

ABSTRACT

DATA MINING ANALYSIS FOR CLUSTERIZATION MEDICAL RECORD DATA USING K-MEANS ALGORITHM IN PORT MEDICAL CENTER, JAKARTA

By
DANTI PUJI YUHERISNA
NIM : 1202160154

Along with the rapid development of Information Technology (IT), all fields need data for information storage. One that requires data to store information as an important part is the hospital. Awareness of the importance of health for one's body provides an increase in the quality of health every year. Hospital is one of the crucial fields that requires data to store all information, such as data about patient data, drug data, etc. Therefore hospitals need data storage to classify what diseases are often plagued by hospital patients and what patients are included in the cluster.

As in one case study that occurred at one hospital in North Jakarta, there was a lack of use of data as a medium for analyzing deeper information. Therefore, hospitals need to use data mining for clustering patients based on medical records. The purpose of using data mining itself is a concept that can be used to group data based on certain categories. Data mining for the Port Medical Center Jakarta Hospital works to dig deeper information and so that the data stored does not only become reports and graphs for hospitals. Data Mining here is also useful to be able to find patterns that are repetitive and valuable that are often hidden in a pile of data using clustering techniques

Clustering technique functions to classify attributes and divide into based on attributes that are owned in medical record data using K-means algorithm. The reason the K-Means algorithm is also a suitable algorithm is because the data obtained by its attributes can be calculated. In addition, the hope is to be able to produce the most disease diagnosis information based on the existing attributes in the medical record. In addition, the purpose of this study using the k-means algorithm is that it is expected that later the output of this study can be a reference for hospitals to group or clustering patient data based on existing attributes so that they can do the handler faster and the handlers can be done correctly because they already know the patient what clusters fall into.

In addition, the purpose of this study using using K-Means algorithm is expected to be a reference for hospitals to group patient data based on the attributes that exist in the medical record at the hospital. Meanwhile, the results of this study are have 4 clusters, the cluster is taken from the smallest bouildin index davies with a value of 0.472. As for the category, cluster 0 is categorized into late adolescents until early adulthood with 2898 patients with a percentage of 31%, for cluster 1 categorized into the elderly with 1696 patients with a percentage of 18%. Cluster 2

falls into the category of infants to early adolescents and has 1565 patients with a percentage of 17%. Cluster 3 is the final adult category with a total of 3148 patients with a percentage of 34%.

Keywords: data mining, K-Means, clustering, patients