

Anekant Education Society's

**Tuljaram Chaturchand College
of
Arts , Science and Commerce**
(Autonomous)

Baramati

Department Of Computer Science


Ms. Kalyani Waman Londhe.
M.Sc. (Computer Science) NET




INTRODUCTION TO WEB MINING

Web Mining Definition

- Web mining is the application of data mining technique to extract useful knowledge and to find information patterns directly from Web data i.e. from World Wide Web.
- Web data is made up of web documents , services, text, images , audios, videos, records, hyperlinks, tags, http logs, app server logs, etc.

- 
- It is also used in finding pattern and knowledge from web contents, web structure, web usage and web dynamics
 - It is a complex method
 - Advanced and Useful data mining process
 - It is development of scalable and effective web data analysis and mining methods

- 
- Web mining is used by Data scientists and data analysts.
 - In Web Mining we have information in the form of structured, unstructured and semi-structured web pages.
 - Special tools for web mining are Scrapy, PageRank and Apache logs
 - Very challenging and fast evolving field in the data mining


Web Mining Tasks

1. Web Usage Mining
2. Web Content Mining
3. Web Structure Mining



- Web Usage Mining

- Web usage data includes http logs, app server logs and all log related data
- Also called as log mining
- Used for mining the web log records
- Used to discover user access patterns of web pages by mining web log records
- These patterns enable you to understand the user behaviors or something like that.
- By analyzing web log records, we can find out access to particular website, its access node, IP address etc.

- 
- **Web Content Mining**
 - Also known as text mining
 - Web page is made up of various content data
 - Web content data includes text, image, records, tables, List, audio, video etc.
 - Process of extracting useful information and pattern from web content data is web content mining.
 - Provide useful, effective and interesting patterns




- **Web Structure Mining**

- Web structure data includes hyperlinks, tags, etc.
- Structure mining basically gives the structured summary of a various websites.
- It makes use of graph theory
- Website is considered as graph where webpages are nodes and hyperlinks are considered as edges
- Very useful in determining the connection between web pages and websites.

Applications of Web Mining

- Used by various Search Engines like google, yahoo etc.
- Used in E-commerce for Customer Behavior Analysis
-
- Used for the Transaction Analysis
- Used in E-Banking
- Very useful for the Web Advertisement

- 
- Used to categorize web documents
 - Used in most of the E- Commerce fields like E-Learning, E-Governance,
 - Used in various Web Page Categorization techniques.
 - Used for Web Spam Filtering
 - To identify topics represented by web documents



Thank You