Asia Pacific Journal of Information Systems
Vol. 29 No. 4 (December 2019), 752-770

ISSN 2288-5404 (Print) / ISSN 2288-6818 (Online)
https://doi.org/10.14329/apjis.2019.29.4.752

Special Issue: Data Analytics in Artificial Intelligence Era

Multi-Purpose Hybrid Recommendation System on Artificial Intelligence to Improve Telemarketing Performance

Hyung Su Kim[a], Sangwon Lee[b,*]

[a] *Associate Professor, Department of Industry & Management Engineering, Hansung University, Korea*
[b] *Associate Professor, Department of Computer & Software Engineering, Wonkwang University, Korea*

ABSTRACT

The purpose of this study is to incorporate telemarketing processes to improve telemarketing performance. For this application, we have attempted to mix the model of machine learning to extract potential customers with personalisation techniques to derive recommended products from actual contact. Most of traditional recommendation systems were mainly in ways such as collaborative filtering, which predicts items with a high likelihood of future purchase, based on existing purchase transactions or preferences for products. But, under these systems, new users or items added to the system do not have sufficient information, and generally cause problems such as a cold start that can not obtain satisfactory recommendation items. Also, indiscriminate telemarketing attempts can backfire as they increase the dissatisfaction and fatigue of customers who do not want to be contacted. To this purpose, this study presented a multi-purpose hybrid recommendation algorithm to achieve two goals: to select customers with high possibility of contact, and to recommend products to selected customers. In addition, we used subscription data from telemarketing agency that handles insurance products to derive realistic applicability of the proposed recommendation system. Our proposed recommendation system would certainly solve the cold start and scarcity problem of existing recommendation algorithm by using contents information such as customer master information and telemarketing history. Also. the model could show excellent performance not only in terms of overall performance but also in terms of the recommendation success rate of the unpopular product.

Keywords: Telemarketing, Recommendation, Artificial Intelligence, Machine Learning, Customer Relationship Management

*Corresponding Author. E-mail: sangwonlee@wku.ac.kr Tel: 82638506566

Ⅰ. Introduction

Past recommendation systems were mainly in ways such as collaborative filtering, which predicts items with a high likelihood of future purchase, based on existing purchase transactions or preferences for products (Ekstrand et al., 2010). Such an algorithm should provide sufficient preference information to present a reliable recommendation to the user. However, new users or items added to the system do not have sufficient information, and generally cause problems such as a cold start that can not obtain satisfactory recommendation items (Hu and Pu, 2010; Reshma et al., 2016). In recent research, a hybrid model has been proposed to improve the problems of the existing recommendation system, and it has been applied to various industries such as movies (Christakouet al., 2007), e-learning (Kardan et al., 2009; Khribi et al., 2011), music (Wang and Wang, 2014), and web services (Burke, 2007).

Telemarketing is one of the most widely used services for marketing campaigns (Moro et al., 2014). The reason why we are actively considering Telemarketing in modern CRM is the strategic importance in modern marketing. In modern CRM, telemarketing is an important consumer connectivity channel and a meaningful e-commerce approach. In making purchases, modern consumers seek the convenience of buying and the diversity of buying methods. In order to satisfy this pursuit, telemarketing serves as a great marketing channel. Therefore, a recommendation system for telemarketing can be an important capability for improving business performance. However, indiscriminate telemarketing attempts can backfire as they increase the dissatisfaction and fatigue of customers who do not want to be contacted. For this reason, telemarketing needs to first consider whether a customer is eligible for telemarketing before using the recommendation system. To this purpose, this study presented a multi-purpose hybrid recommendation algorithm to achieve two goals: to select customers with high possibility of contact, and to recommend products to selected customers. In addition, we used subscription data from telemarketing agency that handles insurance products to derive realistic applicability of the proposed recommendation system. Automated Product Recommendation System is a system that uses information filtering to provide interesting information items to users. If the existing legacy systems are a technique that obtains a user's information profile by questioning the user's personal identity, areas of interest, preferences, etc., this recommendation system is based on this information to recommend or provide information and products that are appropriate for the customer's psychological and preference information. The system can be applied in various ways to search for movies, music, news, books, research topics, search queries and merchandise.

CRM is a marketing methodology that clearly distinguishes itself from mass marketing, segmentation marketing, and niche marketing in the past, and has emerged based on individual marketing, one-to-one marketing, and relocations. CRM actively utilizes customer information through combination with call center and campaign management tools, prioritizing customer profitability, and BPR is implicit in changing thinking within the enterprise. The use of internal and external data related to a company's customers is the same as database marketing. However, in the case of CRM, the customer interface is much more diverse than database marketing, and the acquisition of this diverse information is conducted throughout the enterprise. CRM conducts marketing through strategies that can actively manage and induce customers through cycles such as acquiring new customers, maintaining good customers, enhancing customer value, activating potential customers, and maximizing customer value by performing segmentation of customer data.

In a financial industries such as insurance, face-to-face consulting often requires a case where a plurality of products must be recommended at the same time. For example, a case when the first recommendation product does not show a positive response (Lacerda, 2017). However, in the nature of face-to-face consultations, the difference between the cost of recommending only one product and the cost of recommending multiple products together is not significant, so it is possible to recommend the later products one after another. While the general product recommendation system executes recommendations for a single product, the system proposed in this study recommends a variety of products simultaneously at a time. So we attach the word multiple to the name of this proposed system. Nevertheless, recommending an excessive number of products at the same time can cause complaints to customers. Therefore, this study limited the number of recommended products to a maximum of three. So, we limit our recommendation to only the third rank recommendation, and to evaluate the performance of objective models, we should not only consider first-priority but also the recommended accuracy up to third-priority.

Ⅱ. Literature Review

2.1. Collaborative Filtering-based Approach

Collaborative filtering is one of the most popular recommendation system technologies and is an algorithm that recommends items that users might be interested in based on similarity with other users (Bobadilla et al., 2013; Hu et al., 2010). In general, collaborative filtering is classified as follows (Breese et al., 1998).

(1) Memory-based collaborative filtering: The memory-based collaborative filtering is based on the user-item ratings matrix (Ghazarian and Nematbakhsh, 2015). There is typically user-based collaborative filtering, which analyzes similarities between users, and item-based collaborative filtering, which measures similarities among items. In order to calculate similarity, cosine similarity, Pearson correlation, and Euclidean distance are used (Arsan et al., 2016; Sarwar et al., 2001).

(2) Model-based collaborative filtering: Model-based collaborative filtering is an approach to analyzing user preferences based on models (Langseth and Nielsen, 2015). A representative model for this approach is Bayesian network, clustering model, and latent variable model (Breese et al., 1998; Langseth et al., 2015; Mobasher et al., 2006).

However, as mentioned in the introduction, collaborative filtering in the recommendation system still causes problems such as cold start and scarcity. For this reason, research has been conducted to improve collaboration filtering for a long time. For example, Sarwar (2001) solves scarcity with content-based ratings, SVD(Singular Value Decomposition)-based algorithms and item-based collaborative filtering, and improves scalability by using SVD-based dimension reduction and clustering techniques. Acilar and Arslan (2009) proposed an improved collaborative filtering model including the distribution of space and the interrelationships of clusters. In this way, the collaborative filtering model has been researched in various approaches and used in actual services.

For example, Amazon.com is known to have introduced a successful recommendation system by item-based collaboration filtering, which recommends products that are similar to those purchased or evaluated by users (Linden et al., 2003).

2.2. Rule-based Approach

Association analysis or market basket analysis is an algorithm that extracts items that are related or concurrent in the transaction data, and find the customer's purchase pattern and generate association rules (Chen et al., 2005). Association analysis generally evaluates rules through support, confidence, and lift, and selects important rules through three indicators in the extracted association rule set (Mossong et al., 2008; Ordonez, 2006). In this analysis, as the number of products increases, the amount of association rules increases exponentially, which may degrade the performance of the recommendation system. Therefore, the association analysis should filter the association rules based on the minimum support and the minimum confidence, and this minimum criterion can be determined by the user utilizing the recommendation system (Wang et al., 2004).

As with collaborative filtering, a variety of methodologies are being researched to solve problems such as cold start in a rule-based approach that uses association analysis (Min and Zhu, 2013; Shaw et al., 2010). For example, Voditel and Deshpande (2013) developed an association rule filtering method that reduces the exponential set of items in the stock market based on support and confidence for the stock market portfolio recommender system. In addition, Lin, Alvarez and Ruiz (2002) also considered associations between users along with associations between items. And they proposed a methodology that has the number of rules within a specified range by adjusting the minimum support of the association rule mining process, unlike the existing association analysis, which should set the minimum support in advance.

2.3. Hybrid-based Approach

Hybrid-based recommendation system is a method that combines various approaches or utilizes supplementary data, and it improves the performance of the recommendation system by improving the disadvantages that existed in individual algorithms (Burke, 2002; Wang et al., 2018). Because of these advantages, hybrid-based approach is the most actively studied area in the recommendation system. An example of combining various recommendation algorithms in a hybrid-based recommendation system is as follows. Chen et al. (2014) proposed a two-step process for discovering a set of related items using item-based collaborative filtering and applying it to sequential pattern mining to recommend items to users in an e-learning environment. In addition, Kim et al. (2017) studied hybrid algorithms that recognize contexts of documents using Convolutional Neural Network and Probabilistic Matrix Factorization.

Next, as an example of a hybrid recommendation system utilizing supplementary data, Paradarami, Bastian and Wightman (2017) used review information from the foodservice industry to predict restaurant preference ratings. Yang et al. (2008) also used location information for recommendation in mobile shopping, and Hu et al. (2009) used user preference and music content information for music recommendation.

Ⅲ. Proposed Method

There are two major relaterd works; collaborative filtering based approach and rule-based approach. But, the first appoach as collaborative filtering in the recommendation system still causes problems such as cold start and scarcity. In addition, the second appoach as the association analysis should filter the association rules based on the minimum support and the minimum confidence, and this minimum criterion can be determined by the user utilizing the

<Table 1> Data Schema

Dataset	Description	Type
Customer master information	Age	Numeric
	Segment number	Categorical
	Point	Numeric
	Date of acceptance	Categorical
	Longevity	Numeric
	Number of payment means	Numeric
	Cumulative point	Numeric
	Number of fuel purchases	Numeric
	Number of payment means of fuel	Numeric
	Total fuel purchase amount	Numeric
	Total fuel cumulative point	Numeric
	Major fuel product	Categorical
Telemarketing history	Whether of contact success	Categorical
	Major contact type	Categorical
	Number of contact result types	Numeric
	Average contact time	Numeric
	Average time of contact success	Numeric
	Call success rate	Numeric
	Positive call success rate	Numeric
Insurance subscription history	Major insurers	Categorical
	Number of subscriptions	Numeric
	Number of subscription insurance types	Numeric
	History of home insurance subscription	Categorical
	History of accident insurance subscription	Categorical
	History of cancer insurance subscription	Categorical
	History of silver insurance subscription	Categorical
	History of child insurance subscription	Categorical
	History of driver insurance subscription	Categorical
	History of dental insurance subscription	Categorical
	History of whole life insurance subscription	Categorical
	Whether of home insurance subscription	Categorical
	Whether of accident insurance subscription	Categorical
	Whether of cancer insurance subscription	Categorical
	Whether of silver insurance subscription	Categorical
	Whether of child insurance subscription	Categorical
	Whether of driver insurance subscription	Categorical
	Whether of dental insurance subscription	Categorical
	Whether of whole life insurance subscription	Categorical

recommendation system. So, we propose a new approch as hybrid-based recommendation. Hybrid-based recommendation system is a method that combines various approaches or utilizes supplementary data, and it improves the performance of the recommendation system by improving the disadvantages that existed in individual algorithms.

3.1. Data

This study utilized customer master information, telemarketing history, point history, and insurance subscription history provided by telemarketing agencies in domestic insurance industry to verify the performance of the proposed recommendation system. Among them, 39,839 successful transactions and 39,839 unsuccessful transactions were selected from the telemarketing history between October 2016 and September 2017. At this time, the information of 69,338 customers corresponding to the selected transaction was also collected.

When data collection is complete, configure the dataset for analysis. There were 140 variables in the initial dataset for this study. We removed unnecessary parameters through Feature Selection to improve the speed and accuracy of the recommendation system. As a result, 51 variables were selected as shown in <Table 1>.

3.2. Development of Multi-Purpose Hybrid Recommendation Model

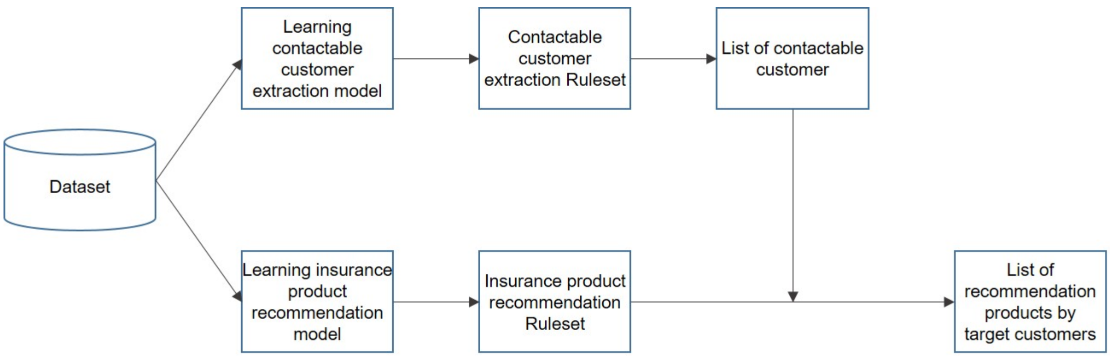
The purpose of this study is to select the customers with high possibility of contact first and to recommend personalized insurance products with high possibility of subscribing to the selected customers. To achieve this, we developed a two-step hybrid recommendation algorithm that combines a model for predicting customers with high contact potential and a model for recommending personalized products, as shown in <Figure 1>.

3.2.1. Contactable Customer Extraction Model

The model to select the prospective customer leads to the process of dataset generation, model learning, model application and prediction. The algorithms for learning the model considered decision tree, artificial neural network, support vector machine, logistic regression, random forest, gradient boosting. 'Contractable customers' means customers who are prospects accessible through this marketing channel.

Create Dataset

First, we configure a dataset for learning the Contactable Customer Extraction Model. The dataset



<Figure 1> Architecture of a Multi-purpose Hybrid Recommendation System

used the customer master information and the point history as independent variables and the success of the telemarketing history as the dependent variables. The dataset is divided into a training set and a test set for consistent performance comparison of the prediction model. Here, the training set is used to learn the model, and the test set is used to evaluate the performance by applying the learned model. The ratio of the training set to the test set is divided into 7: 3.

Learning Contactable Customer Extraction Model

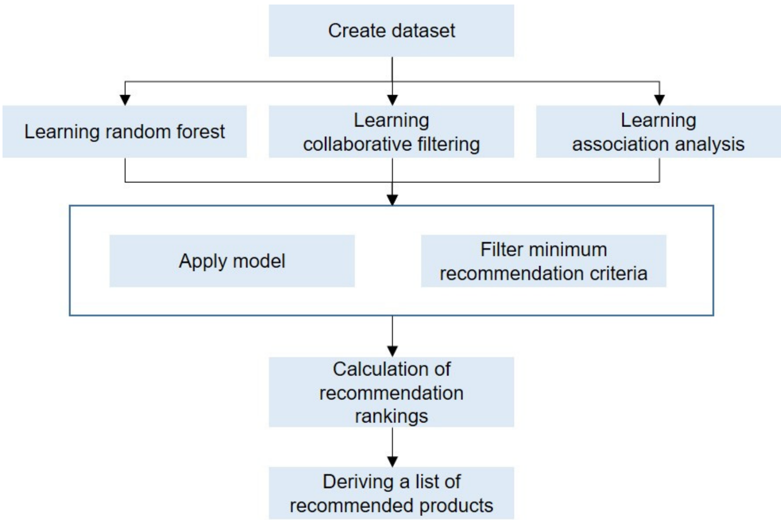
This step is to learn contactable customer extraction model based on machine learning algorithms. The most prominent phenomenon of customer data these days is the occurrence of tremendous data. Machine learning uses this big data to preprocess and optimize the data to be learned so that machine learning results can be practical if learning effects are maximized. Machine learning is mainly used to solve tasks that are difficult to design or difficult to program. Most machine running algorithms first quantify complex relationships by identifying the characteristics of potential mechanisms created by the data, and then use this identified pattern to make predictions of new data. Machine running is associated with data mining, statistical and mathematical optimization issues in that it extracts useful rules, knowledge representation, or judgment criteria from the data. In this study, we evaluate the performance of six candidate algorithms through iterative experiments, and select the algorithm that shows the best performance as the algorithm for the Contactable Customer Extraction Model.

Model Application And Prediction

When the learning of the model is completed, apply a model to the test set to predict the success of the contact. The prediction results are generated with a score between 0 and 1, and the customers to be contacted are selected in descending order of prediction scores.

3.2.2. Insurance Product Recommendation Model

The insurance product recommendation algo-



<Figure 2> Development Process of the Insurance Product Recommendation Algorithm

rithm for this study is developed through the process as shown in <Figure 2>. This process consists of six steps and compares the performance of the generated model-based on random forests, collaborative filtering and association analysis.

Create Dataset

The process of developing a recommendation algorithm and comparing performance begins by creating a dataset for use in the learning model. At this stage, dataset for each model are constructed using customer master information, telemarketing history, and subscription history. At this point, the random forest model should be transformed into a data frame, collaborative filtering should be a ratings matrix, and association analysis should be in the form of transactions. In addition, since the subscription failure transaction is not required when learning the recommendation model, only the subscription successful transactions are extracted and used for model learning.

Learning Recommendation Models

The recommendation algorithms to compare performance in this study are random forest model, collaborative filtering model, and association model. Among them, the association analysis model can select the priority of the association rule based on the confidence, support, and lift, and the performance of the recommendation model is determined according to the selected criteria. Next, the collaborative filtering should select a metric to calculate the similarity, and typically use Pearson correlation, cosine similarity, Euclidean distance, and the like.

Apply Model and Filter Minimum Recommendation Criteria

When the recommendation model for a recommendation system is learned, each model is applied to a test set to predict the recommendation product each transaction. As a result, the random forest model results in a recommendation score between 0 and 1, collaborative filtering is a recommendation rating, and the association analysis generates a association rule. The goal of the insurance product recommendation algorithm is to use the proposed recommendation system to recommend insurance products that are highly available to improve telemarketing performance. However, recommendations based only on priority are considered relative ranking among products, and the absolute recommendation level for each product is unknown. Therefore, ranking-based recommendation systems have problems recommending high-priority products even if they have low recommended scores. This leads to a decrease in the recommendation success rate and a decrease in reliability of the recommendation system itself. For this reason, if the absolute recommendation score falls below a certain threshold, it is necessary to remove it from the recommendation list. In the case of this study, the minimum recommendation criteria can be determined based on recommendation score based on random forest model, recommendation rating based on collaborative filtering model, and support and confidence based on association analysis model. Although there are no metrics or criteria for determining values to filter recommendations, it is possible to set an appropriate criterion according to industry, product, situation, user's experience and the like.

Calculation of Recommendation Rankings

When the recommendation score prediction and the product filtering based on the minimum criteria are completed through the application of the recommendation model to the test set, the recommendation product ranking can be calculated using the recommendation result generated. At this time, the random forest model assigns a recommendation priority from a product having a high recommendation score. Collaborative filtering model is based on recommendation rating by product, and association analysis model will calculate the recommendation rankings of personalised insurance products based on association rule. The products predicted in the same order can be ranked according to the importance of the predetermined goods. In this study, through discussion with current employees, the recommendation priority was determined based on the success rate and unit price of each insurance product.

Deriving a List of Recommended Products

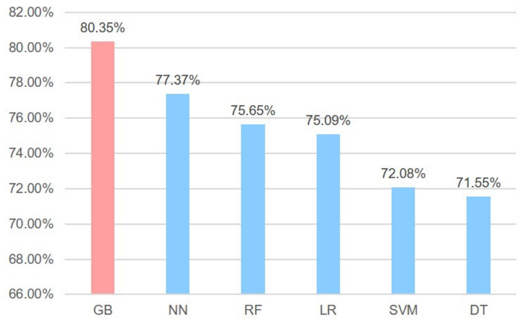
When all of the above steps have been completed, a list of recommended products is derived for each transaction in the test set. Using this method, we compare the performance of the recommendation model and evaluate it, and utilize the best performing algorithm to the insurance product recommendation model.

3.3. Model Performance Evaluation and Final Algorithm Selection

Once the algorithms of the multi-purpose hybrid recommendation system is generated, the final algorithm shall be selected by assessing the performance of the model. Generally, recall, specificity, precision, etc. are used to evaluate the performance of the prediction model. The performance of the model should be compared by determining appropriate evaluation metrics according to the purpose.

3.3.1. Performance Evaluation of Contactable Customer Extraction Model

The number of customers who can contact tele-



<Figure 3> Performance Comparison of Contactable Customer Extraction Model

marketing agencies that provide data for this study for a month is about 20 percent of all customers. Therefore, it is necessary to be able to effectively select the telemarketing target customers through the contactable customer extraction model, and the performance of the model is determined according to how precisely the valid customers are selected. From this perspective, the performance of the contactable customer extraction model can be determined with precision, which is the ratio of the significant customers among the selected customers. As a result of comparing the six candidate algorithms as shown in <Figure 3>, the best performance of the gradient boosting is determined as the algorithm for the contactable customer extraction model.

3.3.2. Performance Evaluation of Insurance Product Recommendation Model

In this study, if the recommended product is hit within third ranks, it is judged as recommendation success. From this point of view, regardless of the recommendation rankings, the recall rate, which considers the recommendation success itself as an important factor, is determined as the performance evaluation measure of the recommendation model.

The association analysis model, which is the first

<Table 2> Accuracy for Association Analysis Models

Metrics of association analysis	Contactable customer extraction model	Random sampling
Confidence	46.85%	40.61%
Support	45.39%	39.72%
Lift	21.52%	18.33%

<Table 3> Accuracy for Collaborative Filtering Models

Metrics of collaborative filtering	Contactable customer extraction model	Random sampling
Pearson correlation coefficient	52.80%	45.57%
Cosine similarity	51.49%	45.27%
Euclidean distance	49.35%	42.60%

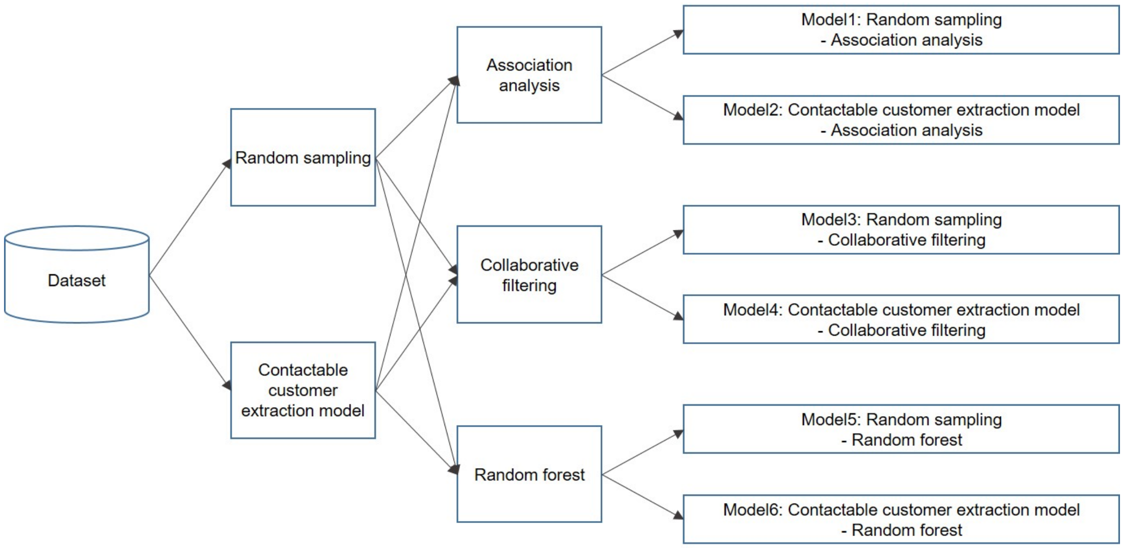
algorithm of the insurance product recommendation model, can prioritize generated associative rules through confidence, support, lift. In this study, the recommended performance of the three indicators is compared through the experiment as shown in <Table 2>. As a result, when the priority of the association rule was determined based on the confidence, 46.85% and 40.61% were the highest in the contactable customer extraction model and random sampling, respectively. The support was 45.39%, 39.72%, and the improvement was 21.52%, 18.33%, respectively, and the recommendation accuracy was lower than that of the confidence based model. Therefore, in this study, confidence is determined as a priority generation criterion of association rules. Confidence refers to the percentage of transactions involving commodity X that also include commodity Y (conditional probability). Support refers to the ratio of the total number of transactions that include both goods X and Y.

Next, user-based collaborative filtering is selected as a model for the recommendation system that is most suitable for utilizing the data set of this research among the collaboration filtering algorithms. In addition, typical Pearson correlation, Cosine similarity, and Euclidean distance were selected as the method for measuring similarity among users in the model.

In this study, the performance of the recommendation model generated by each similarity metric through experiments is compared as shown in <Table 3>. As a result, when the Pearson correlation coefficient was used, the best performance was achieved by 52.80% in the contactable customer extraction model and 45.57% in the random sampling. On the other hand, the cosine similarity was 51.49%, 45.27%, and Euclidean distance was 49.35% and 42.60%, respectively, resulting in lower performance than Pearson correlation coefficient. Therefore, in this study, the Pearson correlation coefficient with the highest recommendation accuracy among the similarity metrics is determined as a metric learning user-based collaborative filtering.

Ⅳ. Evaluation

If the algorithm of both the contactable customer extraction model and the insurance product recommendation model are determined, the performance of the multi-purpose hybrid recommendation model



<Figure 4> Research Model of Multi-purpose Hybrid Recommendation System

combining the two models is evaluated. The step of selecting customers for telemarketing compares the performance of two methods of extracting customers with high possibility of contact and the method of randomly sampling. And the step of recommending insurance products compares the performance of three algorithms of association analysis, user-based collaborative filtering, and random forest. Therefore, a total of six models are generated as shown in <Figure 4>, and the model proposed in this study is the sixth model that combines the contactable customer extraction model with the random forest.

Once all model learning has been completed, a separate test set must be created to evaluate the performance of each model. First of all, we apply the contactable customer extraction model to the whole test set, then select the customers who have the highest possibility of contact with the top 50% and use it as a test set for evaluating the contactable customer extraction model. It also organizes the test set to evaluate the random sampling by screening 50% of customers through random sampling. As a result, the test set of the contactable customer extraction model included 9,570 contact success transactions out of 11,910, showing a contact success rate of 80.35%. And 8,406 cases were included in the test set of the random sampling, resulting in 70.58% contact success rate.

4.1. Model Performance Comparison of All Products

As described in the introduction, face-to-face counseling sometimes recommends a plurality of products continuously without ending with one recommendation. Therefore, we evaluated all of the suggested success rates up to the top three products to compare the exact performance of the model.

As a result of applying the first model, the random sampling-association analysis model to the test set, the recommendation product was estimated to be 2,667 successful (22.39%) when only the first insurance product was considered. Next, after the success in the first recommendation is excluded, the second recommendation product is judged as a recommendation product for the remaining transactions. The number of successes of second rank recommendation is 1,388. A total of 4,055 successful recommendations (34.05%) were obtained by adding the first and second successes. When considering the recommended products of the third rank in the above way, 4,837 (40.61%) recommendation success is shown.

The second model, the contactable customer extraction model-association analysis combined model, evaluates the performance in the same way as the first model. As a result, 3,631 cases (30.49%) were ranked in the first recommendation, 5,067 cases (42.54%) in the second, and 5,580 cases (46.85%) in the third. The results of the second model show better performance than the results of the first model in all sections.

Next, we evaluate the performance of the third and fourth combined models using user-based collaborative filtering. The results of the third model were 2,812 (23.61%) in the first recommendation, 4,446 (37.33%) in the second recommendation and 5,427 (45.57%) in the third recommendation. The results of the fourth model were 3,701 (31.07%) in the first recommendation, 5,325 (44.71%) in the second recommendation and 6,288 (52.80%) in the third recommendation. As in the association analysis, the combined model of user-based collaborative filtering showed better recommendation performance of the contactable customer extraction model than random sampling. In addition, user-based collaborative filtering in all cases has better performance than association analysis model.

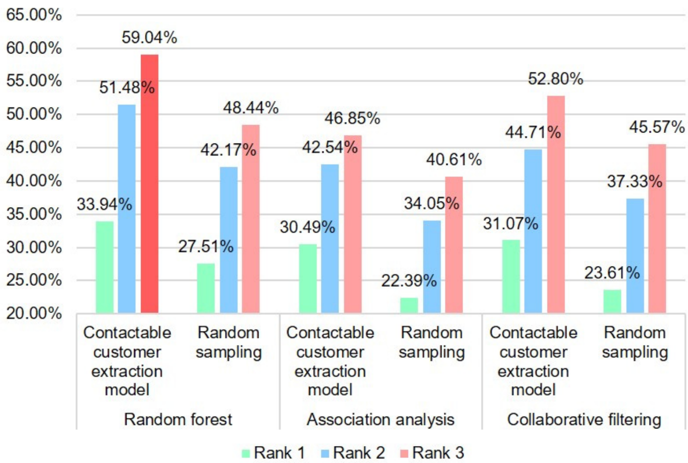
Finally, we evaluate the result of the combined model using random forest. The results of the fifth model, which combines random sampling and random forest, resulted in a recommendation of 3,276 (27.51%) in the first recommendation. This suggests that the recommendation success rate is improved by 22.83% compared to the association analysis model and by 16.50% compared to the user-based collaborative filtering model. Next, the recommendation of 5,022 (42.17%) was successful in the second recommendation, and the recommendation success rate was improved by 23.85% and 12.96%, respectively, compared with the competition model. In the third recommendation, 5,769 (48.44%) were successful in recommendation, and 19.27% and 6.30% of the recommendation performance were improved, respectively, when compared with the competition model.

In addition, the results of the model combining the contact lead extraction model-random forest proposed in this study were recommended 4,042 cases (33.94%) in the first priority recommendation. It is 11.32% higher than the association analysis model, 9.21% better than the user-based collaborative filtering model, and 23.38% higher than the random sampling-random forest combined model. In the second recommendation, the recommendation of 6,161 (51.48%) was successful. The performance of 21.00% and 15.14% was improved compared to the competition model, and the recommendation performance of 22.08% was improved than random sampling. Next, as a result of accumulating all of the third recommendations, 7,032 transactions (59.04%) out of 11,910 transactions succeeded and showed excellent performance. It is an improvement of 26.02% and 11.83% compared to the competition model and 21.89% improvement over random sampling. <Table 4> compares the performance of the above six models.

The results of the comparison of the recommendation success rates of the multi-objective hybrid

<Table 4> Performance Comparison of Multi-purpose Hybrid Recommendation System of All Products

First step	Second step	Rank	Num of target	Num of success	Num of fail	Num of cumulative success	Recommendation success rate
Random sampling	Association analysis	1	11,910	2,667	9,243	2,667	22.39%
		2	9,243	1,388	7,855	4,055	34.05%
		3	7,855	782	7,073	4,837	40.61%
	Collaborative filtering	1	11,910	2,812	9,098	2,812	23.61%
		2	9,098	1,634	7,464	4,446	37.33%
		3	7,464	981	6,483	5,427	45.57%
	Random forest	1	11,910	3,276	8,634	3,276	27.51%
		2	8,634	1,746	6,888	5,022	42.17%
		3	6,888	747	6,141	5,769	48.44%
Contactable customer extraction model	Association analysis	1	11,910	3,631	8,279	3,631	30.49%
		2	8,279	1,436	6,843	5,067	42.54%
		3	6,843	513	6,330	5,580	46.85%
	Collaborative filtering	1	11,910	3,701	8,209	3,701	31.07%
		2	8,209	1,624	6,585	5,325	44.71%
		3	6,585	963	5,622	6,288	52.80%
	Random forest	1	11,910	4,042	7,868	4,042	33.94%
		2	7,868	2,089	5,779	6,131	51.48%
		3	5,779	901	4,878	7,032	59.04%



<Figure 5> Recommendation Success Rate Comparison of Multi-purpose Hybrid Recommendation System of All Products

recommendation models proposed in this study are shown in <Figure 5>. As a result of comparing the performance of each model, it showed superior performance when selecting customers based on the contactable customer extraction model rather than the random sampling. In addition, when the random forest algorithm is compared with the association analysis or the collaborative filtering, a good recommendation success rate is obtained in all cases.

4.2. Model Performance Comparison of Unpopular Products

Of course, the referral success rate for the entire transaction is a very important factor in the recommendation system. However, as can be seen from the Long Tail Law (Brynjolfsson et al., 2011), the success rate of recommendation for unpopular products is also not negligible. In particular, when predicting new data by learning past data such as machine learning, items with a high proportion of dependent variables are easy to predict, but items with low specific weight can be relatively difficult to predict. In this study, injury, cancer, dental, and silver insurance, which have a high proportion of subscription, account for most of the subscription transactions, and the remaining four products account for only about 400 ~ 600 fewer subscriptions among 11,910. Therefore, we also need to confirm the recommendation success rate of drivers, children, housing, and life insurance that have a low proportion of subscriptions. At this time, the random sampling method extracted 408 unpopular products transactions, while the contactable customer extraction model succeeded in extracting 526 cases, which showed better selecting results as in the case of all products.

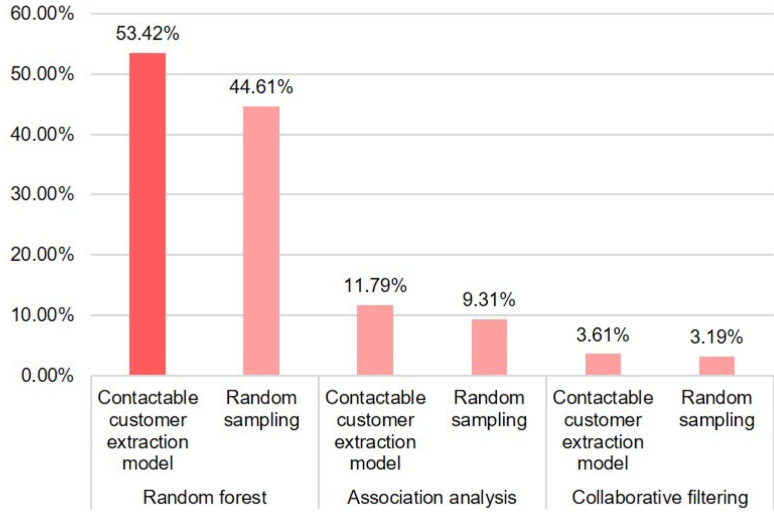
<Table 5> compares the recommended performance of the unpopular products. Unpopular products are often not recommended as first or second rank, so the third-ranked recommended products are considered at once. In the first model, only 38 out of 408 transactions (9.1%) succeeded in recommendation. In addition, the second model based on the contactable customer extraction model succeeded in recommending 62 out of 526 transactions (11.79%). Next, the results of the third and fourth combined model using the user-based collaborative filtering model were recommended 13 (3.19%) and 19 (3.61%) respectively. This is rather lower performance than association analysis.

On the other hand, the fifth combined model achieved 182 (44.61%) of the total 408 transactions and the sixth combined model achieved 281 (53.42%) of the total 526 transactions. The results of the comparison are shown in <Figure 6>, and the random forest model produced superior recommendation results compared to other models.

Thus, the screening of customers based on the contactable customer extraction model showed good results in both the overall product and the unpopular product compared to the random selection method. Therefore, it is considered that the algorithm that

<Table 5> Performance Comparison of Multi-purpose Hybrid Recommendation System of Unpopular Products

First step	Second step	Num of target	Num of success	Num of fail	Recommendation success rate
Random sampling	Association analysis	408	38	370	9.31%
	Collaborative filtering	408	13	395	3.19%
	Random forest	408	182	226	44.61%
Contactable customer extraction model	Association analysis	526	62	464	11.79%
	Collaborative filtering	526	19	507	3.61%
	Random forest	526	281	245	53.42%



<Figure 6> Recommendation Success Rate Comparison of Multi-purpose Hybrid Recommendation System of Unpopular Products

selects customers with high possibility of contact before recommending the personalized product is an approach to effectively utilize recommendation system in telemarketing. In addition, the random forest model using various contents information was selected as the final algorithm of the multi-purpose hybrid recommendation system proposed in this study, with better results than association analysis or collaborative filtering in all cases.

Ⅴ. Conclusions and Future Work

5.1. Conclusions

The proposed recommendation system in this study solved the cold start and scarcity problem of existing recommendation algorithm by using contents information such as customer master information and telemarketing history. In addition, we predicted the possibility of contact, rather than randomly selecting customers for telemarketing purposes to improve business performance, suggesting the use of an approach to select suitable customers for telemarketing. Also, in conducting the research on the Recommendation System, it is practical to improve the Telemarketing Performance by utilizing artificial intelligence, in addition to the traditional statistical method. Of course, machine learning is a kind of black-box model that is insufficient to explain causality, but it is more predictable and descriptive than any other predictive method. Using this technique, together with other statistical techniques that can explain causality, will greatly help to produce marketing results. As a result, the proposed methodology was able to select customers that were effectively valid compared to the existing random sampling method. In addition, the random forest model used as a recommendation algorithm produced better prediction results than the association analysis or collaborative filtering based recommendation system, and the multi-purpose hybrid recommendation system of this study showed that the performance of telemarketing can be greatly improved. At this time, the proposed model shows excellent performance not only in terms of overall performance but also in terms of the recommendation success rate of the unpopular product, so that it can be expected to be highly utilized compared to existing recommendation algorithms. Also, this methodology complements the shortcomings of each existing individual methodology and proposes a converged approach. This is thought to contribute very much to the decision making of the practitioners in charge of the field by enhancing the accuracy and timeliness of the product recommendation algorithm.

5.2. Limitations and Future Work

This study classified hundreds of insurance products into eight highly similar categories. However, even if they are classified in the same category, they excluded that the nature of the products handled by the insurer may be different. Therefore, if sufficient data exist for each insurer, it is possible to expect a quality improvement of the recommendation system by developing a separate model for each insurer.

Also, this study confirmed only the recommendation performance of insurance products in the area of telemarketing, which may be difficult to generalize into all telemarketing areas. Therefore, in order to confirm the generalization of the recommendation system, it is necessary to check the applicability of the proposed recommendation system by using the telemarketing data of the contract service industry such as card and communication.

Finally, we can develop a new hybrid recommendation system using other machine learning algorithms such as SVM in the recommendation algorithms. In addition, the recommendation algorithm can be improved using additional data along with customer master information, point history, and telemarketing history.

Acknowledgement

This research was financially supported by Hansung University.

<References>

[1] Acilar, A. M., and Arslan, A. (2009). A collaborative filtering method based on artificial immune network. *Expert Systems with Applications, 36*(4), 8324-8332.

[2] Arsan, T., Koksal, E., and Bozkus, Z. (2016). Comparison of collaborative filtering algorithms with various similarity measures for movie recommendation. *International Journal of Computer Science, Engineering and Applications, 6*(3), 1-20.

[3] Bobadilla, J., Ortega, F., Hernando, A., and Gutierrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems, 46*, 109-132.

[4] Breese, J. S., Heckerman, D., and Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, 43-52.

[5] Brynjolfsson, E., Hu, Y., and Simester, D. (2011). Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science, 57*(8), 1373-1386.

[6] Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction, 12*(4), 331-370.

[7] Burke, R. (2007). Hybrid web recommender systems. *The Adaptive Web*, 377-408.

[8] Chen, W., Niu, Z., Zhao, X., and Li, Y. (2014). A hybrid recommendation algorithm adapted in e-learning environments. *World Wide Web, 17*(2),

271-284.

[9] Chen, Y. L., Tang, K., Shen, R. J., and Hu, Y. H. (2005). Market basket analysis in a multiple store environment. *Decision Support Systems, 40*(2), 339-354.

[10] Christakou, C., Vrettos, S., and Stafylopatis, A. (2007). A hybrid movie recommender system based on neural networks. *International Journal on Artificial Intelligence Tools, 16*(5), 771-792.

[11] Ekstrand, M. D., Riedl, J. T., and Konstan, J. A. (2010). Collaborative filtering recommender systems. *Foundations and Trends in Human-Computer Interaction, 4*(2), 81-173.

[12] Ghazarian, S., and Nematbakhsh, M. A. (2015). Enhancing memory-based collaborative filtering for group recommender systems. *Expert Systems with Applications, 42*(7), 3801-3812.

[13] Hu, R., and Pu, P. (2010). Using personality information in collaborative filtering for new users. *Recommender Systems and the Social Web*, 17-24.

[14] Hu, B., Guo, M., and Zhang, H. (2009). A hybrid music recommendation system by M-LSA. *The International Conference on Computational Intelligence and Natural Computing, 1*, 129-132.

[15] Kardan, A. A., Abbaspour, S., and Hendijanifard, F. (2009). A hybrid recommender system for e-learning environments based on concept maps and collaborative tagging. *The 4th International Conference on Virtual Learning*, 300-307

[16] Khribi, M. K., Jemni, M., and Nasraoui, O. (2007). Toward a hybrid recommender system for e-learning personalization based on web usage mining techniques and information retrieval. *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, 6136-6145.

[17] Kim, D., Park, C., Oh, J., Lee, S., and Yu, H. (2016). Convolutional matrix factorization for document context-aware recommendation. *Proceedings of the 10th ACM Conference on Recommender Systems*, 233-240.

[18] Lacerda, A. (2017). Multi-objective ranked bandits for recommender systems. *Neurocomputing, 246*, 12-24.

[19] Langseth, H., and Nielsen, T. D. (2015). Scalable learning of probabilistic latent models for collaborative filtering. *Decision Support Systems, 74*, 1-11.

[20] Linden, G., Smith, B., and York, J. (2003). Amazon.com recommendations item-to-item collaborative filtering. *IEEE Internet Computing, 7*(1), 76-80.

[21] Min, F., and Zhu, W. (2013). *Cold-start recommendation through granular association rules*. Retrieved from https://arxiv.org/abs/1305.1372/Accessed 26 September 2018.

[22] Mobasher, B., Burke, R., and Sandvig, J. (2006). Model-based collaborative filtering as a defense against profile injection attacks. *Proceedings of the 21st National Conference on Artificial Intelligence, 2*, 1388-1393.

[23] Moro, S., Cortez, P., and Rita, P. (2014). A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems, 62*, 22-31.

[24] Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G. S., Wallinga, J., Heijne, J., Sadkowska-Todys, M., Rosinska, M., and Edmunds, W. J. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS MEDICINE, 5*(3), 381-391.

[25] Ordonez, C. (2006). Association rule discovery with the train and test approach for heart disease prediction. *IEEE Transactions on Information Technology in Biomedicine, 10*(2), 334-343.

[26] Paradarami, T. K., Bastian, N. D., and Wightman J. L. (2017). A hybrid recommender system using artificial neural networks. *Expert Systems with Applications, 83*, 300-313.

[27] Reshma, R., Ambikesh, G., and Thilagam, P. S. (2016). Alleviating data sparsity and cold start in recommender systems using social behaviour. *2016 International Conference on Recent Trends in Information Technology*, 1-8.

[28] Sarwar, B. (2001). *Sparsity, scalability, and distribution in recommender systems*. PhD thesis, University of Minnesota.

[29] Sarwar, B., Karypis, G., Konstan, J., and Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. *Proceedings of the 10th International Conference on World Wide Web*, 285-295.

[30] Shaw, G., Xu, Y., and Geva, S. (2010). Using association rules to solve the cold-start problem in recommender systems. In *Proceeding of the Advances in Knowledge Discovery and Data Mining (PAKDD 2010)*, LNAI 6118, 340-347.

[31] Paranjape-Voditel, P. and Deshpande, U. (2013). A stock market portfolio recommender system based on association rule mining. *Applied Soft Computing, 13*(2), 1055-1063.

[32] Wang, H. C., Jhou, H. T., and Tsai, Y. S. (2018). Adapting topic map and social influence to the personalized hybrid recommender system. *Information Sciences*, 1-17.

[33] Wang, X., and Wang, Y. (2014). Improving content-based and hybrid music recommendation using deep learning. *Proceedings of the 22nd ACM international conference on Multimedia*, 627-636.

[34] Wang, Y. F., Chuang, Y. L., Hsu, M. H., and Keh, H. C. (2004). A personalized recommender system for the cosmetic business. *Expert Systems with Applications, 26*(3), 427-434.

[35] Yang, W. S., Cheng, H. C., and Dia, J. B. (2008). A location-aware recommender system for mobile shopping environments. *Expert Systems with Applications, 34*(1), 437-445.

◆ About the Authors ◆

Hyung Su Kim

Hyung Su Kim is currently an associate professor in the department of Industry & Management Engineering at the Hansung University and a representative of CMI(Customer Management Institute), a CRM Consulting firm in Seoul, Korea. His main research areas are Customer Management Relationship (CRM), customer analytics and modeling. He served as Chairman of Korea CRM Association and Vice President of Korean Academic Society of CRM. He has published 8 books and more than 15 academic papers on CRM and Big Data Analysis. He received a PhD from the Korea Advanced Institute of Science and Technology in Korea.

Sangwon Lee

Sangwon Lee is currently an associate professor in the department of Computer & Software Engineering at the Wonkwang University in Korea. He is also a non-resident researcher at the Korea European-Union Research Center in Germany. His research interest is data engineering including database and data analytics. His work has been published in 55 journals and presented at 84 academic conferences. He received a PhD from the Korea Advanced Institute of Science and Technology in Korea.

Submitted: May 1, 2019; 1st Revision: June 8, 2019; 2st Revision: July 15, 2019; Accepted: July 19, 2019