

ABSTRAK

Air adalah salah satu unsur penting dalam kehidupan manusia, sehingga harus diperhatikan kualitasnya untuk kehidupan manusia. Air yang layak untuk dikonsumsi harus memenuhi persyaratan fisika, mikrobiologis, kimiawi dan radioaktif. Beberapa orang kurang memahami pengetahuan tentang kualitas air, terlebih berdasarkan karakteristik yang dimiliki. Dataset yang ada belum dapat digunakan secara maksimal tanpa adanya informasi dan Kesimpulan yang diambil dari data tersebut. Dengan demikian, data tersebut akan diolah dengan data mining menggunakan metode K-Means dan DBSCAN. Tujuan dari penelitian ini adalah agar dapat memperoleh pengelompokan air berdasarkan Tingkat Kesehatan dan mengetahui seberapa baik cluster yang terbentuk yang dievaluasi menggunakan Silhouette Index. Melalui penelitian yang telah dilakukan menggunakan 2 algoritma tersebut telah diperoleh algoritma dengan SI terbaik. Algoritma DBSCAN adalah algoritma yang pembentukan clusteringnya mendapatkan SI terbaik sebesar 0.5201. SI terbaik yang diperoleh melalui algoritma DBSCAN adalah ketika menggunakan Epsilon = 0.41 dan MinPts = 15. Dengan demikian cluster yang terbentuk pada algoritma DBSCAN memiliki struktur yang medium.

Kata Kunci : *Data Mining, K-Means Clustering, DBSCAN, Kualitas air, Silhouette Index.*

ABSTRACT

Water is one of the important elements in human life, so its quality must be considered for human life. Water that is suitable for consumption must meet physical, microbiological, chemical and radioactive requirements. Some people lack knowledge about water quality, especially based on its characteristics. Existing datasets cannot be used optimally without information and conclusions drawn from the data. Thus, the data will be processed with data mining using the K-Means and DBSCAN methods. The purpose of this research is to be able to obtain water grouping based on Health Level and find out how well the clusters formed are evaluated using the Silhouette Index. Through research that has been done using the 2 algorithms, the algorithm with the best SI has been obtained. The DBSCAN algorithm is the algorithm whose clustering formation gets the best SI of 0.5201. The best SI obtained through the DBSCAN algorithm is when using Epsilon = 0.41 and MinPts = 15. Thus the cluster formed in the DBSCAN algorithm has a medium structure.

Keywords: Data Mining, K-Means Clustering, DBSCAN, Water quality, Silhouette Index.