

Decision Support System In Determining Class On Accupuncture Clinic

Basiroh¹, Mohammad Nur Hilal²

{basiroh@politeknikcilacap.ac.id¹, najiwaa@yahoo.com²}

¹Departemen Information Technology, Politechnic State of Cilacap, Cilacap, Indonesia

²Department of Mechanical engineering Politechnic state of Cilacap, Indonesia

Abstract: The process of determining promotional class in acupuncture clinics is very common. Clinics also have data i.e form of primary data and customer secondary data. This happens repeatedly and generates a build up of customer data that affects information retrieval of the data. This study aims to grouping the clinic customer data in which those who are a greater contribution value get a valuable promotion as well. The Acupuncture Clinic uses decision support system by utilizing data mining process by using Clustering technique. The method used is research action through Diagnosing process, Action planning, Action taking, Evaluating, Reflection. The algorithm used for cluster formation is the K-Means algorithm. K-Means is one of the non-hierarchical clustering data methods that can group customer data into multiple clusters based on the similarity of the data, so the customer data with similar contribution values are grouped in one cluster and those with different contribution values are grouped into other clusters. Implementation using PHP is used to find accurate values. Attributes used are customer earnings, total clinical outcomes, repeat visits, product purchases, needle types and therapists. The customer cluster formed is four clusters, with the first cluster 5 customers, the second cluster 9 customers and the third cluster a total of 6 customers and the fourth cluster there are 5 customers. The results of this study are used as one of the basic decision-making to determine promotion based on the clusters formed by the administration of acupuncture clinics.

Keywords: DSS, promotion, clustering, k-means algorithm.

1. Introduction

In business an important aspect in determining the viability of a company is customer loyalty. In addition to providing financial benefits over the long term, customer loyalty can build a positive image for the product or service offered. The Acupuncture Clinic should also recognize what customers want through an effort to manage good relationships with customers. Meanwhile, the number of customers is not small and they also have various characteristics. Problems arise when the clinic conducts an anniversary or month promo, In order to increase loyalty to develop customer-centered business strategy or Customer Relationship Management (CRM) (Ivan, 2005). The goal is to create good relationships with customers to enhance mutually beneficial cooperation. This activity has become the routine agenda of Acupuncture Clinic as one way to strengthen and to maintain good relationship of customers, whether they are priority or regular customers, but some customers feel the gap and are not in accordance with the desired promotional gifts by customers. Customers argue that they feel they have contributed the same but are not get the desired promotion.

Based on these problems the clinic applies a decision support system using clustering method. Customer data are grouped so it can be found out any customers who are entitled to the promotion in accordance with the value of the contribution given. From the data they are formed into several groups that are then processed to be a system for decision makers using decision methods.

By k-mean clustering method, the clinic can integrate and automate customer service process in field of promotion, information, and service of clinic-related products or services. K-Means method is used to define the grouping of criteria values from several alternative value groups, The value of this grouping is used to determine the alternate list results that is going to will be used as recommendation for decision support (Supranto, 2000) by the Clinic. So the data mining is only used to help decision support system.

The result of a system is a model or prototype of information service function. The system is, developed periodically to adapt with to the circumstances and development needs of the system. For problems arise, the system produces a systematic effective solution.

2. Research Method

2.1 Decision Support System

The decision support system (DSS) is part of a computer-based information system (including a knowledge system) (Daihani, 2001) used to support decision making within an organization or company (Supranto, 2000). DSS is a combination of individual intelligence resources with component skills to improve the quality of decisions (Turban et al., 2005). Decision Support System is also a computer-based information system for decision-making management that deals with semi-structural problems (Kusrini, 2007).

2.2 Clustering

Basically clustering is a method to search and classify data that has similarity between one data and other. In data mining, there are two types of clustering methods, i.e. hierarchical clustering and non-hierarchical clustering. Based on the data type, the data is divided into several types. The types of data can be divided according to their nature, according to the source, by means of obtaining it and by the time of collecting the data (Kusrini, 2007; Santosa, 2007).

2.3 Qualitative Data

Qualitative data can simply be called non numeric data. The main characteristic of qualitative data is obtained by counting, so it has no decimal value. Besides qualitative data have characteristics, i.e. they cannot be performed mathematical operations, such as addition, subtraction, multiplication and division. Qualitative data can be divided into two types, namely nominal data and ordinal data. Nominal Data is a type of qualitative data used to identify, classify, or distinguish objects. Nominal data is the lowest data in the level of data measurement. All data have an equal position in the sense that no data has more or less level than other data (Santosa, 2007).

2.4 Quantitative Data

Quantitative data can be referred to as data in the form of numbers in the true sense. Thus, various types of mathematical operations can be performed on quantitative data. Quantitative data is data obtained by way of measuring so they have decimal value.

2.5 K-Mean Algorithm

K-means clustering is one of the non-hierarchical data clustering methods that classifies data in the form of one or more clusters. The steps of clustering with K-Means method by (Hazibuan Z. A., 2007) are as follows:

1. Selection the number of clusters k.
2. The initialization of this k center cluster can be done in various ways. But the most often done is in random. Cluster centers are assigned initial values with random numbers.
3. Allocating all data / objects to the nearest cluster. The proximity of two objects is determined by the distance of the two objects. By using Euclidean distance the proximity of each data can be calculated. Some ways of calculating the distance commonly used are:

- a. Euclidean distance

$$C_i = \min + \frac{(i-1) * (\max - \min)}{n} + \frac{(\max - \min)}{2 * n} \quad (1)$$

- b. The spacing formula between two points in one, two and three dimensions is sequentially shown in the following formula 1, 2, 3:

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2} \quad (2)$$

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2} \quad (3)$$

- c. Manhattan Distance

Manhattan distance is also called taxicab distance.

$$d_1(p, q) = \|p, q\|_1 = \sum_{i=1}^n |p_i - q_i| \quad (4)$$

- d. Chebichev Distance

$$D_{cheb}(p, q) = \max(|p_i - q_i|) \quad (5)$$

In Chebyshev distance or Maximum Metric the distance between points is defined by taking the largest difference value of each coordinate dimension.

- Recalculating cluster center with current cluster membership. The cluster center is the average of all data /objects in a particular cluster. It can also use the median of the cluster, so the mean is not the only size that can be used.
- Reassigning each object using the new cluster center. If the center of the cluster does not change anymore, then the clustering process is complete. Alternatively, return to step number 3 until the center of the cluster does not change anymore.

2.6 Promotions

According to Kotler, promotion, the fourth marketing mix tools, stand for various activities, the company undertakes to communicate its products merits and to persuade target customers to buy them. The definition has the sense that promotion includes all the tools contained in the promotion mix, in which the main role is to make a persuasive communication (Kotler, 1997).

In this study, the data are primary data and secondary data, The method used is Action research that aims to develop new approaches in problem solving with direct application in cyberspace. Action research includes Diagnosing, Action planing, Action taking, Evaluation, Reflektng (Kusrini., 2007). Primary data are taken from research object. This research starts from the data collection it is done through interview and observation. In this case the researchers obtained the results about the general description of the object , the input and the idea of the object, The direct statement from the patient. Then it formulate the problems and limit so the explanation does not come out of the scope of the problem. a literature study is done to get theories about the problems (Turban et al., 2005).

The data used in this study are obtained from the administration in the form of patient data in acupunctre clinic since 2015. These data contain self-data from patients, patient medical records, This study only uses of the data, such as patient name, profession, visitation and total value of patient transactions. The data are initialized into figure form through several steps so the data can be processed by using K- means Clustering algorithm. The workplot design of the class promotion in acupunctre clinics can be seen.

2.6.1 Data Transformation

Numeric Data such as profession have to be initialized first into the form of numbers / numerical. To initialize the profession can be done with:

- Initializing of data

Table1. Intizializing of patien Profession

No	Profession	Freque ncy	Initializati on
1	Businessman	8	1
2	Employees	6	2
3	Teacher	5	3
4	Enterpreneur	5	4
5	Labour	2	5

- The profession that has the largest frequency is initialized with the number 1 and the profession which has the second largest frequency is initialized with the number 2, and so on.

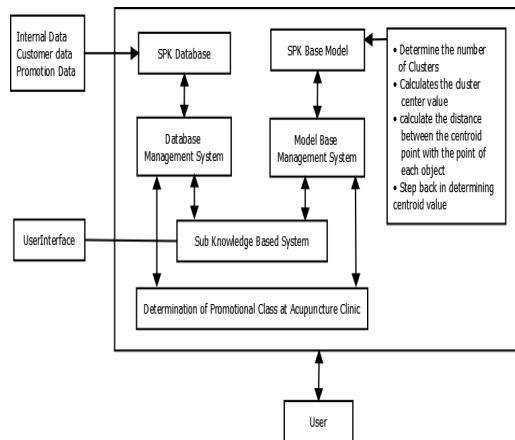


Fig. 1. The design of the Promotion workflow.

Table 2. Customer / patience dataset present at clinic

No	Nama	Profesi	Repeat	Total Result
1	Harjono	2	4	1.160.00
2	Mulyono	3	4	1.420.000
3	Mujilah	4	3	750.000
4	Suyono	2	4	1.160.000
5	Gunadi	2	2	530.000
6	Widadi	3	3	1.410.000
7	Marsono	1	5	2.175.000
8	Muharwani	1	4	1.880.000
9	Mulyadi	3	3	1.425.000
10	Dwi M	1	4	1.880.000
11	Sumedi	3	3	1.485.000
12	Junaedi	2	3	1.005.000
13	Jumiyatun	3	4	1.452.000
14	Kamarun	2	2	1.060.000
15	Joko W	4	4	750.000
16	Rubimin	2	3	1.160.000
17	Eko	3	3	1.350.000
18	Ponirah	3	3	1.305.000
19	Sumirah	3	4	1.305.000
20	Jumingan	1	4	1.900.000
21	Siska	1	3	1.900.000
22	Retno	2	3	1.065.000
23	Susilo	4	3	795.000
24	Marto	4	3	795.000
25	Ngatiyo	3	3	1.425.000

- **Processing Data**
After all customer data are transformed into numeric form, then the data can be grouped by using K-Means Clustering algorithm.

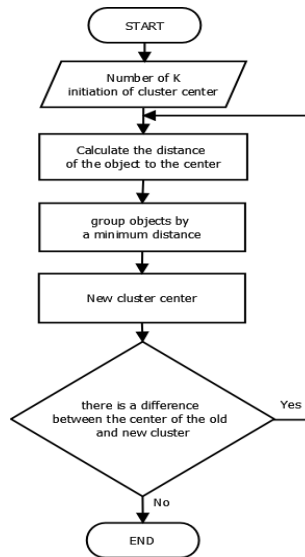


Fig. 2. Flowchart K-Mean.

To group the data into several clusters needs to do some steps, namely:

- Determining of the number of clusters (K), i.e. four clusters.

Table3. Central Starting Point for each Cluster.

Start	Name	Profession	Repeat	Total
C1	Marsono	1	5	2.175.000
C2	Widodo	2	3	1.410.000
C3	Junaidi	3	3	1.005.000
C4	Rubimin	3	3	1.160.000

- Calculating the distance of each data to the center of the cluster between objects to the centroid by calculating the Euclidean distance. The equation used is:

$$d(p,q) = \sqrt{(p1-q1)^2 + (p2-q2)^2 + (p3-q3)^2} \quad (6)$$

The above 2-3 equations are used because the attributes used are 3.

For example, it calculate the distance from the first customer data to the center of the first cluster with the equation:

$$d(p,q) = \sqrt{(1-1)^2 + (1-5)^2 + (2175000-1160000)^2} = 1015,0078 \quad (7)$$

The distance of the first customer data to the second cluster center:

$$d(p,q) = \sqrt{(3-1)^2 + (4-3)^2 + (1420000-1410000)^2} = 3162,2784 \quad (8)$$

The first customer data distance to the third cluster center:

$$d(p,q) = \sqrt{(4-2)^2 + (3-3)^2 + (795000-1005000)^2} = 210000 \quad (9)$$

The first customer data distance to the fourth cluster center:

$$d(p,q) = \sqrt{(2-3)^2 + (4-3)^2 + (1160000-1160000)^2} = 1,414215 \quad (10)$$

- Grouping data into clusters with minimal distances

- Once all the data is placed into the nearest cluster, then recalculating the new cluster center based on the average members in the cluster.
- After obtaining a new center point from each cluster, going back from the third step until the center point of each cluster does not change again and no more data is moved from one cluster to another cluster.

3. Result And Discussion

3.1 Result

Based on the result of data grouping using k-means clustering method, the clustering result showed that when it reached 7th iteration, the center point no longer changed and no data moved between clusters. The Results where show in table 4.

2.2 Discussion

From cluster 1 it could be seen that it consistence consisting of 5 customers whose transaction value here large and all work as entrepreneurs, then for cluster 2 it had customers with average profession as Teachers and with 3 time in a week visit. For cluster 3 the average consisted of employees with an average 4 times visit, while for cluster 4, it consisted of customers who work as entrepreneurs and had average visit of 3 times in a week.

Table 4. Cluster Analysis Results.

Cluster 1		Cluster 2	
Marsono	=Total results 2.175.000	Mulyono	=Total results 1.420.000,
Profession 1		Profession 3	
Muharwani	=Total results 1.880.000,	Widodo	=Total results 1.410.000
Profession 1		Profession 3	
Dwi	=Total results 1.880.000, Profession 1	Sumarsih	=Total results 1.305.000
Jumingan	=Total results 1.900.000,	Profession 4	
Profession 1		Mulyadi	=Total results 1.325.000
SiskaW	=Total results 1.900.000,	Profession 3	
Profession 1		Ngatiyo	=Total results 1.425.000
		Profession 3	
		Sumadi	=Total results 1.485.000
		Profession 3	
		Jumiyatun	=Total results 1.425.000
		Profession 1	
With an average of 4 visits in week		Eko budi	=Total results 1.350.000
		Profession 3	
		Ponirah	=Total results 1.305.000
		Profession 3	
		With an average of 3 visits in 1 week	

Cluster 3		Cluster4	
Harjono	=Total results 1.160.000,	Mujilah	=Total results 795.000, Profession 4
Profession 2		Gunadi	=Total results

Suyono	=Total	530.000, Profession 5
results	1.160.000,	Profession 2
.....

4. Conclusions

From the results of research can be concluded that the value of the transaction by the patient affects the value to get the promotion.

- Promoting continuously to maintain customer value to perform the most preferred treatments.
- Conducting data collection or doing recap on customer data each day so active customer data are not re-mixed with non-active data

Acknowledgements

The Author thank to Departemen of Informatic Engineering and departemen mechanical engineering, Politechnic State of Cilacap for their support to this work, Author would like to acknowledge the support of minister of research and Height education.

References

- [1] Daihani, D. U. (2001) *Komputerisasi Pengambilan Keputusan., PT. elex Media Komputindo*. Bandung: Andi The publishing company.
- [2] Hazibuan Z. A. (2007) *Metodelogi Penelitian Pada Bidang Ilmu Komputer dan Teknologi Informasi*. Jakarta.
- [3] Ivan, A. H. (2005) *Desain target optimal, Laporan Penelitian Hibah Bersaing,Proyek Multitahun*. Jakarta: Dikti.
- [4] Kotler, P. (1997) *Manajemen Pemasaran: Analisis, Perencanaan, Implementasi dan Kontrol*. 1st Editio. Jakarta.: Prenhallindo The publishing company.
- [5] Kusriani. (2007) *Konsep dan aplikasi Sistem Pendukung Keputusan. Andi*. Yogyakarta: Andi The publishing company.
- [6] Santosa, B. (2007) *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu.
- [7] Supranto, J. M. A. (2000) *Statistik: Teori dan Aplikasi Edisi Keenam*. Jakarta: Erlangga The publishing company.
- [8] Turban et al. (2005) *Decicion Support Systems and Intelligent Systems*. Yogyakarta: Andi The publishing company.