INTERNET TRAFFIC CLUSTERING AND ANALYSIS BASED ON FUZZY C MEAN WITH DATA FEATURE EXTRACTION

Nama Mahasiswa : Adi Suryaputra Paramita
NRP : 5110 202 018
Supervisor : Dr.Eng. Febriliyan Samopa, S.Kom, M.Kom
Co-Supervisor : Bekti Cahyo Hidayanto, S.Si, M.Kom

ABSTRACT

This study is purpose to test the internet bandwidth clustering method modified by adding feature extraction pre-processing the data, the process of extraction is to make the selection key features Internet bandwidth and then extraction with search features that are correlated. The feature extraction process is necessary to produce a clustering that has a higher level of accuracy and can put the data in the right cluster.

The research methodology in this study start with main features selection of the internet bandwidth usage data using Principal Component Analysis (PCA), after the main feature is found then the next process is the extraction of feature data using the Correlation Feature Selection (CFS). After the extraction process is complete dataset will be incorporated into Fuzzy C-Mean Algorithm for diklasterisasi. Results of clustering the Internet bandwidth will be evaluated for accuracy, class and class precision and recall of the results of the evaluation will obtain the appropriate method to perform clustering of Internet bandwidth.

The results obtained from this study is that the method of extracting the main features of the selection process dengen using Principal Component Analysis (PCA) prior to feature extraction data using Correlation Feature Selection (CFS) can increase the number of classes identified in the clustering and also can improve the accuracy of clustering that generated. The level of accuracy produced by this method is 88.49% and can put the data into the correct cluster 7 of 11 available cluster, this method is better than just feature extraction data using Correlation Feature Selection (CFS), in which this method has an accuracy of 88.10 % and can only put the correct data into 5 clusters of 11 clusters available.

Kata Kunci : Traffic, Internet, Fuzzy C-Mean, Clustering, Extraction, Feature