Unit-I INTRODUCTION TO WEB DATA MINING

* Need

* Importance

* Applications of Web Data mining

* Capturing-users web

* activities

* Client side v/s middleware v/s server side-data and usage logging

* Web

* Mining and its types

* Web Usage Mining

* Web Structure Mining

* Web Content

* Mining

Unit-II WEB USAGE MINING

* Learning from Browser

* Server Logs

* Identifying frequent item sets

* pattern

* identification

* representing patterns in form of relations/Graphs

* Understanding

* web application or website- Usage

* Heat maps

* Using statistical tools for usage

* analysis and machine learning for prospective improvements

*

Unit-III WEB STRUCTURE MINING

* Understanding link structure of the web

* Static v/s dynamic linking

* representing

* the link structure asgraphs

* identifying most / least used links

* paths

* Categorizing

* links based on required attributes

* Clustering links based on required attributes

* Web as a graph

* identifying nodes

* edges

* in-degree

* out- degree

* HITS Algorithm

* Page Rank algorithm

Unit-IV WEB CONTENT MINING

* Storing web content as text

* database

* various document types

* generating meta-

* information of web documents

* labelling

* tagging

* identifying feature sets

* Representing web documents

* Vector Space Model

* TF-IDF

* web-page

* summarization

* tokenization

* n-gram analysis

* categorizing web pages based on

* required attributes

* Clustering web pages based on required attributes