ISSN No:-2456-2165

Business Intelligence Analysis of Blood Donor Data Management (Case Study UDD PMI Cirebon City)

Sendy Hendriawan
Department of informatics Widyatama
University Bandung, Indonesia

Sunjana
Department of informatics Widyatama
University Bandung, Indonesia

Abstract:- In business intelligence, we can analyze data, one of which is by using clustering. Clustering can be used to group data into groups, the dataset is divided into groups.

Keywords:- *Business Intelligence, Clustering, K-Means.*

I. INTRODUCTION

Business Intelligence is a series of activities to understand the business situation by conducting various types of analysis of data held by organizations and external data from third parties to help determine strategies, tactical business decisions, and operations and take the necessary actions to improve business performance [1]. In Business Intelligence, there are several methods that can be used in clustering, including K-Means.

K-means can be defined as a grouping algorithm. The purpose of this algorithm is to divide data into groups. This algorithm accepts input in the form of data without class labels. This is different from supervised learning which accepts input. This algorithm will classify data or objects into groups of k. And in each cluster there is a center point (centroid) which represents the cluster, K-Means. This algorithm is often used as an efficient clustering method [2].

Business intelligence here displays the data that has been clustered. This will determine how data that was previously still irregular will appear more organized because it has been clustered into groups.

II. RELATED WORK

There have been several studies using clustering and k-means in various fields, for example to process blood donor data [3], or in analyzing the performance of clustering and k-means using Rapid Miner or Pentaho. In this research, there are several libraries that are used in the manufacture is Pentaho as a program or algorithm, Rapid Miner to display the results.

III. DATASET

In this study, using a dataset from UDD PMI Cirebon City with 6719 data managed by UDD PMI Cirebon City in which there are various kinds of data variables. For this study, only in year 2015 data were used.

IV. CLUSTERING

Clustering is a way to solve complex data, it requires algorithms in grouping new data sets such as ensemble grouping, computing software and graphical algorithms [4].

Clustering can be called a data analysis method that is often used for data mining which has the aim of grouping data with the same characteristics into the same 'region' and data with different characteristics into another 'region'. One of the clustering algorithms is called K-Means. K-Means is an algorithm that is commonly used in the clustering process. The K-Means method is able to divide data into several groups.

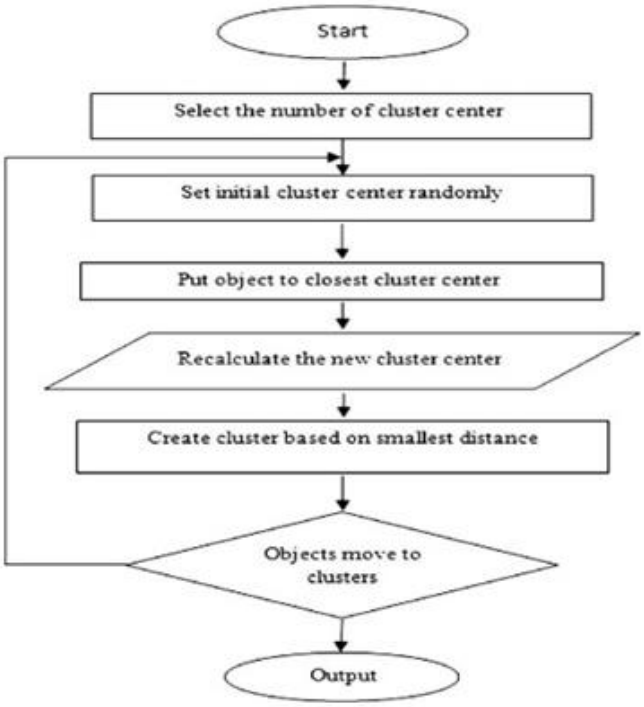

Fig 1:- Flowchart Clustering

V. K-MEANS

K-Means is a grouping algorithm. Because this algorithm is responsible for dividing data into groups. This algorithm accepts input in the form of data without class labels. This is different from supervised learning which accepts input. grouping data into its own input without first knowing the target class, Input received is the data or object and the desired cluster. This algorithm will classify data or objects into groups of k. In each cluster there is a center point (centroid) which represents the cluster, K-Means Algorithm is used as an efficient clustering method [2].

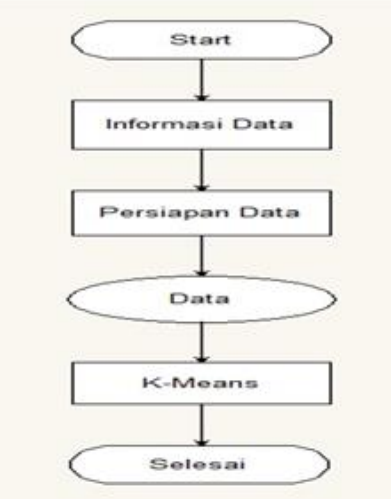

Fig 2:- K-Means flow

In Figure 3, The K-Means workflow is explained that there are four flows that begin with data information, we must know what information about our data looks like, continue with data preparation, after we know the data information we immediately prepare what to do with the data, continue the data at if using the k-means algorithm in its clustering

VI. RESULT

A. K-Means

Attribute	Cluster_0	Cluster_1	Cluster_2	Cluster_3
JK	1.612	1.617	1.628	1.643
Jenis Darah	2.062	2.002	2.040	2.040
Gol	2.562	2.573	2.638	2.584
Rumah Sakit	13.903	9.545	2.028	5.668

Fig 3:- Results of K-Means

Before getting the results from K-Means, the data is entered for clustering using K-Means, after processing the centroid results can be seen.

Row No.	id	cluster	JK	JENIS DAR...	GOL	RUMAH SA...
1	1	cluster_2	1	3	1	1
2	2	cluster_2	2	3	1	2
3	3	cluster_2	1	1	3	3
4	4	cluster_2	1	3	4	3
5	5	cluster_3	1	3	2	4
6	6	cluster_3	1	3	2	5
7	7	cluster_3	2	1	2	6
8	8	cluster_3	2	1	2	4
9	9	cluster_2	2	3	2	3
10	10	cluster_0	1	3	4	15
11	11	cluster_3	1	1	4	7
12	12	cluster_2	1	1	3	1
13	13	cluster_1	2	1	1	8
14	14	cluster_1	1	1	1	9
15	15	cluster_1	2	1	1	10
16	16	cluster_3	1	1	2	7
17	17	cluster_1	1	3	2	10
18	18	cluster_2	2	3	4	2

Fig 4:- Result K-Means

This is a display of the clustering results when viewed from a data view, the number of clusters consists of cluster_0 to cluster_4 for significant differences seen in the hospital data.

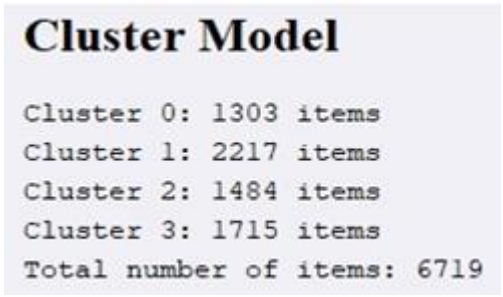

Fig 5:- Result Cluster Model

Here it can be seen from the results of the data as much as 6719 data, on cluster_0 as much 1303 items, on cluster_1 as much 2217 items, on cluster_2 as much 1484 items, on cluster_3 1715 items.

B. Conclusion

To conclude, after the data input was successfully clustered using k-means it produced as many as 4 clusters, the most visible difference was in the hospital data, while for gender, blood type, blood group were all in each cluster, while for there are only a few hospitals in each cluster. And it is proven that clustering using k-means can be done even though it uses a very large amount of data.

REFERENCES

[1]. Rahmat Tri Yunandar, Amir and Khairul Rizal, “Perancangan Data Warehouse Untuk Informasi Strategi Studi Kasus Penerimaan Siswa Baru STIE Binaniaga Bogor”, Journal of Computer Engineering AMIK BSI, Indonesia, Vol. 06, No.1, January 2020.

[2]. Kumari, R., Singh, M., Jha, R., Singh, N., et al. “Anomaly detection in network traffic using k-means clustering, ”In Recent Advances in Information Technology (RAIT), 2016 3rd International Conferences on, IEEE (2016), 387-393

[3]. Maryam Ashoori, Zahra Taheri, "Using Clustering Methods for Identifying Blood Donors Behavior", 5th Iranian Conference on Electrical and Electronics Engineering (ICEEE 2013).

[4]. I. Lahmar, A. Ben Ayed, M. Ben Halima and A.M. Alimi, “Cluster Forest Based Fuzzy Logic for Massive Data Clustering”, Ninth International Conference on Machine Vision (ICMV’2016), Nice, France, Nov 18-20, 2016, International Society for Optics and Photonics, SPIE Vol. 10341, pp. 103412J-1-103412J-5