

## **Data Warehousing and Data Mining Lab**

### **Assignments No. 7 : Web Mining**

1. Implement the PageRank algorithm to calculate the rank of each page in the file. The output should be the 10 pages with the highest rank, together with their rank values.
  2. Implement the HITS algorithm to calculate the hub and the authority weight of each web page in the data set. The output should be the 10 most authoritative pages and 10 most hubby pages.
- Use Web Graph datasets from “Stanford Large Network Dataset Collection”  
<http://snap.stanford.edu/data/#web>
  - Tabulate the results.

**Deadline : 04/11/2019**

**Dr. B. F. Momin**  
**Subject Teacher**