ABSTRACT

CLUSTERING ANALYSIS ON CUSTOMER DATA IN PT PLN (Persero) DISTRIBUTION WEST JAVA AND BANTEN APJ BANDUNG

By

Berry Zulkarnaen

10108123

PT PLN (Persero) introduced a system of classification of customers based on the type and capacity utilization are connected. There are six the current classification of the Household, Industrial, Social, Business, Government, and street lighting. Each faction has a base rate of electricity and different subscription. However, over time and technological developments are common similarities in the amount of electricity consumption between groups. This happens because there is the possibility of errors in the utilization goals. It is necessary for the analysis of customer grouping using clustering method.

Clustering method is one of the techniques of data mining that has a function to find a group / group naturally using similarities as well as dissimilarities benchmark. The application of the method of clustering is to classify customers based on similar amounts of electricity each month (KWh). To measure the similarity distance between the data used K-means algorithm. In the k-means algorithm, the proximity principle represented by the value of the centroid (mean), after calculating distance using Euclidian Distance technique the data that has the closest distance grouped in a single cluster, and apply to all data.

After testing the clustering process will be visible customer groups formed. It will be seen that the intensity of the class of customer electricity consumption is almost similar to other customers but there are different groups that have been set previously. Thus it can be concluded that the clustering method able to find or create groups of customers based on the intensity of the electricity consumption (KWh).

Keywords*: classification, group, clustering, k-mean, centroid, Euclidean Distance, and clusters.*