Unit-8

Advance topics

Outline

- Spatial Mining
- Web Mining
- Text Mining
- Temporal Mining
- Multimedia Mining
- Information privacy and data mining

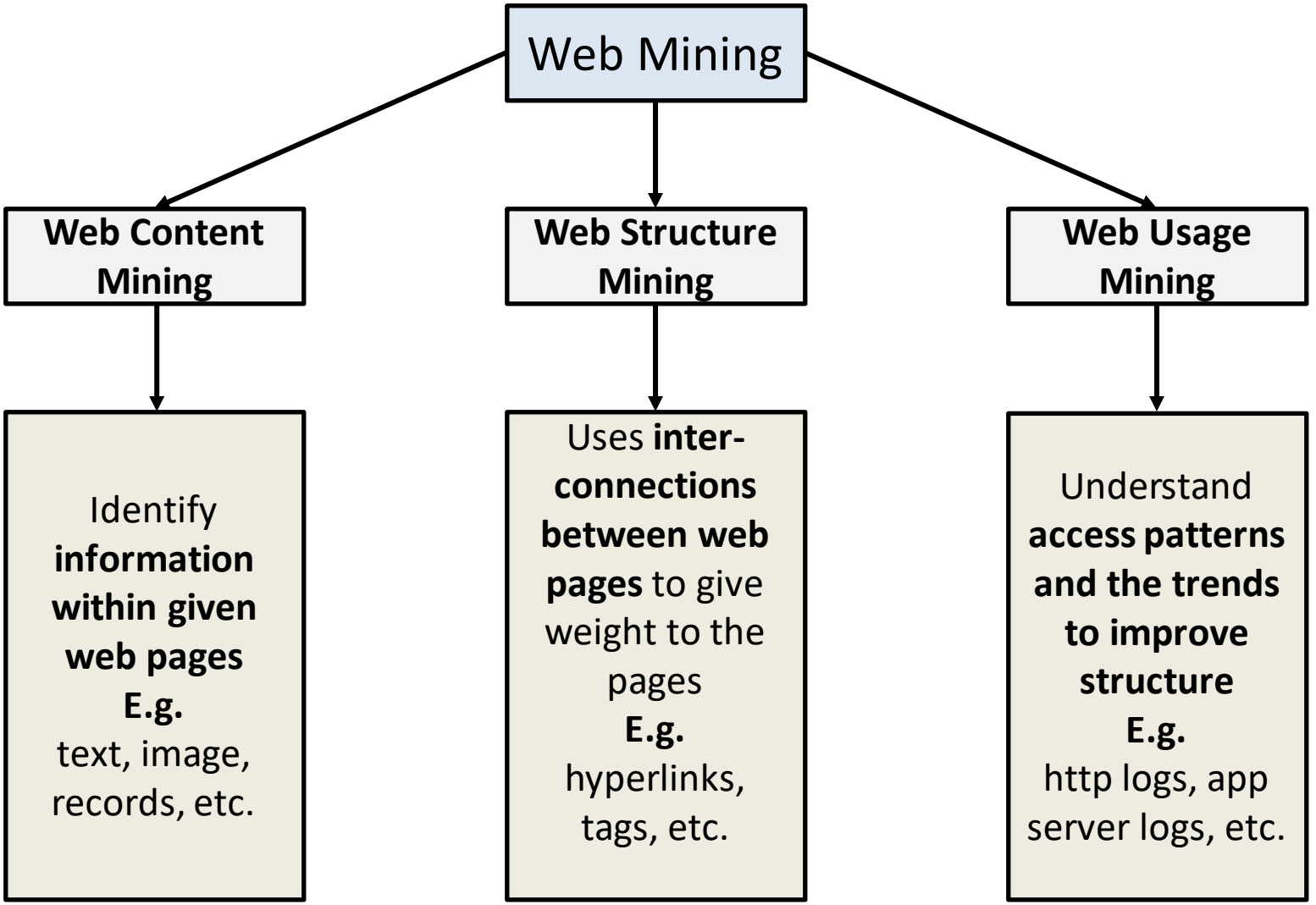
Spatial mining

- Spatial data mining is the application of **data mining** to **spatial models**.
- Spatial data mining is based on **geographical analysis.**
- In spatial data mining, **analysts use geographical or spatial information to produce business intelligence** or other results.
- It requires **specific techniques and resources to get the geographical data** into relevant and useful formats.
- The task is to search for **spatial patterns**.

Web mining

- Web mining is the use of data mining techniques to automatically discover and **extract information from web documents** and services.
- There are general classes of information that can be discovered in web mining: **web activity, from server logs** and **web browser activity tracking**.
- Web mining can be broadly divided into three categories, according to the kinds of data to be mined.
  - **Web content Mining**
  - **Web structure Mining**
  - **Web usage Mining**

Web mining (Cont..)



Text mining

- **Text mining**, also referred to as **text data mining**, roughly equivalent to text analytics, is the process of deriving **high-quality information from text**.
- With the advancement of technology, more and more data is available in digital form, among them, **most of the data** (approx. 85%) is in **unstructured textual form**.
- Compared with the kind of data stored in databases, text is unstructured, ambiguous, and difficult to process.
- Nevertheless, in modern culture, **text** is the **most communal way for the formal exchange of information**.
- It has become essential to develop better techniques and algorithms to **extract useful and interesting information** from this large amount of **textual data**.

Temporal Mining

- Temporal data mining defines the process of extraction of non-trivial, implicit, and potentially essential data from large sets of temporal data.
- Temporal data are a series of primary data types, generally numerical values, and it deals with gathering beneficial knowledge from temporal data.

Temporal Mining

- Temporal data mining is composed of three major works such as the description of temporal data, representation of similarity measures, and mining services.
- Temporal Data Mining includes processing time series, generally sequences of data, which compute values of the same attribute at a sequence of multiple time points.
- Pattern matching using such information, where it is searching for specific patterns of interest, has attracted considerable interest in current years.

Temporal Mining

- Temporal Data Mining can include the exploitation of efficient techniques of **data storage, quick processing, and quick retrieval methods that have been advanced for temporal databases**.

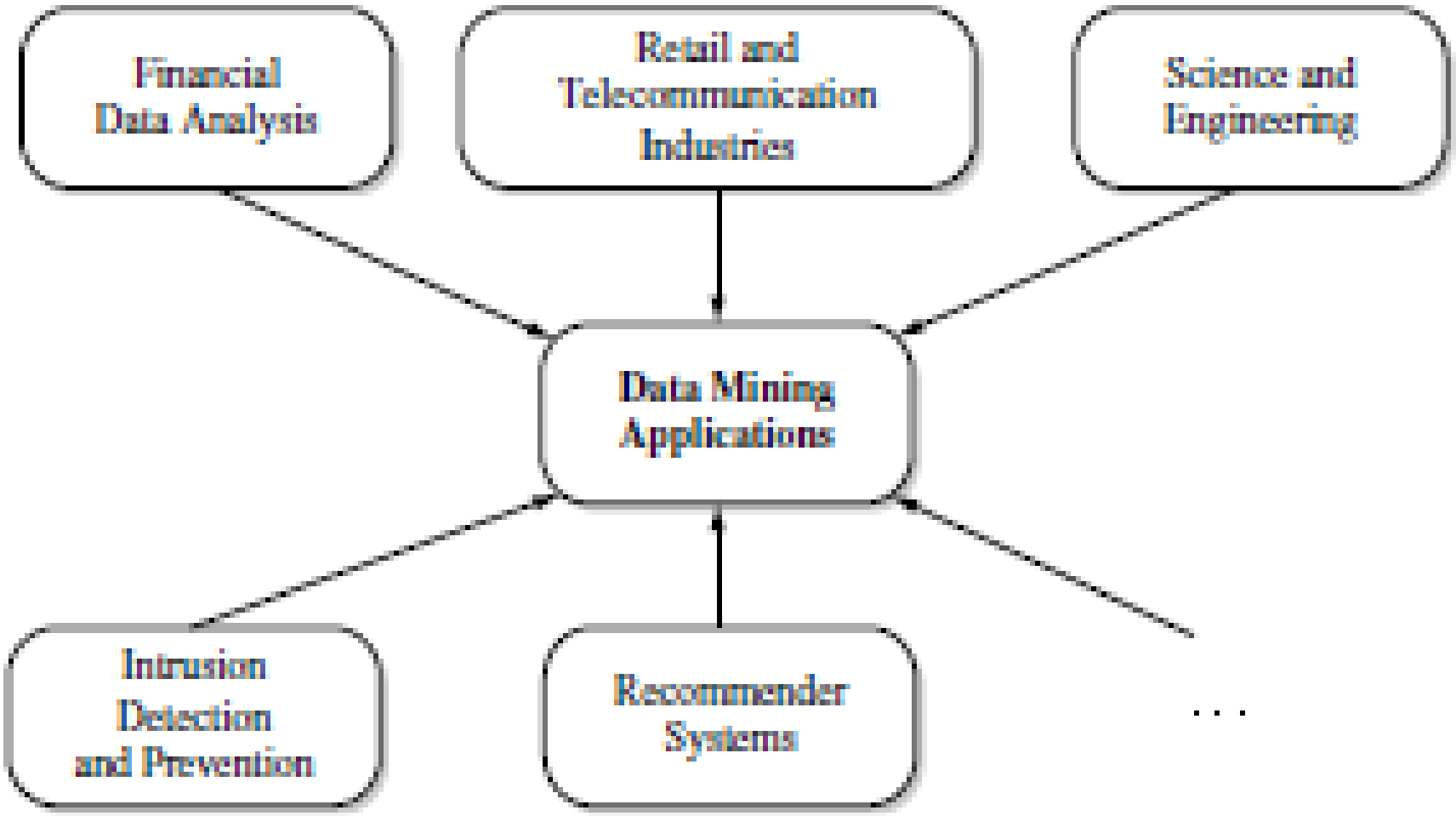
Multimedia Mining

- Multimedia mining is a subfield of data mining that is used to find interesting information of implicit knowledge from multimedia databases.
- Mining in multimedia is referred to as automatic annotation or annotation mining.
- Mining multimedia data requires two or more data types, such as text and video or text video and audio.

Multimedia Mining

- Multimedia data mining is an interdisciplinary field that integrates image processing and understanding, computer vision, data mining, and pattern recognition.
- Multimedia data mining discovers interesting patterns from multimedia databases that store and manage large collections of multimedia objects, including image data, video data, audio data, sequence data and hypertext data containing text, text markups, and linkages.

Application of DM:



Banking & Finance

- The financial data in the banking and financial industry is generally reliable and of high quality which facilitates systematic data analysis and data mining.
- Some of the typical cases are as follows:
  - ✓ Design and construction of data warehouses for multidimensional data analysis and data mining.
  - ✓ Loan payment prediction and customer credit policy analysis.
  - ✓ Classification and clustering of customers for targeted marketing.
  - ✓ Detection of money laundering and other financial crimes.

Retail and Telecommunication Industries

- The retail industry is a well-fit application area for data mining, since it collects huge amounts of data on sales, customer shopping history, goods transportation, consumption, and service.
- The quantity of data collected continues to expand rapidly, especially due to the increasing availability, ease, and popularity of business conducted on the Web, or e-commerce.

Retail and Telecommunication Industries

- Retail data mining can help identify customer buying behaviours, discover customer shopping patterns and trends, improve the quality of customer service, achieve better customer retention and satisfaction, enhance goods consumption ratios, design more effective goods transportation and distribution policies, and reduce the cost of business.

Retail and Telecommunication Industries

- A few examples of data mining in the retail industry are outlined as follows:
- Design and construction of data warehouses
- Multidimensional analysis of sales, customers, products, time, and region
- Analysis of the effectiveness of sales campaigns
- Customer retention—analysis of customer loyalty

Thank you!