Web Usage Mining Using Clustering andAssociationRule Mining

Aye Theingi Thwin, Nang Saing Moon Kham
University of Computer Studies, Yangon
aye.theingithwin.5cs96.20batch@gmail.com,
moonkhamucsy@gmail.com

Abstract

Data mining methods are used to discover the behaviour of the users. Therefore, the data used for the mining purpose must be qualified for the data cleaning stage and must be considered and planned efficiently to meet the requirement. For this reason, the data cleaning of the pre-processing stage becomes the essential key. Similarity measurement method is used to discover web usage data that have same category or usage purpose for clustering. Association rule mining uses the clustered data to generate rules that discover the patterns of interest.

This proposed system presents web usage mining using data mining methods. The main components that are included in this system are the preprocessing of web access log, computing similarity measurement using Jaccard coefficient, clustering the web pages using K-Mean Algorithm and finally the generation of rules for frequent pattern of web pages using Apriori Algorithm for interesting relationships among web pages in given web usage data set.