

DATA MINING AND WAREHOUSING

WEB MINING

DHANYAJA N

Assistant Professor
STM Kannur

INTROUCTION

- Introduction
- Web Content Mining
- Web Structure Mining
- Web Usage Mining.

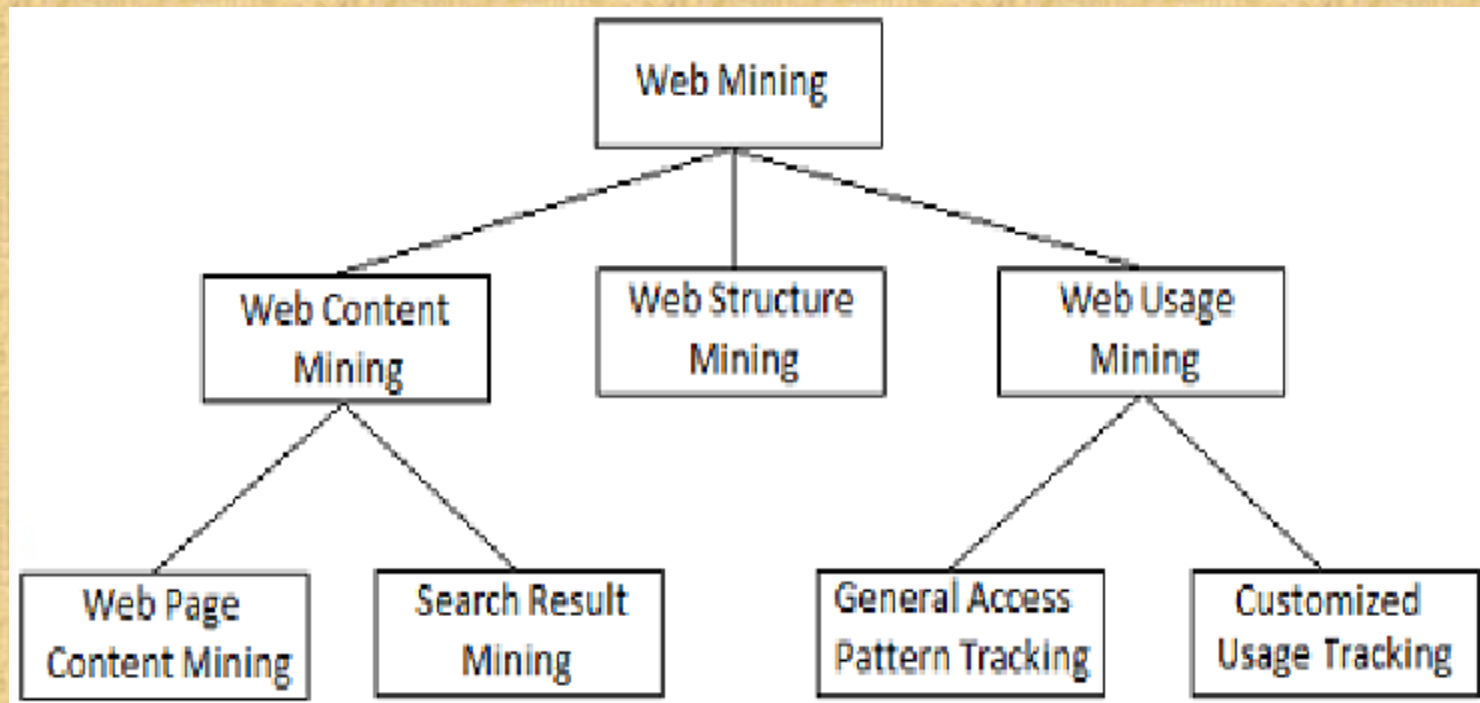
Web Mining

- **Web Mining** is the process of data mining techniques to automatically discover and extract information from Web documents and services.
- The main purpose of web mining is discovering useful information from the World-Wide Web and its usage patterns.

Web Mining

- Web mining extracts knowledge from web data.
- Web Data is,
 - 1.web content- text, image,record etc.
 - 2.web structure-hyperlinks,tags etc

Types: Web Mining



Web Content Mining

- Web Content Mining is the process of extracting useful information from the contents of Web documents.
- Content data corresponds to the collection of facts a Web page was designed to convey to the users.
- It may consist of text, images, audio, video, or structured records such as lists and tables.

Web structure mining

- Web structure mining is the application of discovering structure information from the web.
- The structure of the web graph consists of web pages as nodes, and hyperlinks as edges connecting related pages.
- Structure mining basically shows the structured summary of a particular website.

Web structure mining

- It identifies relationship between web pages linked by information or direct link connection.
- To determine the connection between two commercial websites, Web structure mining can be very useful.
- This type of mining can be performed either at the (intra-page) document level or at the (inter-page) hyperlink level.

Web-graph

- Web-graph: A directed graph that represents the Web.
-
- Node: Each Web page is a node of the Web- graphLink:
- Each hyperlink on the Web is a directed edge of the Web-graph.

Web Usage Mining

- Web usage mining is the process of extracting useful information from server logs. i.e. User's history.
- Web usage mining is the process of finding out what users are looking on internet.
- User identification, session creation, robot detection. and filtering, and extracting usage path patterns.

•

Applications

- Web mining helps to improve the power of web search engine by classifying the web documents and identifying the web pages.
- It is used for Web Searching e.g., Google, Yahoo etc and Vertical Searching
- Web mining is used to predict user behavior.
- Web mining is very useful of a particular Website and e-service e.g., landing page optimization.

Challenges in Web Mining

- The web is too huge
- Complexity of Web pages
- Web is dynamic information source
- Diversity of user communities
- Relevancy of Information

Challenges in Web Mining

- The web is too huge
- Complexity of Web pages
- Web is dynamic information source
- Diversity of user communities
- Relevancy of Information

Thank You !!!

DHANYAJA N

Assistant Professor
STM Kannur