

## Abstract

Clustering is one of data mining's methods that used to grouped a set of objects into a cluster based on the similarity of its characteristics. In a large dataset, there is a possibility that more objects have difference characteristic with another, this object called *outlier*. The small number of *outlier* from a lot of objects causes the *outlier* difficult to detect. It is oftenly there are some important information can used from that anomaly.

Clustering can used to detect the existence of *outlier*. In this study of final task, *outlier* get from postprocessing technique of result by CLARANS (*Clustering Large Application based on RANdomized Search*) algorithm. CLARANS applying principle of k-medoid and combining the sampling technique with PAM to find the best medoid from each cluster. Because of clustering CLARANS disable to detect the *outlier*, then it need postprocessing technique after clustering. There are two postprocessing techniques, Clustering Removal *Outlier* (COR) and Separation.

Result of system obtain that accuration of detection *outlier* from clustering CLARANS is more better by COR in condition ideal number of cluster. Whereas in separation technique, value of accuracy influenced by number of cluster and distribution model of data.

**Keywords:** *clustering, outlier detection, k-medoid, CLARANS, clustering outlier removal, separation*