

Clustering of Web Usage Data Using Enhanced K-means Clustering Algorithm

Hnin Yadana Than, May Tar Hla Myint

University of Computer Studies, Yangon

hninyadana8892@gmail.com

ABSTRACT

Clustering is the process of partitioning or grouping a given set of patterns into disjoint clusters. Clustering web usage data requires developing specialized techniques based on the web log data. The methodology has to improve the data preprocessing as well as the quality of the clusters. Traditional K-means algorithm is widely used clustering algorithm with wide range of application. Since K-means algorithm has some disadvantages. The proposed system intends to implement an enhanced K-means clustering algorithm to get more accurate and effective cluster results. The algorithm will carry out in two steps which include finding initial centroids and clustering stage. The testing of the system uses web log data from NASA web usage data for clustering. The performance and accuracy of the system are also measured.