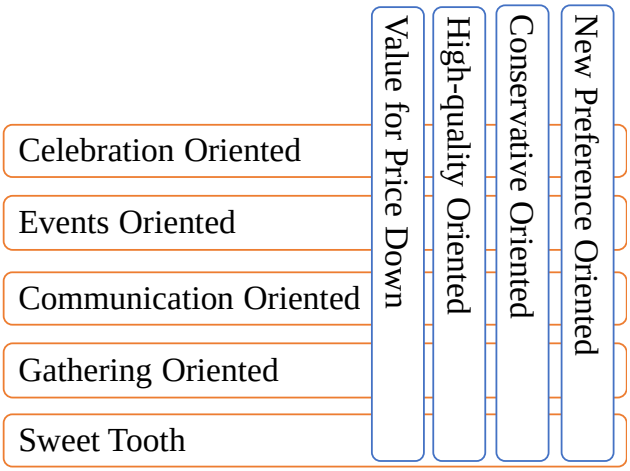



Figure 5. Customer demand value for confectionery.

TABLE IV. THE METHOD OF CALCULATION FOR VALUES

| Customer Demand Value | Method of Calculation |
|---|---|
| New Preference Oriented | Characteristic value is the total number of purchasing new products in 1 month from first sales. |
| Conservative Oriented | Characteristic value is the maximum purchasing number for each product. |
| High-quality Oriented | Characteristic value is the total number of purchasing 'Gift' is more than 2,520 yen and 'Cake' is more than 525 yen. |
| Value for Price Down | Characteristic value is the total number of purchasing 'Gift' is less than 864 yen and 'Cake' is less than 376 yen. |
| Celebration Oriented | Characteristic value is the total 1), 2), & 3) (unique date).<br>1) The purchase frequency of 'Decoration Cake', 'Event Decoration Cake', 'Event Gift' & 'Cake' with candles.<br>2) The purchase frequency of products are dedicated for congratulation on entrance and Shichi-go-san (the traditional event to celebrate the healthy growth of children).<br>3) The total number of purchasing 'Decoration Cake', 'Event Decoration Cake', 'Event Gift', 'Cake', & 'Event' in around 10 days of Coming of Age Day (second Monday at January). |
| Events Oriented | Characteristic value is the total 1) & 2) (unique date).<br>1) The purchase frequency of products are dedicated for below events.<br>2) The total number of purchasing 'Decoration Cake', 'Event Decoration Cake', 'Event Gift', & 'Cake' in around 10 days of below events.<br><Event><br>@Setsubun; 3/Feb (the traditional event, the close of winter) @Hina-matsuri; 3/Mar (the traditional event, Girl's Day) @Easter; the date depends to year @Tango-no-sekku; 5/May (the traditional event, Boy's Day) @Tanabata; 7/Jul (the traditional event, Star Event) @Keiro-no-hi; 15/Sep (the traditional event, Aged People's Day) @Halloween; 31/Oct @Christmas; 25/Dec |
| Communication Oriented | Characteristic value is the total 1), 2), 3) & 4) (unique date).<br>1) The purchase frequency of 'Gift' & 'Event Gift'.<br>2) The total number of purchasing 'Gift' & 'Event Gift' in around 10 days of below @ events.<br>3) The total number of purchasing 'Decoration Cake', 'Event Decoration Cake', 'Event Gift', 'Cake', & 'Gift' in around 10 days of below # events.<br>4) The total number of purchasing 'Gift' & 'Event Gift' in below $ events.<br><Event><br>@Valentine's Day; 14/Feb @White Day (the traditional event) #Mother's Day; second Sunday at May #Father's Day; second Sunday at June $Ochugen; Jul (the traditional event, Midsummer gift) $Osebo; 15/Nor~15/Dec (the traditional event, Year-end gift) $Onenga; 25/Dec~5/Jan (the traditional event, New year Gift) |
| Gathering Oriented | Characteristic value is the purchase frequency of three and more 'Cake', 'Baked confectionery' per one purchase (Unique date). |
| Sweet Tooth | Characteristic value is the purchase frequency of 'Cake', 'Baked confectionery' (Unique date). |

1) As for the preference of new things, characteristic value is the total number of purchasing new products in 1 month from first sales. 2) As for the conservative-oriented, the maximum purchasing number for each product is taken as characteristic value. 3) & 4) As for the high-quality oriented and "value for price down" oriented, characteristic value is taken from number of purchases of 'Gift' & 'Cake' products that are many in products and relatively higher/lower in price. The minimum value of price for gifts are 50 yen, the first quartile is 864 yen, followed by the average price of 1,874 yen, third quartile of 2,520 yen and the maximum price of 39,260 yen. Hence, gifts that are categorized as high-quality, are ones with price higher than the third quartile (i.e., 2520 yen), and gifts categorized as value-for-price are ones with price lower than the first quartile (i.e., 864 yen). Furthermore, the minimum price of the cake are 100 yen, the first quartile is 376 yen, followed by the average price of 410 yen, third quartile of 525 yen and the maximum price is 30,000 yen. Thus, high-quality cake are priced above the third quartile (i.e., 525 yen), and ones that are categorized as value-for-price are ones with price lower than the first quartile (i.e., 376 yen). 5) For celebration oriented, characteristic value is taken from the purchase frequency of candles and products that are dedicated for anniversary celebration. 6) Event oriented uses the sum of the number of purchases of goods for events throughout the year and the number of purchases near events. The sum is then taken as the characteristic value. 7) As for communication orientation, the characteristic value is derived from the number of purchases of gifts and event gifts. 8) Gathering oriented: the characteristic value is taken as a sum of number of purchases of 'Cake' and/or 'Baked confectionery'. Here, only pastries that have been purchased for three or more times are counted towards the feature value. The reason for this constraint is that the average number of people per household in Japan was 2.4 according to the Ministry of Internal Affairs and Communications (2016)[14]. 9) Sweet Tooth: the

characteristic value is calculated based on the purchase frequency of 'Cake' and/or 'Baked confectionery'. 10) Sense of security: it was considered to be difficult to conduct calculation, since there were no strong indication of the production location of material displayed in stores. 11) Lastly, as for sharing oriented, the calculation was deemed to be impossible since it is difficult to identify consumer's SNS accounts individually.

## V. Proposal of DM Distribution Method Based on Customer Demand Value

### A. Factorization Machines

This section explains the framework of the DM distribution method using Factorization Machine (i.e., FM). FM is a model proposed by Rendle (2010)[15] that combines SVM (i.e., Support Vector Machine) and Matrix Factorization. Unlike SVM, FM specializes in dealing with very sparse data. It also deals with many problems that matrix decomposition models often have (e.g., not being able to deal with general prediction problem). Hence, it is a method that specializes in solving general prediction problems corresponding to high dimensional and very sparse data. Equation (1) shows a FM model in the case where the term of interaction is set to two dimensions.

$$\hat{y}(x) := w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} \langle v_i, v_j \rangle x_i x_j \qquad (1)$$

$$\langle v_i, v_j \rangle := w_0 + \sum_{f=1}^{k} v_{i,f} \cdot v_{j,f} \qquad (2)$$

where the model parameters that have to be estimated are:

$$w_0 \in \mathbb{R}, \mathrm{W} \in \mathbb{R}^n, \mathrm{V} \in \mathbb{R}^{n \times k} \qquad (3)$$

$k \in \mathbb{N}_0^+$ is a hyperparameter that defines the dimensionality of the factorization.

### B. Extending the DM Distribution Method

#### 1) DM distributuion

DM distribution method is suggested in this research has two phases. 1st phase is selecting the customers who will get the DM. 2nd phase is deciding the contents of DM for each customer.

In the 1st phase, not only good customers but also prospective customers for shops have to be selected. In the 2nd phase, it is important for customers to be decided sending the DM which can promote sales while increasing the value of usage in customers' living. Which means that the DM which strongly recommends products based on customer's demand, makes customer's purchasing purpose increase.

In addition, DM has just two recommended products because of space limitations of DM.

#### 2) Extending the method using FM

The step of deciding recommended products in this research is explained below (figure6, 7).

Phase1: The most strongly customer demand is extracted using Chapter IV.

Phase2: Purchasing number of each product that the customer has never purchased before is predicted in each customer using FM.

Phase3-1(for customers who have not achieved all purchasing purposes): (1) The top score of purchasing purposes is elected in purchasing purpose which the customer has never purchased before (yellow). (2) The top score of purchasing purposes is elected in purchasing purpose which the customer has purchased before (blue). (3) The recommended products which are based on customer demands (phase1) are decided in each (1) & (2).

Phase3-2(for customers who have achieved all purchasing purposes): (1) The top 2 score of purchasing purpose is elected in purchasing purposes. (2) The recommended products which are based on customer demands (phase1) are decided in (1).

| Customer Demand Value | New Preference Oriented | Conservative Oriented | High-Quality Oriented | Value for Price Down | | Strong Customer Demand Value |
|---|---|---|---|---|---|---|
| Customer 1 | 0.030303 | 0.121212 | 0.090909 | 0.60606 | → | Value for Price Down |
| Customer 2 | 0.038961 | 0.045454 | 0.324675 | 0.045454 | → | High-Quality Oriented |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | | |
| Customer 17410 | 0.218181 | 0.072727 | 0.236363 | 0.036363 | → | New Preference Oriented |

Figure6. Phase1: Step of deciding strong customer demand value.

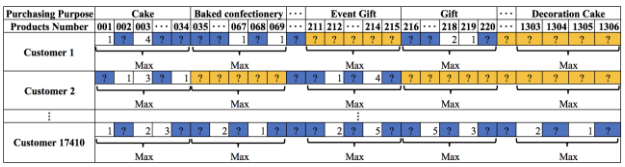



Figure7. Phase2;2-1;2-2: Step of deciding recommended purchasing purpose.

The step from model creation to model evaluation using FM is shown below and in Figure8.

Step1 (Preparing training and test data): As training data, customer's each purchasing products are calculated per one entering the shop in the period A. As test data, customer's each purchasing products are calculated per one entering the shop in the period B.

Step2 (Model creation): The prediction model (explanatory variables: User/Item, target variables: purchasing number of each products) using FM is constructed. The optimal model is decided with cross validation.

Step3 (Prediction): As explanatory variables are User/Item, the purchasing number of each products are predicted using the step2's model.

Step4 (Model evaluation): The hyperparameter is decided so that mean squared error between predicted value and actual measured value (i.e., MSE) becomes smallest.

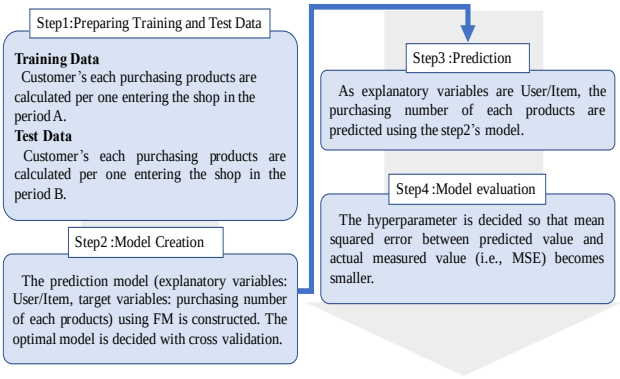



Figure8. Step from model creation to model evaluation.

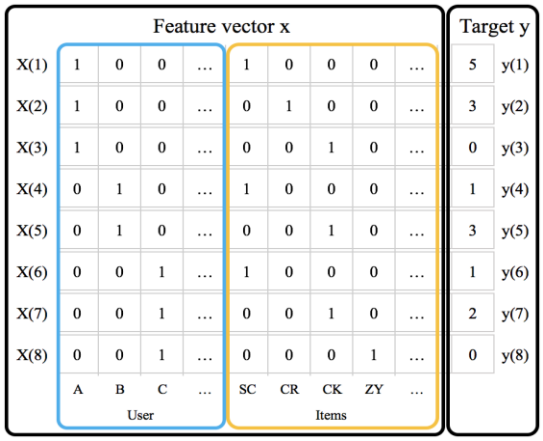



Figure9. Data format for model construction.

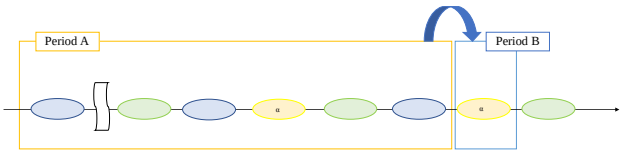



Figure10. Prediction model.

## VI. Conclusion

This study aims to compare the customer sales data of two different types of stores in the confectionery industry, using the framework which is "Sources of Communication Messages in a Relationship", and this study proposed the method to categorize the customer's pursuit values for each store based on sales data. This paper proposes the new approach that can promote sales while increasing the value of usage in customers' living, based on devising a method to improve the efficiency of customer direct mail delivery, using Factorization Machines.

In comparison of two shops, the results indicated the strong points in each two shops are different from on-site investigation. Moreover, the value desired by customers to each two shops is also different from customer sales data. Therefore, the value customers desire in practice are defined as "customer demand value".

In customer direct mail delivery, customer purchase behavior prediction makes 2 parts below possible. 1) recommends purchasing purposes which the customer has never purchased before, 2) recommends purchasing purpose which the customer has purchased before. Further, this paper proposed adding the information of customer demand value to DM which enables recommend more customized products for customers. It can be considered that distribution of DM is able to promote sales while increasing the value of usage in customers' living.

## References

[1] C. Grönroos, *Service Management and Marketing Customer Management in Service Competition*, 3rd ed., John Wiley & Sons Limited, 2007, ch.10.
[2] E. S. Don, "Integrated marketing communications," *Journal of Promotion Management*, vol. 1, no. 1, pp. 99–104, 1992.
[3] C. Grönroos, "Creating a relationship dialogue: Communication, interaction, value," *The Marketing Review*, vol. 1, pp. 5–14, 2000.
[4] T. Duncan and S. Moriarty, *Driving Brand Value. Using Integrated Marketing to Manage Profitable Stakeholder Relationships*, New York: McGraw Hill, 1997.
[5] F. Simon and A. Lynda, "A relational approach to direct mail consumption: The perspective of engagement regimes," *European Journal of Marketing*, vol. 49, no. (9/10), pp. 1527–1562, 2015.
[6] Canada Post Corporation, "Understanding the impact of physical communications through neuroscience," *Canada Post / True Impact Marketing*, 2015.
[7] J. C. Gázquez-Abad, M. H. Canniére, De and F. J. Martínez-López, "Dynamics of customer response to promotional and relational direct mailings from an apparel retailer: The moderating role of relationship strength," *Journal of Retailing*, vol. 87, no. 2, pp. 166–181, 2011.
[8] V. L. Miguéis, A. S. Camanho, and J. Borges, "Predicting direct marketing response in banking: Comparison of class imbalance methods," *Service Business*, vol. 11, no. 42, pp. 1–19, 2017.
[9] C. Chen, C. Hou, J. Xiao, and X. Yuan, "Purchase behavior prediction in e-commerce with factorization machines," *IEICE Transactions on Information and Systems*, E99D(1), pp. 270–274, 2016.
[10] B. Piyathasanan, C. Mathies, M. Wetzels, P. G. Patterson, and K. De Ruyter, "A hierarchical model of virtual experience and its influences on the perceived value and loyalty of customers," *International Journal of Electronic Commerce*, vol. 19, no. 2, pp. 126–158, 2014.
[11] C. Kim, R. D. Galliers, N. Shin, J. H. Ryoo, and J. Kim, "Factors influencing Internet shopping value and customer repurchase intention," *Electronic Commerce Research and Applications*. Elsevier B.V., vol. 11, no. 4, pp. 374–387, 2012.
[12] Y. Miyamoto and M. Tsubaki, "A study on the promotion of Purchase based on analyzing the effect of the service by combination between customers type and service providers type using the bayesian network," in *Proc. International Social Sciences Academic Conference*, pp. 12-17, 2016.
[13] H. Watabe and M. Tsubaki, "A study on the matching between customers and service providers using the system for analyzing the effects of the service by customer type," *The Japan Society for Management Information*, vol. 24, no. 4, pp. 231–238, 2016.
[14] Ministry of Internal Affairs and Communications (2016), Family Income and Expenditure Survey (Result of Two-or-more-person Households) 2016 Yearly Average, [Online][Accessed 25th August 2017], Available at: <http://www.stat.go.jp/data/kakei/2016np/index.htm>
[15] S. Rendle, "Factorization Machines," in *Proc. IEEE International Conference on Data Mining Factorization*, pp. 995−1000, 2010.

**Yuya Miyamoto** is a graduate student in the Department of Informatics, Graduate School of Informatics and Engineering, the University of Electro-Communications, Japan. He received his BS in Informatics and Engineering from the University of Electro-Communications, Japan His research interests relate to services marketing, various data analysis. Recent his research is a project in collaboration between Academic and Industry.

**Michiko Tsubaki** is a professor in the Department of Informatics at the University of Electro-Communications, Japan. She received her BS in Applied Mathematics and her MS and DS in Management Science from Tokyo University of Science, Japan. She was a Visiting Scholar at Oxford University in 1992. Her recent research interests focus on service marketing, marketing strategy, sales skills, big data analysis. She was the Associate Editor of the *Journal of the Japanese Society for Quality Control* (JSQC) from 1990 to 1997, the Associate Editor of the *Journal of the Japan Industrial Management Association* from 2000 to 2001 and the Associate Editor of the *Journal of the Japanese Society of Applied Statistics* from 2000 to 2008. She has been a Program Committee Member of the World Multi-Conference on Systemics, Cybernetics and Informatics since 2005 and a Program Committee Member of the International Symposium on Academic Globalization since 2007.