**DATA MINING APPLICATIONS**

Here is the list of areas where data mining is widely used −

* Financial Data Analysis
* Retail Industry
* Telecommunication Industry
* Biological Data Analysis
* Other Scientific Applications
* Intrusion Detection

Financial Data Analysis

The financial data in 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 Industry

Data Mining has its great application in Retail Industry because it collects large amount of data from on sales, customer purchasing history, goods transportation, consumption and services. It is natural that the quantity of data collected will continue to expand rapidly because of the increasing ease, availability and popularity of the web.

Data mining in retail industry helps in identifying customer buying patterns and trends that lead to improved quality of customer service and good customer retention and satisfaction. Here is the list of examples of data mining in the retail industry −

* Design and Construction of data warehouses based on the benefits of data mining.
* Multidimensional analysis of sales, customers, products, time and region.
* Analysis of effectiveness of sales campaigns.
* Customer Retention.
* Product recommendation and cross-referencing of items.

Telecommunication Industry

Today the telecommunication industry is one of the most emerging industries providing various services such as fax, pager, cellular phone, internet messenger, images, e-mail, web data transmission, etc. Due to the development of new computer and communication technologies, the telecommunication industry is rapidly expanding. This is the reason why data mining is become very important to help and understand the business.

Data mining in telecommunication industry helps in identifying the telecommunication patterns, catch fraudulent activities, make better use of resource, and improve quality of service. Here is the list of examples for which data mining improves telecommunication services −

* Multidimensional Analysis of Telecommunication data.
* Fraudulent pattern analysis.
* Identification of unusual patterns.
* Multidimensional association and sequential patterns analysis.
* Mobile Telecommunication services.
* Use of visualization tools in telecommunication data analysis.

Biological Data Analysis

In recent times, we have seen a tremendous growth in the field of biology such as genomics, proteomics, functional Genomics and biomedical research. Biological data mining is a very important part of Bioinformatics. Following are the aspects in which data mining contributes for biological data analysis −

* Semantic integration of heterogeneous, distributed genomic and proteomic databases.
* Alignment, indexing, similarity search and comparative analysis multiple nucleotide sequences.
* Discovery of structural patterns and analysis of genetic networks and protein pathways.
* Association and path analysis.
* Visualization tools in genetic data analysis.

Other Scientific Applications

The applications discussed above tend to handle relatively small and homogeneous data sets for which the statistical techniques are appropriate. Huge amount of data have been collected from scientific domains such as geosciences, astronomy, etc. A large amount of data sets is being generated because of the fast numerical simulations in various fields such as climate and ecosystem modeling, chemical engineering, fluid dynamics, etc. Following are the applications of data mining in the field of Scientific Applications −

* Data Warehouses and data preprocessing.
* Graph-based mining.
* Visualization and domain specific knowledge.

Intrusion Detection

Intrusion refers to any kind of action that threatens integrity, confidentiality, or the availability of network resources. In this world of connectivity, security has become the major issue. With increased usage of internet and availability of the tools and tricks for intruding and attacking network prompted intrusion detection to become a critical component of network administration. Here is the list of areas in which data mining technology may be applied for intrusion detection −

* Development of data mining algorithm for intrusion detection.
* Association and correlation analysis, aggregation to help select and build discriminating attributes.
* Analysis of Stream data.
* Distributed data mining.
* Visualization and query tools

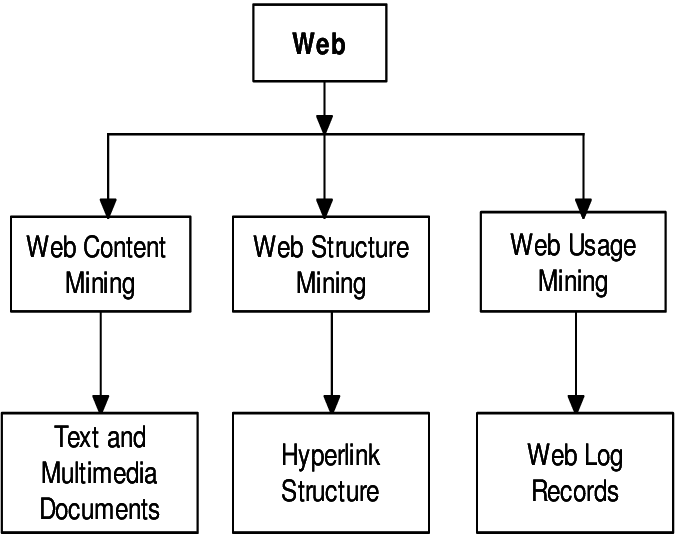
**WEB MINING:**

**Web Mining** is the process of [Data Mining](https://www.geeksforgeeks.org/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.

**Applications of Web Mining:**

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

Web mining can be broadly divided into three different types of techniques of mining: Web Content Mining, Web Structure Mining, and Web Usage Mining. These are explained as following below.



1. **Web Content Mining:**  
   Web content mining is the application of extracting useful information from the content of the web documents. Web content consist of several types of data – text, image, audio, video etc. Content data is the group of facts that a web page is designed. It can provide effective and interesting patterns about user needs. Text documents are related to text mining, machine learning and natural language processing. This mining is also known as text mining. This type of mining performs scanning and mining of the text, images and groups of web pages according to the content of the input.
2. **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. 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.
3. **Web Usage Mining:**  
   Web usage mining is the application of identifying or discovering interesting usage patterns from large data sets. And these patterns enable you to understand the user behaviors or something like that. In web usage mining, user access data on the web and collect data in form of logs. So, Web usage mining is also called log mining.

**Comparison Between Data mining and Web mining:**

|  |  |  |
| --- | --- | --- |
| **POINTS** | **DATA MINING** | **WEB MINING** |
| Definition | Data Mining is the process that attempts to discover pattern and hidden knowledge in large data sets in any system. | Web Mining is the process of data mining techniques to automatically discover and extract information from web documents. |
| Application | Data Mining is very useful for web page analysis. | Web Mining is very useful for a particular website and e-service. |
| Target Users | Data scientist and data engineers. | Data scientists along with data analysts. |
| Access | Data Mining is access data privately. | Web Mining is access data publicly. |
| Structure | In Data Mining get the information from explicit structure. | In Web Mining get the information from structured, unstructured and semi-structured web pages. |
| Problem Type | Clustering, classification, regression, prediction, optimization and control. | Web content mining, Web structure mining. |
| Tools | It includes tools like machine learning algorithms. | Special tools for web mining are Scrapy, PageRank and Apache logs. |
| Skills | It includes approaches for data cleansing, machine learning algorithms. Statistics and probability. | It includes application level knowledge, data engineering with mathematical modules like statistics and probability. |

**Spatial Data Mining :**

Spatial data mining is the process of discovering interesting and previously unknown, but potentially useful patterns from spatial databases. In spatial data mining analyst use geographical or spatial information to produce business intelligence or other results. Challenges involved in spatial data mining include identifying patterns or finding objects that are relevant to research project.

Spatial data is associated with geographic locations such as cities,towns etc. A spatial database is optimized to store and query data representing objects. These are the objects which are defined in a geometric space.

## Characteristics of Spatial Database

A spatial database system has the following characteristics

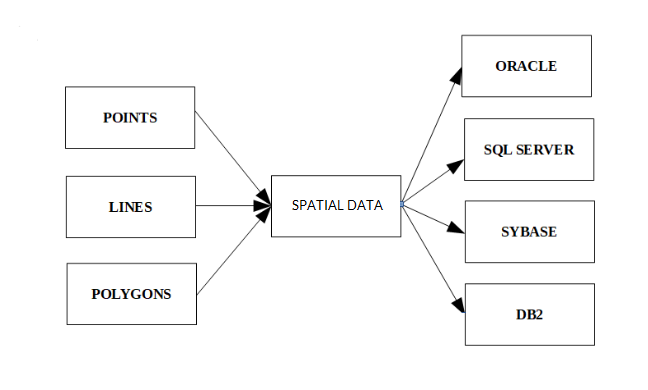
* It is a database system
* It offers spatial data types (SDTs) in its data model and query language.
* It supports spatial data types in its implementation, providing at least spatial indexing and efficient algorithms for spatial join.

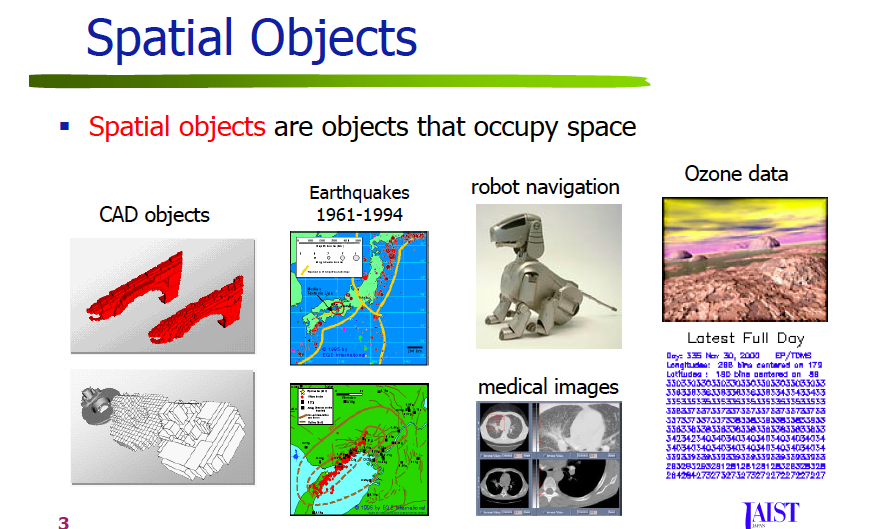
## Example

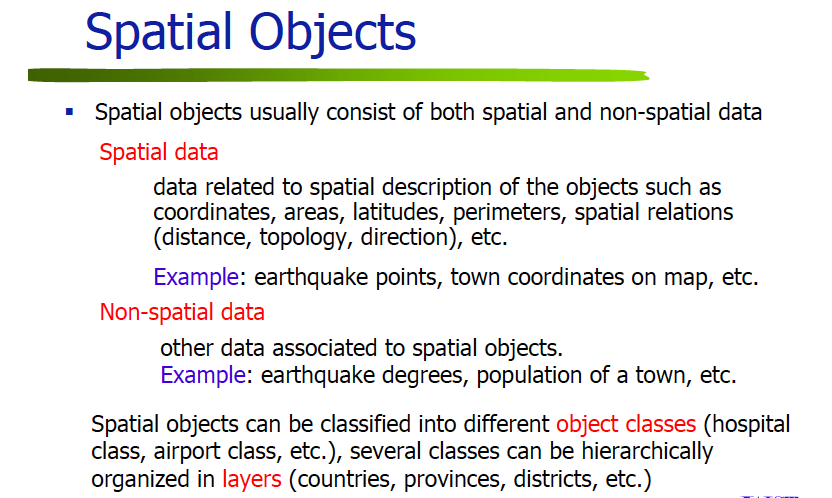
A road map is a visualization of geographic information. A road map is a 2-dimensional object which contains points, lines, and polygons that can represent cities, roads, and political boundaries such as states or provinces.

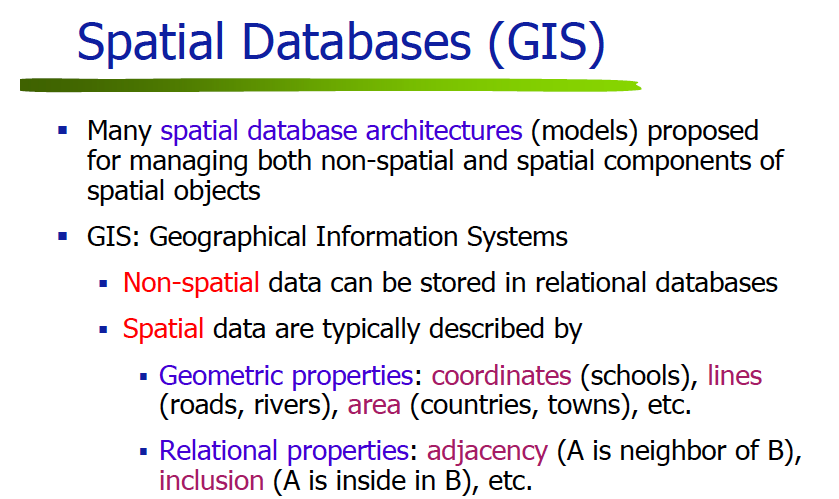
**In general, spatial data can be of two types −**

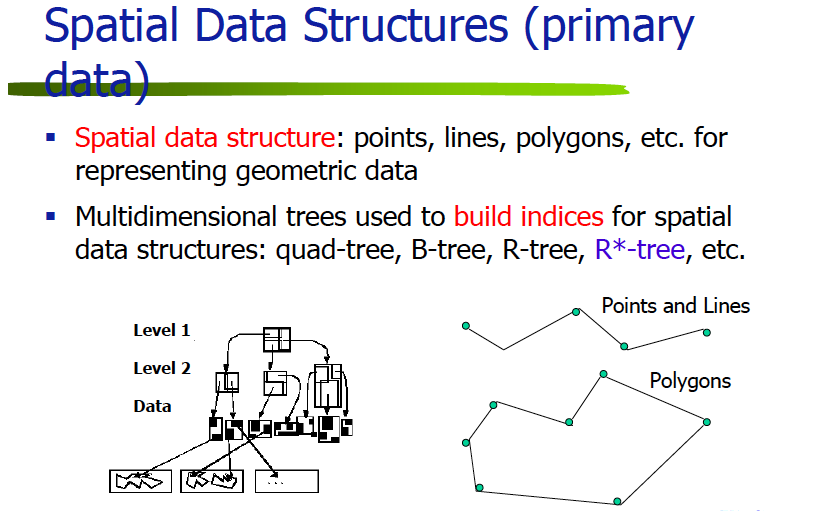
* Vector data: This data is represented as discrete points, lines and polygons
* Rastor data: This data is represented as a matrix of square cells.

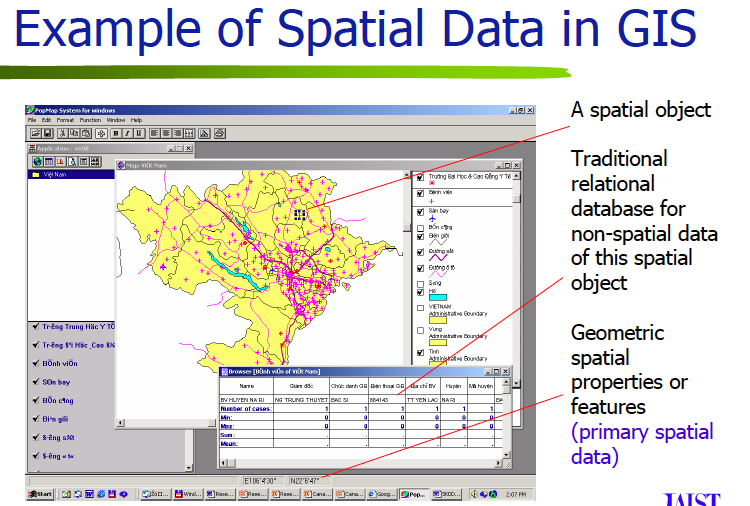
****

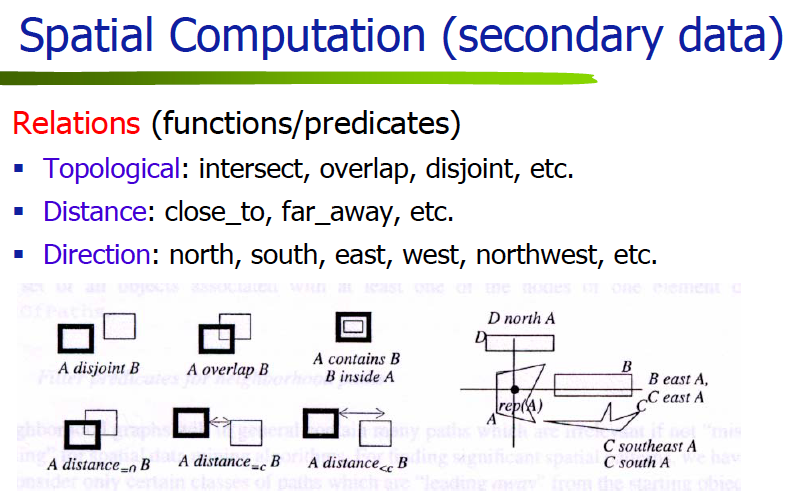
****

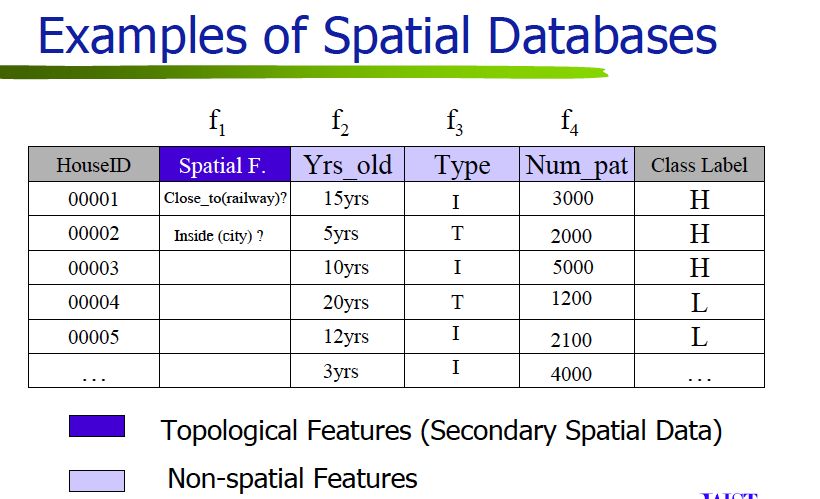
****

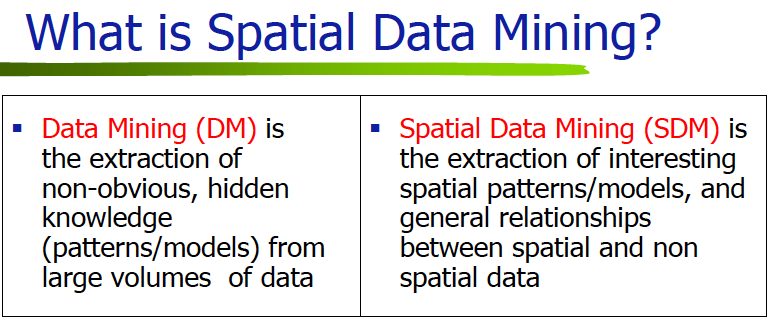
****

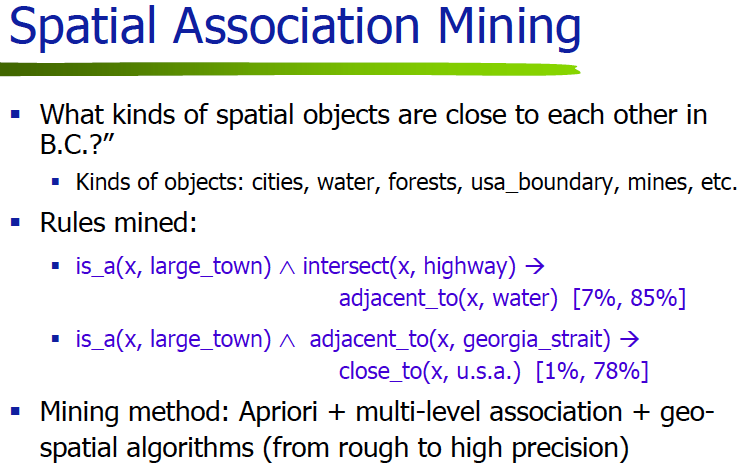
****

****

****

****

****

****

**Temporal Data Mining :**

Temporal data refers to the extraction of implicit, non-trivial and potentially useful abstract information from large collection of temporal data. It is concerned with the analysis of temporal data and for finding temporal patterns and regularities in sets of temporal data. Temporal reasoning is added by storing separate snapshots of the rule set over time which are then compared to draw conclusions regarding the change in data over time. This technique could be applied to any non-temporal database to allow some temporal reasoning.

**Examples of temporal rules:**

* Some patients tend to develop reactions after two months with this combination of drugs.
* There is a higher incidence of earthquakes during and soon after periods of higher atmospheric pressure;
* The sequence Committee ! Board ! Council occurs approximately every month;

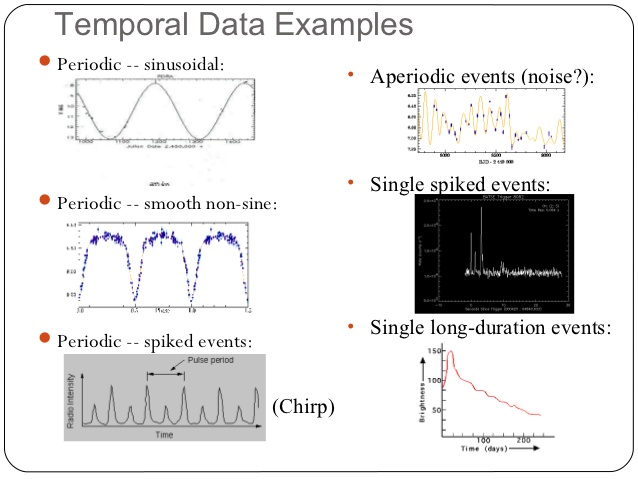
**Types of Temporal Data:**

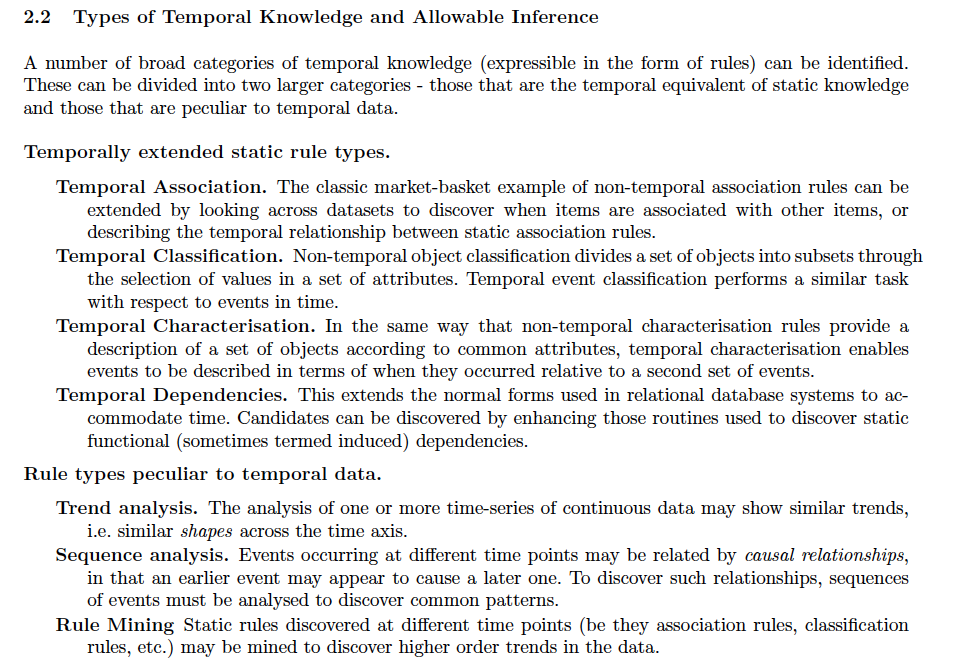
**Static.** No temporal context is included and none can be inferred. Occasionally, some temporal inference can be made through reference to transaction-time by referring to audit trails or transaction logs.

**Sequences.** Ordered sequences of events. The category would include ordered, but not timed, collections of events such as a series of market-basket transactions. While most collections are often limited to the sequence relationships before and after, this category also includes the richer relationships described by Allen and others [9, 26, 67] such as meets, overlaps, contemporary of, etc.

**Timestamped.** A timed sequence of static datasets taken at more or less regular intervals. Examples include census and satellite meteorological data.

**Fully temporal.** Each tuple in a time-varying relation in the database has the capacity to have one or more dimensions of time, such as either or both of a transaction-time or valid-time history5. In practice, the valid-time histories are likely to be the more common form of source data of interest to the user.





**Tasks of temporal data mining are –**

* **Data Characterization and Comparison**
* **Cluster Analysis**
* **Classification**
* **Association rules**
* **Prediction and Trend Analysis**
* **Pattern Analysis**

**For more details: https://www.slideshare.net/patilbharat7/temporal-data-mining**

**Difference between Spatial and Temporal Data Mining :**

|  |  |  |
| --- | --- | --- |
| **SNO.** | **Spatial data mining** | **Temporal data mining** |
| **1.** | **It requires space.** | **It requires time.** |
| **2.** | **Spatial mining is the extraction of knowledge/spatial relationship and interesting measures that are not explicitly stored in spatial database.** | **Temporal mining is the extraction of knowledge about occurrence of an event whether they follow Cyclic , Random ,Seasonal variations etc.** |
| **3.** | **It deals with spatial (location , Geo-referenced) data.** | **It deals with implicit or explicit Temporal content , from large quantities of data.** |
| **4.** | **Spatial databases reverses spatial objects derived by spatial data. types and spatial association among such objects.** | **Temporal data mining comprises the subject as well as its utilization in modification of fields.** |
| **5.** | **It includes finding characteristic rules, discriminant rules, association rules and evaluation rules etc.** | **It aims at mining new and unknown knowledge, which takes into account the temporal aspects of data.** |
| **6.** | **It is the method of identifying unusual and unexplored data but useful models from spatial databases.** | **It deals with useful knowledge from temporal data.** |
| **7.** | **Examples –**  **Determining hotspots , Unusual locations.** | **Examples –**  **An association rule which looks like – “Any Person who buys a car also buys steering lock”. By temporal aspect this rule would be – ” Any person who buys a car also buys a steering lock after that “.** |