Title : Model-Based Cluster Analysis for Web Users Sessions

Roll No : 16PD22, 16PD36

Name : Nehal M, Subashini S

Abstract:
➤ One of the main issues in Web usage mining is the discovery of patterns in the navigational behavior of Web users.
➤ Standard approaches, such as clustering of user's sessions and discovering association rules or frequent navigational paths, do not generally allow to characterize or quantify the unobservable factors that lead to common navigational patterns.
➤ Therefore, it is necessary to develop techniques that can discover hidden and useful relationships among users as well as between users and Web objects.
➤ Correspondence Analysis(CO-AN) is particularly useful in this context, since it can uncover meaningful associations among users and pages.
➤ We present a model-based cluster analysis for Web users sessions including a novel visualization and interpretation approach which is based on CO-AN.
➤ Then we do a group wise analysis of page hits which can be used to understand the needs of each segment.
➤ Additionally we build a user network based on sessions and do community detection to identify similar users for targeted advertisements.
➤ Also a web-page network is built in order identify important pages in a website. Based on the TPM formed we find the probability of page transitions.

Link to Data, Code, Paper and Result:
https://drive.google.com/drive/folders/1w0iranKcYbIRqZZRUqGQ1JDExUNVYyXO?usp=sharing
To view the final output download the "**result**" folder from the above link and open the "**result.html**" page.