

A Project Report On
IDENTIFICATION AND CLASSIFICATION OF THE
USERS ACCESS BEHAVIOURAL PATTERNS USING
WEB MINING TECHNIQUES

K S C S T

6881

submitted to



VISVESVARAYA TECHNOLOGICAL UNIVERSITY,
BELGAUM- 590014.

in partial fulfillment of the requirement for the award of the degree of
Bachelor of Engineering
in
Computer Science and Engineering
by

Abhisha Purushottam Tarimane	2KE08CS002
Annapurna M Sidnal	2KE08CS007
Deepa S Mulimani	2KE08CS015
Sachin Katarki	2KE08CS040

Under the guidance of

Prof. Sindhu. P. Menon
Asst. Professor, Dept of CSE, KLEIT, Hubli



K. L. E. S.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

K. L. E. SOCIETY'S
K. L. E. INSTITUTE OF TECHNOLOGY,
HUBLI-580030.
2011-2012

ABSTRACT

The World Wide Web is the largest and most widely known and used document repository, but in some aspects it has become a victim of its own success. World Wide Web has become a powerful platform to assimilate, disseminate and retrieve information as well as mine useful knowledge. Due to the properties of the huge, diverse, dynamic and unstructured nature of the Web, the Web data research has been facing a lot of challenges. As a result, Web users are always getting drowned in an “ocean” of information and face the problems of information overload while interacting with the Web. Here the issue of quality of service(QOS) crops up. To improve the internet service quality, it is, thus, necessary for a web developer to know what the user really wants to do, predict which pages the user is interested in and present the Web pages to the user by the knowledge of user navigational patterns.

This project mainly focuses on cleaning the data i.e. sever web log file, processing the data according to some specific strategy, identifying the users using maximal forward reference algorithms and classifying them into predefined classes. Here supervised learning is used to train the classifier. We have carried out this project using an educational institute’s log file as input data.