****

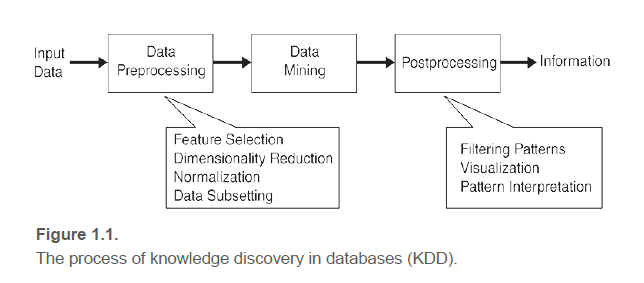
**Ch 1 : Introduction**

* customer profiling
* targeted marketing
* workflow management
* store layout
* fraud detection
* automated buying and selling

Data Mining Techniques : scour large sets of data to find novel and useful patterns

**KDD – Knowledge discovery in databases :** overall process of converting raw data into useful information

Data preprocessing -> data postprocessing



**Closing the loop :** integrating the results of data mining into decision support systems

Example of postprocessing is visualization.

**Factors motivating the need of advanced data mining techniques :**

1. Scalability
2. High Dimensionality
3. Heterogenous and Complex Data
4. Data Ownership and Distribution
5. Non-traditional Analysis

* Web mining
* Natural language processing (NLP)
* Network analysis
* Data mining
* Statistics
* Deep learning

**Data Mining Tasks :**

**(2 major categories)**

1. Predictive Tasks
2. Descriptive Tasks

Four of the core data mining tasks :

1. Predictive Modelling
2. Classification – for discrete target variables
3. Regression – for continuous target variables
4. Association Analysis – patterns
5. Cluster Analysis – group of closely related observations
6. Anomaly detection :

* Detection of fraud
* Network intrusions
* Unusual patterns of disease
* Ecosystem disturbance