The Application of Data Mining Techniques to Create Promotion Strategy for Mobile Phone Shop

A U Khasanah, K S Wibowo, and H F Dewantoro

Department of Industrial Engineering
Universitas Islam Indonesia
Indonesia
annisa.uswatun@uii.ac.id, 14522280@students.uii.ac.id, 14522297@students.uii.ac.id

Abstract. The number of mobile shop is growing very fast in various regions in Indonesia including in Yogyakarta due to the increasing demand of mobile phone. This fact leads high competition among the mobile phone shops. In these conditions the mobile phone shop should have a good promotion strategy in order to survive in competition, especially for a small mobile phone shop. To create attractive promotion strategy, the companies/shops should know their customer segmentation and the buying pattern of their target market. These kind of analysis can be done using Data mining technique. This study aims to segment customer using Agglomerative Hierarchical Clustering and know customer buying pattern using Association Rule Mining. This result conducted in a mobile shop in Sleman Yogyakarta. The clustering result shows that the biggest customer segment of the shop was male university student who come on weekend and from association rule mining, it can be concluded that tempered glass and smart phone “x” as well as action camera and waterproof monopod and power bank have strong relationship. This results that used to create promotion strategies which are presented in the end of the study.

Keyword: promotion strategy, mobile phone shop, agglomerative clustering analysis, association rule mining

1. Introduction
Mobile phone today is no longer just be a tertiary need, but it has become a primary need for almost everyone. Reported by Ericsson Mobility, in 2014 there were at least about 7 billion mobile users worldwide and this number is increasing by 7 percent annually. It is predicted that the number of smartphone users will reach 5.6 billion in the world [1]. In Indonesia, data extracted from the US Cencus Bureau by 2015 explains that mobile phone users have exceeded 281 million where as early as 2014 the number of Indonesians has reached 251 million [2]. This fact indicates that the need for communication devices especially mobile phones in Indonesia is very high and it makes Indonesia as the fourth largest mobile phone user in the world [3].

Along with the growing number of mobile phone users in Indonesia, the mobile industry is also increasingly stretched. The number of shops that sell mobile phones is growing very fast in various regions in Indonesia including in Yogyakarta. Now the store that sells mobile phones not only can be found in the city centre but almost in all corners of the city. This fact leads high competition among the mobile phone shops. The shops do not only sell mobile phone but they also sell mobile phone accessories such as power bank, action camera, monopod, case, tempered glass, earphone, and so on. Due to the high demand for mobile phones, the demand for those accessories is also high. In these conditions the mobile phone shop should have a good promotion strategy in order to survive in competition, especially for a small mobile phone shop.
Promotion is an activity that included in marketing strategy. This activity is very essential to inform the detail information about product or service and encourage the customer to buy or use it. Based on Chance [4] one important thing to be succeed in marketing including in promoting product or service is knowing the market target. This step can be conducted through customer segmentation. Segmentation is greatly embedded in a marketing strategy because different customer groups will require a different marketing mix [5]. Customer or market segmentation is the division of the market into different customer groups where specific customer groups can be selected as market targets to be achieved with a specific marketing mix [6]. After knowing the target market, a specific promoting strategy can be obtained for this specific target market.

For many shops, including mobile phone shops, customer segmentation is also essential step that must be done to create an attractive promotion strategy. Attractive promotion will encourage customer to buy their product. Usually, information that appeals to consumers is related to discounts, bonus, or product package with special prize. To create this kind of promotion, a company or shop should analysis the buying pattern of their target market, for example the company should know what kind of item that usually bough by their consumer in conjunction with another product.

The purpose of this study is to create an attractive promotion strategy for a mobile phone shop in Sleman Yogyakarta by conducting customer segmentation and analysing the product that commonly bought by the customer as the same time. This kind of analysis can be conducted using data mining techniques. Data mining is the process of automatically discovering useful information in large data repositories [7]. Data mining has several popular techniques such as clustering analysis and association rule mining. There are many methods that can be used in clustering analysis, one of them is agglomerative hierarchical clustering. In this study agglomerative hierarchical clustering, is used to conduct the customer segmentation, and association rule mining is used to discover an interesting pattern about the products that commonly bought by the customer at the same arrival time, then those results will be used as a recommendation to make an attractive promotion strategy. The rest of this paper is organized as follows, Section 2 present the literature review, Section 3 presents the research method related to this study, Section 4 shows the results and discussion. The concluding is finally made in Section 5.

2. Literature Study

2.1. Agglomerative Hierarchy Clustering
Hierarchical clustering analysis is the oldest clustering method when compared to other methods, however this method is still widely used. There are two basic approach to generate this method, agglomerative approach and divisive approach, but agglomerative approach is more commonly used. Agglomerative hierarchical clustering starts with the point as individual clusters and each step, merge the closest pair of clusters until only one cluster remains [7]. This method uses a bottom up strategy and the single cluster become the hierarchy’s root. In the merging step, two clusters which are close to each other are determined according to some similarity measure, then combine the two into one cluster [8]. A hierarchical clustering is usually displayed graphically using tree-like diagram called dendogram. The similarity measurement can be obtained by several methods such as min (single link), max (complete link), and group average.

There have been a lot of scholars in many different field of studies apply agglomerative hierarchical clustering. Iqbal et al [9] applied agglomerative hierarchical clustering to cluster 47 tomato genotypes into 5 clusters. Kang and Landry [10] use agglomerative hierarchical clustering to characterize, compare, and analyze eye movement sequence that occur during visual tracking of multiple moving targets. The eye movement data was collected from 25 expert air traffic controller and 40 novices through difficult scenarios on simulated radar display. The data was clustered into three different clusters, easy, moderate. While Kumar and Toshniwal [11] apply hierarchical clustering to identify the different reasons behind road accidents such as traffic characteristics, weather characteristics, road characteristics and etc. In this study trend analysis was also used to identify the increasing and decreasing accident rate in different reasons in Gujarat state India. Ip et al [12] implemented hierarchical clustering to cluster demand of electricity source for electric vehicle (BEV), the result than used to maximized the usability of battery.
recharging infrastructure. The plan should be designed where the charging station should serve the commuters demand and it can minimize the relevant operating cost.

2.2. Association Rule Mining

Association rule mining can be illustrated as follows, suppose that the set $I = \{i_1, i_2, \ldots, i_k\}$ is the total number of items. The item set $X$ and $Y$ are part of the total set $I$. If $Y$ happen in one transaction when $X$ happens, as association rule can be expressed as $X \Rightarrow Y$ [13]. In market basket analysis, association rule can be can be express as “A customer who buys product $X_1$ and $X_2$ will also buy product $Y$ with probability $c\%$” are found [14]. For example, in a restaurant egg and coffee may be ordered together primarily between 7 am and 11 am, and maybe turkey and pumpkins are sold at the same time in the week before Thanksgiving [15]. In service industry, association rule mining can be express as a customer who use a certain service will also use another service at the same service time.

The strength of an association rule can be measured in terms of its support and confidence. Tan et al. [7] explain that support determines how often a rule is applicable to a given data set. A low support rule is uninteresting from a business perspective and sometimes it will be eliminated. While, confidence determines how frequently items in $Y$ appear in transaction that contain $X$, the higher the confidence the more likely it is for $Y$ to present in transaction that contain $X$. The formal definitions of these metrics are,

$$
Support, s(X \Rightarrow Y) = \frac{\sigma(X \cup Y)}{N}
$$

$$
Confidence, c(X \Rightarrow Y) = \frac{\sigma(X \cup Y)}{\sigma(X)}
$$

Some scholars have applied association rule mining. Abdulsalam et al [16] implement association rule mining which is also known as market basket analysis to reveal sales pattern in a supermarket. This study showed rules of purchasing one product would lead to the purchase of another product. This study concluded that apple and chocolate had strong correlation. This information helps the company to create marketing and advertising strategies that outshine the competitors. Verma et al [17] apply association rule mining to find out the pattern incidents in a steel plant in India. Czibula et al [18] used association rule to predict defect of a software system in order to conduct continuous improvement. Kumar, & Toshniwal [19] apply data mining techniques to analyse data related with road accident in India. Cluster analysis was used to cluster the accident data, then association rule mining was used to identify the situation when an accident occurred in each cluster. The results can be used to make some preventions for each different categories, thus it can minimize the number of accidents happened.

3. Research Method

The research was conducted using Cross Industry Standard Process for Data Mining (CRISP-DM) framework. The CRISP-DM methodology is described in terms of a hierarchical process model, consisting of sets of tasks. CRISP-DM provides a non-proprietary and freely available standard process for fitting data mining into the general problem-solving strategy of a business or research unit. This framework was firstly developed in late 1996 [20]. CRISP-DM is divided into six phases Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment as shown in Figure 1.

The business process describes the background and objectives of this study as discussed in introduction in Section 1. Data understanding explain the initial data. The data was customer data from “ABC” mobile shop in Sleman Yogyakarta including the gender, occupation, age, arrival date, and the product that bought by the customer. The data were collected using questionnaire, then the data was cleaned and transformed in Data Preparation Step. After data cleaning, only 50 data left. The initial data was nominal and it need to be transformed into numeric. The data for clustering analysis must be numeric since it measures distance between data. The next step is modeling step. Agglomerative Hierarchical Clustering (AHC) and Association Rule Mining are used to create the model. Clustering is
used to segment the customer and association rule mining is used to know the several products that commonly bought by the customer at the same arrival time. This result that will be evaluated whether the model in fact achieves the objectives set or not. The modeling and evaluations step will further discuss in Section 4. The result from clustering and association rule will be used to create suitable promotion strategy for a certain customer segment and it will be discussed in deployment step in the last section (conclusion). Software SPSS was used to apply agglomerative clustering analysis and Rapid Miner was used to apply association rule mining.

4. Result and Discussion

4.1. Clustering Analysis

Modeling Step
In this step mobile shop customer was segmented based on several attributes including the gender, occupation, age, and arrival date. The 50 data were clustered using AHC, and some measurement distance including single linkage, complete linkage, and average linkage were used. Each measurement distance showed different output and based on the outlook of the dendogram, the output from average linkage was the most obvious to be interpreted.

Evaluation Step
The result of agglomerative hierarchical clustering is shown in dendogram (Figure 2.) The dendogram using average linkage was cut in 9.5 and it resulted 3 clusters and 1 outlier. The profile for each cluster is represented in Table 1. Cluster 1 consisted of 17 members, cluster 2 consisted of 25 members, and cluster 3 consisted of 7 members, thus cluster 2 had the biggest cluster members. There was no difference in average of age for all cluster. For cluster 1 the number of female and male were slightly difference, and it was dominated by university student who mostly come on weekdays. For cluster 2 all of the member were male, and it was dominated with university student, student, and businesswoman who come on weekend. While all members in cluster 3 were female, dominated with university student who mostly come on Sunday. From this result it was known mostly customer come on weekend and they rarely come on Monday. Then based on this result, the objective of clustering analysis in this study to know the specific customer segment was achieved.
Table 1. Cluster Profiling

| Cluster | Average Age | Gender | Occupation | Arrival date |
|---------|-------------|--------|------------|--------------|
| 1 (17)  | 21.76       | 59%    | 76%        | 18% mostly on Thursday, Wednesday, Tuesday |
|         |             | 41%    | 0%         | 6%           |
| 2 (25)  | 21.76       | 100%   | 32%        | 32% mostly on Friday, Saturday, Sunday |
|         |             | 0%     | 32%        | 4%           |
| 3 (7)   | 21.57       | 0%     | 43%        | 14% mostly on Sunday |
|         |             | 100%   | 29%        | 14%           |

Figure 2. Dendogram using average linkage

4.2. Association Rule Mining

Modeling Step

This step was conducted to reveal the customer buying pattern on the “ABC” mobile phone shop. To get the information, the customer was asked related with what kind of products that they bought. Then the data was transformed into binary form, 0 for not buying the product and 1 for buying the product. Then, the data was put into Rapid Miner software. Minimum support was adjusted at 5% and confidence was 5%. There were lots of rules generated but only 18 rules had lift ratio bigger than 1. Lift ratio measure the correlation between item set and it can use to measure the importance of the rule. Greater lift ratio indicates stronger associations, the highest lift which means highest correlation [21,22]. From the 8 rules, there are 3 interesting rules, as follows

- tempered glass ⇒ smart phone "x" (s:16%, c:53%)
- action camera ⇒ waterproof, monopod (s:16%, c:61.5%)
- action camera ⇒ power bank (s:24%, c:50%)
Evaluation Step
The first rule means 16% of all transaction contain tempered glass and smart phone “x” and 53% of transaction have the item tempered glass also contain smart phone “x”. The second rule means 16% of the transaction contain action camera, waterproof camera case, and monopod, and 61.5% of the transaction have item action camera also contain waterproof camera case and monopod. While the third rule means 24% transaction contain action camera and power bank, and 50% of the transaction have action camera also contain power bank. Based on this result, the objectives of association rule to analyse the product that strongly related and commonly bought by the customer as the same time was achieved.

5. Conclusion
From the clustering analysis the customer segmentation can be known and from association rule mining, the important customer buying pattern can be revealed. This information can be used to make suitable promotion strategy. Not all the information is used to make the promotion strategy. From the clustering result it can be concluded that the biggest customer segment come to the shop were male university student who come on weekend, and customer rarely come Monday. While from association rule mining, it can be concluded that tempered glass and smart phone “x” as well as action camera, waterproof camera case, monopod and also action camera with power bank have strong relationship. Thus, the shop can make promotion to give free tempered glass for university student who buy smart phone “x” on Monday, or discount for waterproof camera case for university student who buy action camera and monopod on Monday, or it can be also give special price for university student who buy action camera and power bank on Monday. The promotion strategies are illustrated in Figure 3.

![Figure 3. Promotion strategy](Source for mobile phone picture: [23]), (source for camera picture: [24]), (Source for power bank picture: [25]), (source for monopod picture: [26])

6. References
[1] Nistanto R K 2014 2015 Pengguna “Mobile” Lampau Jumlah Penduduk Dunia online access on 10 September 2017 URL:// http://teknokompas.com/read/2014/06/04/1025003/2015.Pengguna.Mobile.Lampau.Jumlah.Penduduk.Dunia
[2] Bursadana 2016 Jumlah Pengguna Ponsel di Indonesia online access on 10 September 2017 URL:// http://media.bursadana.co.id/2016/03/18/jumlah-pengguna-ponsel-di-indonesia.html
[3] Wahyudi A 2015 Indonesia Raksasa Teknologi Digital Asia online access on 10 September 2017 URL:https://www.tempo.co/read/kolom/2015/10/02/2310/indonesia-raksasa-teknologi-
digital-asia
[4] Chance J 2017 8 Keys to a Strong Marketing Strategy online access on 10 September 2017 URL://
https://www.businessknowhow.com/marketing/blocks.htm.
[5] Kahreh M S, Tive M, Babania A and Hesan M 2014 Analyzing the applications of customer
lifetime value (CLV) based on benefit segmentation for the banking sector Procedia-Social
Behav. Sci. 109 590–4
[6] Kotler P 1980 Marketing Management: Analysis, Planning, and Control, 4th edition (Englewood
Cliffs, N.J: Prentice-Hall)
[7] Tan P N, Steinbach M and Kumar V 2006 Introduction to Data Mining (MA: Pearson
Education)
[8] Han J, Kamber M and Pei J 2006 Data Mining: Concepts and Techniques, 2nd edition (MA: Elsevier)
[9] Iqbal Q, Saleem M Y, Hameed A and Asghar M 2014 Assessment Of Genetic Divergence In
Tomato Through Agglomerative Hierarchical Clustering And Principal Component Analysis
Pak. J. Bot 46 1865–70
[10] Kang Z and Landry S J 2015 An Eye Movement Analysis Algorithm for a Multielement Target
Tracking Task: Maximum Transition-Based Agglomerative Hierarchical Clustering IEEE
Trans. Human-Machine Syst. 45 13–24
[11] Kumar S and Tosniwal D 2016 Analysis of hourly road accident counts using hierarchical
clustering and cophenetic correlation coefficient (CPCC) J. Big Data 3
[12] Ip A, Fong S and Liu E Optimization for Allocating BEV Recharging Stations in Urban Areas by
Using Hierarchical Clustering 460–5
[13] Han H K, Kim H S and Sohn S Y 2009 Sequential association rules for forecasting failure patterns
of aircrafts in Korean airforce Expert Syst. Appl. 36 1129–33
[14] Nakhaeizadeh G, Hipp J and Güntzer U 2000 Algorithms for association rule mining - a general
survey and comparison ACM sigkdd Explor. Newsl. 2 58–64
[15] Lia Y, Ningb P, Wanga X and Jajodiaa S 2003 Discovering Calendar-based Temporal Association
Rules G Discovery.Csc.Nesu.Edu 44 193–218
[16] Abdulsalam S O 2014 Data Mining in Market Basket Transaction : An Association Rule Mining
Approach International Journal of Applied Information Systems 7 15–20
[17] Verma A, Khan S D, Maiti J and Krishna O B 2014 Identifying patterns of safety related incidents
in a steel plant using association rule mining of incident investigation reports Saf. Sci. 70 89–98
[18] Czibula G, Marian Z and Czibula I G 2014 Software defect prediction using relational association
rule mining Inf. Sci. (Ny). 264
[19] Kumar S and Tosniwal D 2015 Analysing road accident data using association rule mining
Computing, Communication and Security (ICCCS) (IEEE) pp 1–6
[20] Chapman P, Clinton J, Kerber R, Khabaza T, Reinartz T, Shearer C and Wirth R 2000 CRISPDM
1.0 step-by-step data mining guide (Technical report: CRISP-DM)
[21] Hahsler M, Grün B and Hornik K 2007 Introduction to arules—mining association rules and
frequent item sets. SIGKDD Explor 2 4
[22] El-Halees A 2009 Introduction to arules—mining association rules and frequent item sets SIGKDD
Explor 2
[23] Ilnanny 2012 online accesses in October 2017 URL: https://openclipart.org/detail/168116/nexus-
one
[24] Sanity D 2014 online access in October 2017 URL: https://dribbble.com/shots/1741674-Camera-
Icon
[25] Olike 2017 online access in October 2017 URL: http://olike.com.my/products-power-bank/
[26] Banggood 2017 online access in October 2017 URL: https://www.banggood.com/Original-
Bluetooth-Remote-Control-Monopod-Selfie-Stick-for-Xiaomi-Yi-Sports-Camera-p-982396.html