Product Recommendation in Offline Retail Industry by using Collaborative Filtering

Bayu Yudha Pratama[1], Indra Budi[2]*, Arlisa Yuliawati[3]
Faculty of Computer Science
Universitas Indonesia, Depok
West Java, Indonesia

Abstract—**The variety of purchased products is important for retailers. When a customer buys a specific product in a large number, the customer might get benefit, such as more discounts. On contrary, this could harm the retailers since only some products are sold quickly. Due to this problem, big retailers try to entice customers to buy many variations of products. For an offline retailer, promoting specific products based on the markets' taste is quite challenging because of the unavailability of information regarding customers' preferences. This study utilized four years of purchase transaction data to implicitly find customers' ratings or feedback towards specific products they have purchased. This study employed two Collaborative Filtering methods in generating product recommendations for customers and find the best method. The result shows that the Memory-based approach (k-NN Algorithm) outperformed the Model-based (SVD Matrix Factorization). Another finding is that the more data training being used, the better the performance of the recommendation system will result. To cope with the data scalability issue, customer segmentation through k-Means Clustering was applied. The result implies that this is not necessary since it failed to boost up the models' accuracy. The result of the recommendation system is then applied in a suggested business process for a specific offline retailer shop.**

Keywords—Recommendation system; offline retail store; memory-based collaborative filtering; customer segmentation

I. INTRODUCTION

Recommendation system is a collection of tools and techniques to provide products or services suggestions for users [1]. The existence of this system allows companies to develop a marketing strategy, attract more customers, and increase sales. Therefore, many companies try to implement a recommendation system for their business interest. Recommendation system has been applied in a variety of industries. It can be found in the entertainment domain (music, movies, TV shows, books), news or tourism sites, e-commerce, e-library, and e-learning systems [2].

Even though the recommendation system has been extensively used in e-commerce domain as described in [2] and [3], research in [4] argued that it can also be implemented in the traditional retail stores. They suggested personalization as the next possible strategy for this kind of retailers. Personalization establishes a one-to-one relationship between the retailer and the customer. By using a one-to-one relationship, a retailer can remember details and preferences for each customer. These preferences can be utilized to identify customer personal needs, wants, and demands. This personalization strategy can be realized by the implementation of recommendation system. E-commerce has already implemented the recommendation system with many benefits such as boosting up customer level of interaction, increasing sales, the diversity of items sold, customer satisfaction or loyalty, and also understanding customers' demand better [5]. Such benefits are expected to be achieved in traditional or offline retail stores.

A traditional or offline retail store differs from e-commerce in several aspects. The first is that traditional retail store still having a physical store for storage, display, and transaction. This is costly for them to keep rarely sold items in inventory [4]. On the other hand, the diversity of customers' demands always increasing. Retailers must be able to correctly identify customers' demand as well as offering a variety of products. This is a way to keep the goods in inventory to keep moving. It differs from the e-commerce setting on which they do not always have physical storage to keep their products.

The other distinction between offline and online commerce is that in an offline retail store, it is often difficult to access customers' purchase history and observe their purchase behavior. In some cases, the offline stores only keep the transactions without knowing who the buyers are. Of course, this is different from the online based stores, which the customers' identities and activities are recorded in the system. From this situation, there is a problem related to the unavailability of the user related data. Moreover, it is difficult for all customers to give feedback towards items they have bought. The impact is that the predictions are often poor when other users or customers are looking for recommendations for the rarely rated items [4]. This is different from e-commerce business where users usually explicitly asked to give a rating when their transaction is completed. The availability of such feedbacks or ratings is crucial for recommendation system studies since customers' feedback or ratings becomes one of available sources to derive any information that may be useful for other customers [6].
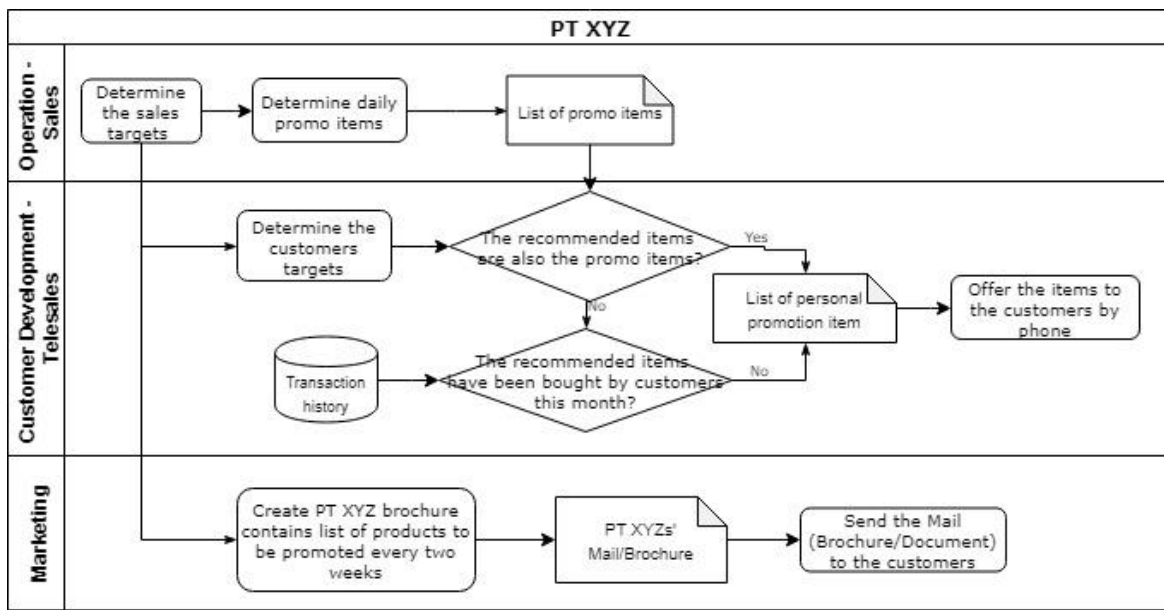
---

*Corresponding Author

Fig. 1. The Business Process of PT XYZ in Promoting their Products.

This study is conducted based on the situation in a wholesales retail store in Indonesia, called PT XYZ. This retailer is categorized as a Broad-Deep-Mix retail store, which has large variety of products and heterogeneous customers [4]. This retailer has a personalized marketing scheme, by offering specific items either by phone or by sending product brochure to specific customers as presented in Fig. 1. Based on previous explanation, there are problems related to personalization in offline retail store. Inaccurate promotion target, either the items to be offered or the targeted customers can lead to the failure in selling those items. Hence, it will fail to fulfill their purpose to sell various products to customers, as well as fail to improve their profit. Therefore, the aim of this paper is to find a suitable approach to give better recommendations for customers on an offline retailer (specifically for PT XYZ) both empirically and practically, as well as considering the personalization approach that has already been used.

Following section will describe the related works and summarized proposed works about recommendation system in offline retailer. It continued to the next two sections with the materials and the research methodology. The result and discussion are presented in following section, and at the end of this paper the conclusion of this study will be presented.

II. LITERATURE REVIEW

A. The Problem of Limited Customers' Data

Specific for offline retailer, the main issue of this study is related to the absence of explicit product ratings or feedback by customers. As the main input for the process of the recommendation system, it is important to define such things. A research in [7] summarizes about some experiments to deal with this problem. Many of them use the association rule that ended up with the lack of personalization. Study by [7] itself utilized the smart fitting room, i.e. the IT artifact that gives product recommendation to the customers through a screen stored in the individual cabin. Started from the use of Association Rule Mining, their study shows that combine the

information from customers' interaction to the screen with the contextual information about products could improve the product recommendation in fashion stores. Another study about fashion retailer as written in [8] combines the online product click data and offline product sale data to reflect the preference of the customers. This experiment concludes that it is better to substitute than complement the products in the recommendation system. The percentage of purchase by using the former approach is higher than the latter. Nevertheless, there is less information related to the use of online and offline data combination. Recalling the situation of PT XYZ, those used in these previous studies did not owned by PT XYZ. But the idea is that these two retailers utilized the availability system that interacts with their customers, e.g. the smart fitting room and the online system to combine with the offline product sale. In PT XYZ, there is a membership system. Customers who register for membership in this retail store have their transaction history recorded in the stores' database. This study utilizes this data to generate the customers' feedback.

B. Techniques in Recommendation System Study

There are several techniques to identify which items are recommended for specific users as summarized in [1]. They are distinguished based on the domain, knowledge, or the algorithm being used.

1) Collaborative Filtering (CF): This approach gives a recommendation of items based on the similar preferences of other users in the past. The similarity between users/items can be inferred from their previous behavior such as rating or buying history.

2) Content-based Filtering (CBF): This approach recommends items that are similar to what the user liked in the past. The similarity of items is obtained from their feature and description. The recommendation is compiled from the attribute information of items.

3) Demographic recommender system: The recommendation is given based on the demographic information, such as location, language, and age of user or customers. This approach implies that people with different demographic background should not receive the same recommendation.

4) Knowledge-based recommender system: By using this approach, system gives a recommendation based on specific knowledge over items from experts. It is then matched up to the items' benefit for users. The similarity is implied from the user needs and items' function. This approach identifies the similarity based on match "answer/solution" to users' "question/problem description".

5) Community-based recommender system: The recommendation is implied from the preference of users' circle friend. This approach is popular on social network-based system.

Based on the previously mentioned techniques, there are some drawbacks on some techniques to be applied to an offline retailer, specifically for the case of PT XYZ. CBF approach is not applicable for this case since traditional retailers usually do not store comprehensive description about their products rather than only consists of name, price, main categories, and sub-categories (e.g. dry-food/fresh-food/non-food). There is also limited information about customers' profile except for members of the stores (if any), so approach based on demographic is also not suitable for this case. Similarly, offline stores also do not maintain how their customers connected each other. By this condition it is difficult to get the recommendation based on customers' circle or community. Lastly, offline stores usually provide various kinds of items, so find experts for various kinds or categories of items is another problem for employing knowledge-based approach. From this analysis, it implies that CF is the most suitable approach for the case of PT XYZ.

C. Data Scalability vs Customer Segmentation

Another issue in the recommendation system in general is the data scalability. It is caused by the huge amount of data that leads to the accuracy problem of the recommendation system [9], [10]. In some cases, this issue is related to the algorithm or approach being used to build the system, i.e. the use of Collaborative Filtering. Its performance on scalability is still poor given a huge user and item base [11]. The previous study in [9], [10], [12], as well as [13], try to include the customer segmentation process to cope with this challenge. This process is also used to identify profitable customers [9] [10], not only to make the data become smaller.

Commonly used approach to differentiate the customers into several segments is the simple-yet-powerful RFM model (Recency, Frequency, and Monetary). Recency is defined as the last time (in a month) a customer completed a transaction, Frequency describes how many total transactions for each customer, and Monetary calculates how much they buy in value [13]. Some examples of algorithms that can be used for segmenting customers are Artificial Neural Network (ANN) [10], *k*-Means clustering [9], Expectation-Maximization (EM) [13]. Both studies by [9] and [13] use this RFM model as the

segmentation method and shows a satisfiable result. The difference is the former use Association Rules and hybrid method, while the later use *k*-NN. The performance of former study was affected by using hybrid method while the later by customer segmentation. Nevertheless, the study by [9] is based on homogeneous store, i.e. only sell one kind of item. It is different from the case used in [13], as well as PT XYZ that are selling various items (heterogeneous retailer).

D. Proposed Work

Based on the previous analysis in offline retailer about the problem of limited customers' data, various techniques to use, and the challenge in the data scalability, this study propose some steps to be applied in the case of PT XYZ as a wholesale offline retailer. The first, to cope with the unavailability of the customers' ratings data, this study takes advantage of the membership system applied in PT XYZ. Basically, customers' activities recorded in the system is elaborated, and the rating data is generated implicitly from the customers' purchase pattern based on a specific transformation metric. An assumption is made related to the result of this transformation: "The more frequently customers purchase an item, the higher the rating they implicitly give". The second, Collaborative Filtering will be used in this study. Since there are two approach in this technique, this study also tries to find the best approach. The last, to cope with the data scalability, this study adapt the use of RFM model to apply the customer segmentation. An experiment is employed to elaborate whether this approach also gives better performance compared to the original process without segmenting the customers.

III. COLLABORATIVE FILTERING

CF is the most successful and widely used recommendation technique [14], [15]. CF utilizes a user-item matrix to make the recommendations. Suppose there is a set of m users $U = \{u_1, u_2, ... u_m\}$ and a set of n items $I = \{i_1, i_2, ... i_n\}$, CF constructs an $m \times n$ matrix R representing the preference of users to items. For each user, the list of relevance items can be viewed from the descending order of matrix values related to the user. If there are two users giving the same rating to an item, then it can be implied that they have the same taste or preference. As an example, from Table I, the relevance item for User_C is Item_3 and Item_1, while another information is that User_A and User_C have the same taste or preference toward Item_3. There are two approaches that are commonly used in CF, they are Memory-based and Model-based [16]. A study by [17] specifically compared these two approaches in the e-commerce domain. The result indicates that the Model-based is better than Memory-based not only in the accuracy and the relevancies of the recommendation but also in the computational time. Following subsections will explain more about these approaches.

TABLE I.   AN ILLUSTRATION OF USERS PREFERENCE TOWARD ITEMS AS INPUT FOR USER-ITEM MATRIX

	Item_1	Item_2	Item_3	Item_4
User_A	5	-	4	-
User_B	-	3	-	1
User_C	2	-	4	-

A. Memory-based Collaborative Filtering

This is also called Neighborhood-based CF. This is the most popular method in the recommendation system domain [15]. It generally shows that similar users have similar rating behavior, so do with similar items, they receive similar ratings [16]. The similarity can be defined among users (User-based) or items (Item-based). The distinction is that in the former case, the ratings are predicted using those of neighboring users, while in the latter case, they are predicted using the users' own ratings on neighboring or closely related items [16]. In this study, the User-based CF is chosen instead of Item-based CF since based on the experiment the former is better in term of accuracy, time, and space complexity. This might be caused by the number of items that is greater than the number of users in the case of PT XYZ. It is also known that User-based CF is one of the most widely used among CF approaches [18].

This approach uses k-Nearest Neighbor (k-NN) algorithm to find the top k similar users and predict the rating for specific items that have been bought by those k users. The similarity between users can be calculated by using distance metrics, such as Cosine Similarity or Pearson Correlation. The recommendation then will be given based on the rating calculation of items obtained from each of the k users. Generally, Memory-based or Neighborhood-based is a simple and straightforward approach yet still have an accurate prediction. Nevertheless, it has some limitation such as a scalability issue in a large matrix and the cold-start problem where the model cannot recommend a new user/item [10], [12].

B. Model-based Collaborative Filtering

Basically, this model tries to find the hidden factors in the original/initial matrix [1]. To create the prediction model, it applies various data mining technique, such as the Decision Tree, Bayesian, and Latent Factor model. This approach is better than the Memory-based approach in several ways [19]. For example, it has better scalability since it has a good performance in a large matrix as well as the better accuracy, i.e by using Singular Value Decomposition (SVD) as the Latent Factor model [20]. SVD runs by decomposing an initial n users × m products matrix into three matrices: U, λ, and V as illustrated in Fig. 2. These three matrices are updated continuously until the result of their multiplication is approaching the initial matrix. In Model-based CF approach, these matrices are equivalent to users U(n users × r concepts), concepts λ (r concepts × r concepts), and products V (r concepts × m products) matrices. One of the important parameters in SVD is the number of the Latent Factor, which refers to the number of concepts that are hidden in the initial matrix.
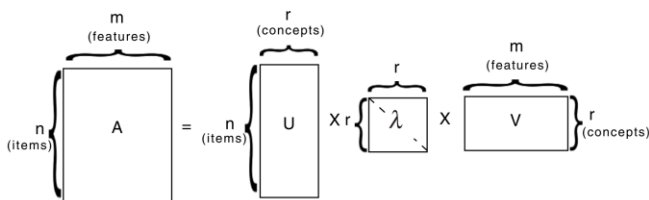

Fig. 2. Illustration of SVD [1].

IV. RESEARCH METHODOLOGY

Fig. 3 shows the research methodology of this study. Data sources are provided by PT XYZ as one of the largest retail chains in Indonesia. PT XYZ has a unique business model, it combines retail and wholesale sales. They serve both professional customers and end-user/individual consumers. The data set contains 2.5 years of transactions data from one of their branches. This research is mainly divided into two experiments. The first experiment examines the performance of both Memory-based and Model-based CF. The former approach employs k-NN algorithm and the later uses SVD matrix factorization. This study also observes the optimal value of k and the number of the latent factor for k-NN algorithm and SVD matrix factorization respectively. The second experiment focuses on the application of customer segmentation. k-Means algorithm is then used to differentiate customer into several categories based on the RFM model. The result of both experiments is compared based on the Root-Mean-Square-Error (RMSE) metric, as it is a stable means of comparison between models [21]. The better approach based on the minimum RMSE value is then applied for the recommendation system. Following subsections explain each process in this research.

A. Data Collecting and Clean Up

The data was taken in one of PT XYZ branch store. This data consists of information about users, products, and transactions from professional customers, including smaller retailer, hotel, restaurant, or catering, who buy products for their business necessities. This study only use the transactions from professional customers and omitting those from end-users because PT XYZ wants to test personalization to their professional customers first, before moving on to end-user later. Besides, PT XYZ already has a direct relation to professional customers by offering products via telephone. The expectation is to expand the relationship into one-to-one personalized marketing.
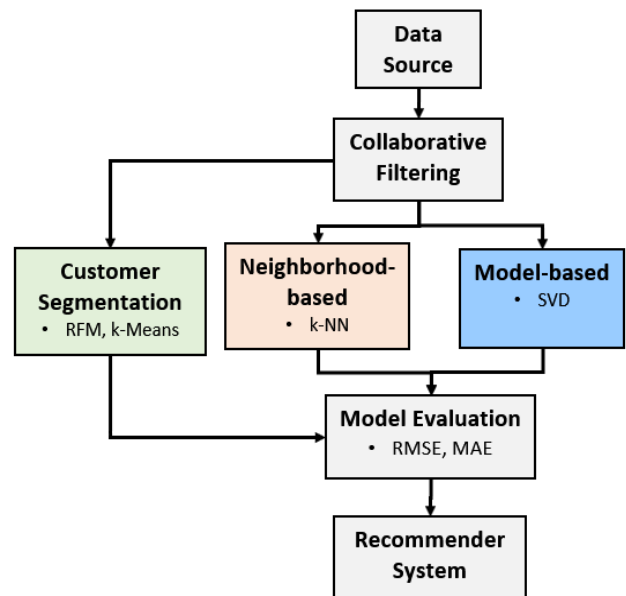

Fig. 3. Research Methodology.

The information about the user consists of customer number (cust_no) and customer name (cust_name). Product data also consists of product code (prod_cd) and product name (prod_name). While the transaction data is the purchase of an item (prod_cd) by a customer (cust_no) in a specific time (sale_day). The data collected from 2.5 years of transactions from one store within time windows from January 2015 to September 2017. There are numbers of transactions that are not relevant for this study such as the internal purchase or transactions involving non-trade items such as insurance, administration fee, and shipping charges. Clean-up process is carried out by removing this type of transactions. From this data collection, a total of 8,515 customers, 23,532 items, and over 2,1 million transaction records are obtained.

B. Data Transformation

The Collaborative Filtering method utilizes the user-item rating matrix for making predictions. This matrix describes user u giving item i a rating value r_{ui}. Since this kind of data is not available in the case of a traditional retail store, the customers' purchase history is utilized such that it implicitly represents customers' feedback towards what they have bought. This approach is suitable for the characteristics of professional customers. Basically, they tend to continuously buy products for their day-to-day business operation. If they repeatedly buy a specific item, it is an implicit indication that they like that item because it shows that the more frequently they purchase an item, the higher the rating they give.

From the provided data, a customer did a transaction of a specific product in a specific time. From this data the number of purchases by customer for each item is extracted by using aggregate function. Table 2 presents an example of some complete transaction data, while Table 3 show the result of the aggregate process. The quantity and value of a transaction to create a level playing field between transactions in big and small companies are ignored. This is because the big company usually has bigger transaction value than the small company. From this process, one purchase of an item is considered as one transaction regardless to its total value of purchase.

The user-item matrix come from a purchase frequency matrix consists of user u buy item i for f_{ui} times. This information is converted into a rating value by using a min-max scaling algorithm as shown in Equation (1). The variable f_{ui} represents the frequency of purchase while $f_{\min}$ and f_{max} represent the minimum and maximum purchase respectively for each item. The variable r_{ui} represent the transformed rating value ranged from 1 ($r_{\min}$) to 5 (r_{max}).

$$r_{ui} = \left[\left(\left(\frac{f_{ui} - f_{min}}{f_{max} - f_{min}}\right) \times (r_{max} - r_{min})\right) + r_{min}\right] \quad (1)$$

Table 4 shows the example of original data, consist of the frequency of purchasing an item by a customer. The definition of the value of $f_{\min}$ and f_{max} are based on the information of each column that represent each item instead of considering the minimum and maximum frequency of the whole matrix. Therefore, the sale rate of items can be analyzed whether they are fast-moving items or slow-moving items. Then Table 5 shows the transformation result, which is the implicit rating value of an item given by a user.

In some cases, the purchase frequency of specific items is too high, far from the purchase frequency on average case. This can result in the skewness of the rating value gather in a lower value. To overcome this condition, the data with too high purchase frequency are removed. The removal process is conducted by modelling the purchase distribution of frequency of each item as a normal distribution. When the purchase frequency (f_{ui} value) exceeding the normal distribution limit resulted from Equation (2), then this outlier is removed. The limit (T_i value) is obtained from the mean value of purchase frequency of item i (μ_i) added by 3 times the standard deviation (σ_i) value of item is' purchase frequency.

$$T_i = \mu_i + (3 \times \sigma_i) \quad (2)$$

C. Model Development

In this study, two collaborative-filtering approaches is compared, they are Memory-based by using k-NN and Model-based through SVD matrix factorization. These models have some adjustable parameters to obtain optimal performance as described below. The model for both k-NN and SVD is developed by using the optimal parameter values found in this experiment and evaluate their performance.

TABLE II.   AN EXAMPLE OF COMPLETE TRANSACTIONS

cust_no	prod_cd	freq	time
6001000580465	1029504000	1	2015Q1
6001000580465	1029504000	1	2015Q1
6090001779264	1020248000	1	2016Q2
6090001779264	1020248000	1	2016Q3
6090001779264	1020248000	1	2016Q4
6090001779264	0032080000	1	2017Q2

TABLE III.   THE RESULT OF THE AGGREGATE PROCESS

cust_no	prod_cd	freq
6001000580465	1029504000	2
6090001779264	1020248000	3
6090001779264	0032080000	1

TABLE IV.   AN EXAMPLE OF USER-ITEM PURCHASE FREQUENCY MATRIX

	Item_1	Item_2	Item_3	Item_4	Item_5
User_A	2	-	-	7	-
User_B	-	1	-	-	5
User_C	8	-	3	2	1
User_D	-	20	-	3	-
User_E	10	-	-	-	-

TABLE V.   THE DATA TRANSFORMATION RESULT: USER-ITEM RATING MATRIX

	Item_1	Item_2	Item_3	Item_4	Item_5
User_A	1	0	0	5	0
User_B	0	1	0	0	5
User_C	4	0	5	1	1
User_D	0	5	0	2	0
User_E	5	0	0	0	0

Evaluation for all experiments are based on the Root-Mean-Square-Error (RMSE) error metric, as shown in Equation (3). In this equation, $\hat{R}$ is the collection of rating prediction, while r_{ui} is the actual rating in testing data set and $\hat{r}_{ui}$ is the predicted rating from the model.

$$RMSE = \sqrt{\frac{1}{|\hat{R}|} \sum_{\hat{r}_{ui} \in \hat{R}} (r_{ui} - \hat{r}_{ui})^2} \qquad (3)$$

1) In general, k-NN algorithm finds k most similar users $(N_i^k(u))$ based on the previous buying or rating pattern. The similarity between two users $(sim(u,v))$ is calculated through *cosine similarity* as shown in Equation (4) and the rating prediction ($\hat{r}_{ui}$) is computed by using Equation (5). The variable I_{uv} represents the item that is rated by user u and v, r_{ui} is the rating prediction value from user u for item i, and r_{vi} is the actual rating value from user v for item i. One of the important parameters for k-NN is the number of nearest/similar users (the value of k). Fig 4 presents the result of an experiment to find the optimal value of k. The experiment is performed with training data collected from 2015 Q1 to 2017 Q2 and use the testing data from transactions in 2017 Q3. RMSE metric is used to compare the result. k-NN delivers best result when the value of k is greater or equals to 80.

$$sim(u,v) = \frac{\sum_{i \in I_{uv}} r_{ui} \cdot r_{vi}}{\sqrt{\sum_{i \in I_{uv}} r_{ui}^2} \cdot \sqrt{\sum_{i \in I_{uv}} r_{vi}^2}} \qquad (4)$$

$$\hat{r}_{ui} = \frac{\sum_{v \in N_i^k(u)} sim(u,v) \cdot r_{vi}}{\sum_{v \in N_i^k(u)} sim(u,v)} \qquad (5)$$

2) SVD decompose original user-item matrix into user, concept, and item matrices. As explained in [1], the decomposition or the factorization process maps the users and items into latent factors space. This latent space explains ratings by characterizing both items and users on factors that are inferred from users' feedback. Rating prediction that describes the overall interest of the user in characteristics of the item is computed within Equation (6). Each item i is associated with a vector q_i that measure the extent of item i possesses those factors. While each user u is associated with a vector p_u that measure the interest of user u towards factors. This study perform an experiment with SVD to evaluate the optimal number of factor, and the result is shown in Fig 5. The best result is achieved by using factors of less than or equals to 10. The result shows that SVD is able to find few factors in the original matrix, presumably because of the heterogeneous nature of items offered.

$$\hat{r}_{ui} = q_i^T p_u \qquad (6)$$

After obtaining the optimal parameter values, then evaluation towards the use of Model-based CF by using k-NN and Memory-based CF through SVD matrix factorization in product recommendation is conducted. Each model is built based on those previously found parameter values, i.e. k = 80 for the implementation of k-NN algorithm and the SVD matrix factorization is trained by using the number of factors = 10.

The data set contains time-series data and divided into quarter year data. There are eleven quarters from 2015 Q1 until 2017 Q3. The transaction data from 2015 Q1 to 2017 Q2 are used as the training data set, while the transaction data in 2017 Q3 are used as the testing data set. The variant of each training set is made by removing data from the oldest quarter. This approach is used to determine the amount of data and the extent to which optimal training data points required. Is it sufficient to train only the newest data from the last quarter of the year? Or is it necessary to train as much data as possible? The models process the user-item rating matrix and produce rating prediction for every pair (item u, user i) in the testing data set that is 2017 Q3 transactions (105,177 records). The rating prediction from each model is then compared to the factual sales data in this testing data set.

The better CF approach will be used to apply customer segmentation based on customer lifetime value using the RFM model. This to answer whether customer segmentation has a positive impact on the development of the product recommendation system. The segmentation process employs k-Means clustering algorithm with $k = 3$. The model is then tested by using the training data starting from 2015 Q1.
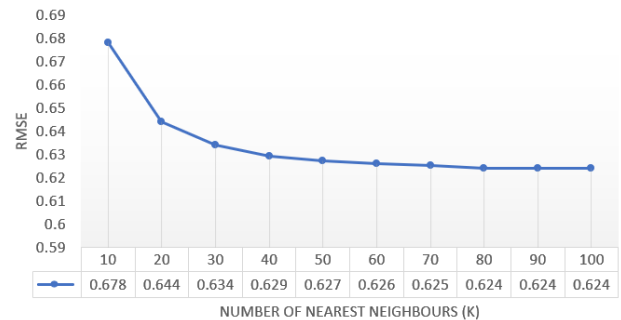
	10	20	30	40	50	60	70	80	90	100
	0.678	0.644	0.634	0.629	0.627	0.626	0.625	0.624	0.624	0.624

NUMBER OF NEAREST NEIGHBOURS (K)

Fig. 4. The Experiment Result of Finding the Optimal Value of k.

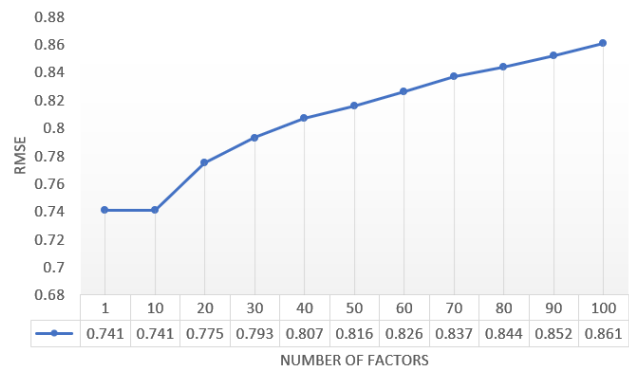|	1	10	20	30	40	50	60	70	80	90	100
	0.741	0.741	0.775	0.793	0.807	0.816	0.826	0.837	0.844	0.852	0.861

NUMBER OF FACTORS

Fig. 5. The Experiment Result of Finding Optimal Value of the Latent Factors.

V. RESULT AND DISCUSSION

A. Memory-based vs Model-based Collaborative Filtering

The experiment in comparing the performance of using Memory-based by using k-NN and Model-based through SVD matrix factorization generally shows that the former is slightly outperforms the latter. As presented in Table 6, by using k-NN a better prediction is obtained compared to the use of SVD matrix factorization. Among the ten variants of data training, the error value by using k-NN is always below the error value of SVD, as long as the number of training data being used come from more than two quarters. This result is caused by the size of the matrix that is not exceptionally high and is still sufficient to be performed in the Memory-based approach.

Meanwhile, Table 6 also shows that the smallest value of RMSE for both approaches is obtained when all of the ten quarters training data are utilized. By comparing the result from each training data set, as the number of training data is decreased, the RMSE values are increasing consistently. From this point, it implies that the older the starting point of the training data or the more data available for training process, the less error value will be obtained, so the prediction result would be better.

B. Customer Segmentation

This experiment tries to differentiate the customers into several segments before applying the better model, which is Memory-based by using k-NN. Table 7 presents the characteristics of each segment and the result of this experiment. It implies that Recency is the most important variable since most of the transactions come from customers who have low Recency or recently purchased items. These customers tend to have high Frequency and Monetary value because they are buying products frequently (repeat buying). This customer segment is defined as "active". On contrary, customers who have high Recency are rarely done transactions. This is based on their level of Frequency and Monetary that are ranged from medium to low. They called as "inactive" customers, and the remaining cluster between the "active" and "inactive" ones are called "semi-active" customers.

The other finding is that based on the RMSE value, the use of training data that only come from "active" customer results in a slightly better performance of the recommendation system compared to those from both "semi-active" or "inactive" customer segments. While the use of training data that come from both of these last two segments, yield the significantly decreasing performance compared to both of the use of only "active" customer or without clustering process.

Although the data from "active" customer yield good performance, the error value is still higher than the excluding of the segmentation process. This result implies that segmenting customers into several segments is failed to improve the performance of the models. One factor that causes this is the size of the data training. By dividing the data based on a specific customer, each cluster of training data has a fewer amount of data than the combined data. This result strengthens the previous verdict from the first experiment that the more data training being used, the better model will result.

C. Implementations

Experiment in this study basically shows that it is better to use Memory-based CF with the use of as many as training data. The recommendation provided by this model consists of the list of relevant products related to specific customers. This list is generated by sorting the prediction rating value obtained from the model. From this list of recommendations, the additional steps are inserted in the promotion flow of PT XYZ as the offline retailer so this would help them to give better promotion result as well as increasing their profit.

Fig. 6 shows the suggested business model for PT XYZ. The sales division promote their products by phone but based on the recommended products obtained from the output of the recommendation system. If these products are included in the promo items, then they can be immediately offered to the customers. The combination of recommended products with specific discounts will attract customers' intention to buy them. This can improve the success rate of the promotion. If customers have already purchased a specific item on the recommendation list, the sales team should skip or remove this item then move to the next item on the list. This approach can improve the variety of items to be purchased by the customers.

Another advantage of the recommendation systems' output is for the marketing division. The actions are similar to the previous promotion by phone. Furthermore, they can utilize PT XYZs' personal mail to create personalization towards each customer. The content of the mail or brochure consists of relevant products for the specific customer. In the normal pipeline, they send this mail to the customer once in two weeks with the same content in one segment. Now they can adjust the content and the frequency of the mailing. This approach will decrease the number of human resources needed to make promotions by phone, in case there are only a few numbers of resource available. Sending this mail to all customers also effective to improve the sale rate.

TABLE VI.   MODEL EVALUATION RESULT BASED ON RMSE VALUE

Training Data	Number of Quarter	RMSE	
		k-NN	SVD
2015 Q1 – 2017 Q2	10	0.624	0.750
2015 Q2 – 2017 Q2	9	0.639	0.753
2015 Q3 – 2017 Q2	8	0.653	0.762
2015 Q4 – 2017 Q2	7	0.673	0.770
2016 Q1 – 2017 Q2	6	0.699	0.788
2016 Q2 – 2017 Q2	5	0.727	0.802
2016 Q3 – 2017 Q2	4	0.762	0.824
2016 Q4 – 2017 Q2	3	0.818	0.850
2017 Q1 – 2017 Q2	2	0.890	0.887
2017 Q2 – 2017 Q2	1	0.974	0.926

TABLE VII.   THE RESULT OF CUSTOMER SEGMENTATION EXPERIMENT

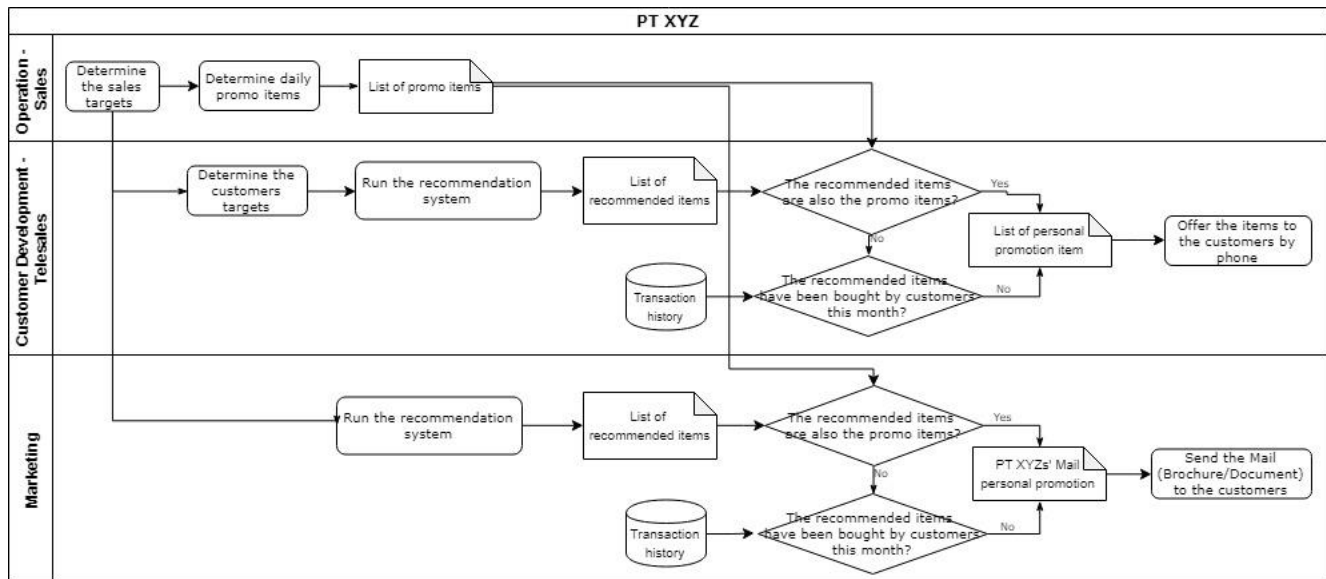| Cluster	Recency	Frequency	Monetary	Number of customers	% of transactions	RMSE
1 – "active"	Low	High, Medium	High, Medium	3,805	83%	0.636
2 – "semi-active"	Medium	Medium, Low	Medium, Low	2,513	12%	1.454
3 – "inactive"	High	Medium, Low	Medium, Low	2,198	5%	1.515
No cluster	-	-	-	8,515	100%	0.625

Fig. 6.   Suggested Pipeline for the Promotion Process in PT XYZ.

VI. CONCLUSION

The goal of this research is to develop a better approach to be implemented in a recommendation system for a traditional or offline retail store. Before suggesting the new pipeline for the promotion process in PT XYZ, some experiments are conducted. The first, related to one of the problems in an offline retail store is the unavailability of customers' rating/feedback data towards products. This study deals with this problem by constructing a user-item matrix based on the number of purchases by the user as an implicit feedback score. The higher number of purchases implies the higher feedback or rating being given by a customer. Nevertheless, this approach only covers the data from customers who join the membership of PT XYZ.

The second, the widely used Collaborative Filtering approach is applied in this study. An experiment to find a better approach by using Memory-based and Model-based Collaborative Filtering is conducted to predict the rating given by the target customer. The result shows that Memory-based CF with k-NN outperforms the Model-based CF through SVD. Regarding the amount of training data, it can be concluded that more data training is always resulting in a better prediction. This result is concluded by conducting an experiment to find the optimal value of the parameters being used in k-NN algorithm and SVD matrix factorization. In k-NN model. The neighborhood size (the value of k) is directly proportional with the model accuracy. The higher the neighborhood size, the better accuracy is reached. On contrary, the number of factors that are used in SVD matrix factorization is inversely proportional, where best performance is found on the use of the fewest number of factors.

The last, related to the data scalability problem, some previous studies to cope with this problem by creating a smaller size of data is adapted. One of the approaches in previous studies is using customer segmentation based on specific criteria. Their studies show that it was better to add this segmentation process. This study applied this approach to differentiate the training data based on RFM model before running the recommendation system model, and the result implies that in the term of accuracy, it is not necessary to do this since the performance of the model is failed to be improved. Related to the case being used in this study, the more training data is needed to build a better model. While by segmenting the customers into three categories (active, semi-active, inactive) reducing the size of the training data.

As the contribution for the promotion process in PT XYZ, the best model to create a personalized list of recommended products for each targeted customer is utilized. The recommendation list can be generated by sorting the product by rating prediction. This list then can be integrated into current sales and marketing strategy, for example: creating promotional products, clearing stock or cross-selling. By offering the relevant product, a retailer can reap the benefit of the recommendation system such as personalized marketing, improving customer loyalty, and increasing cross-selling. The

main difference from previous pipeline is the list of recommended products has been adjusted to the target customers. It is more useful for customers since they are likely to receive recommendation of products they need.

REFERENCES

[1] F. Ricci, L. Rokach and B. Shapira, Recommender System Handbook, Springer, 2011.

[2] J. Lu, D. Wu, M. Mao, W. Wang and G. Zhang, "Recommender system application developments: A survey," Decision Support Systems, vol. 74, pp. 12-32, 2015.

[3] S. Sivapalan, A. Sadeghian, H. Rahnama and A. M. Madni, "Recommender Systems in E-Commerce," in World Automation Congress, Waikoloa, HI, USA, 2014.

[4] F. E. Walter, S. Battiston, M. Yildirim and F. Schweitzer, "Moving recommender systems from on-line commerce to retail stores," Information System and Business Management, vol. 10, no. 3, pp. 367-393, 2012.

[5] M. Kang, D.-H. Shin and T. Gong, "The role of personalization, engagement, and trust in online communities," Information Technology and People, vol. 29, no. 3, pp. 580-596, 2015.

[6] T. Donkers, B. Loepp and J. Ziegler, "Explaining Recommendations by Means of User Reviews," in Workshop on Explainable Smart Systems (ExSS) 2018, 2018.

[7] J. Hanke, M. Hauser, A. Dürr and F. Thiesse, "Redefining the Offline Retail Experience: Designing Product Recommendation Systems for Fashion Stores," in 26th European Conference on Information Systems, Portsmouth, United Kingdom, 2018.

[8] H. Hwangbo, Y. S. Kim and K. J. Cha, "Recommendation system development for fashion retail e-commerce," Electronic Commerce Research and Applications, vol. 28, pp. 94-101, 2018.

[9] F. Rodrigues and B. Ferreira, "Product Recommendation based on Shared Customer's Behaviour," Procedia Computer Science, vol. 100, pp. 136-146, 2016.

[10] Z. Shi, Z. Wen and J. Xia, "An Intelligent Recommendation System based on Customer Segmentation," International Journal of Research in Business Studies and Management, vol. 2, no. 11, pp. 78-90, 2015.

[11] M. Singh, "Scalability and sparsity issues in recommender datasets: a survey," Knowledge and Information Systems, pp. 1-43, 2018.

[12] G. K. Kishore and D. S. Babu, "Recommender System based on Customer Behaviour for Retail Stores," IOSR Journal of Computer Engineering, vol. 19, no. 3, pp. 06-17, 2017.

[13] S. M. Rezaenia and R. Rahmani, "Recommender system based on customer segmentation (RSCS)," vol. 45, no. 6, pp. 946-961, 2016.

[14] Y. Shi, M. Larson and A. Hanjalic, "Collaborative filtering beyond the user-item matrix: A survey of the state of the art and future challenges," ACM Computing Surveys, vol. 47, no. 1, 2014.

[15] J. Bobadilla, F. Ortega, A. Hernando and A. Gutiérrez, "Recommender systems survey," Knowledge-Based Systems, vol. 46, pp. 109-132, 2013.

[16] C. C. Aggarwal, An Introduction to Recommender Systems. In Recommender System, Springer, 2016.

[17] P. H. Aditya, I. Budi and Q. Munajat, "A Comparative Analysis of Memory-based and Model-based Collaborative Filtering on the Implementation of Recommender System for E-commerce in Indonesia : A Case Study PT X," in International Conference on Advanced Computer Science and Information System, Malang, Indonesia, 2016.

[18] J. Park and K. Nam, "Group recommender system for store product placement," Data Mining and Knowledge Discovery, vol. 33, pp. 204-229, 2019.

[19] S. G. Moghaddam and A. Selamat, "A Scalable Collaborative Recommender Algorithm based on User Density-Based Clustering," in The 3rd International Conference on Data Mining and Intelligent Information Technology Applications, Macao, China, 2011.

[20] S. Chan, P. Treleaven and L. Capra, "Continuous hyperparameter optimization for large-scale recommender systems," in Internationa Conference on Big Data, Silicon Valley, CA, USA, 2013.

[21] P. Cremonesi, F. Garzotto, S. Negro, A. V. Papadopoulos and R. Turrin, "Looking for "Good" Recommendations: A Comparative Evaluation of Recommender Systems," in IFIP Conference on Human-Computer Interaction, 2011.