**Page-1**

**Data Mining in a Nutshell:**



Page-2

**Types of Methods**

Supervised Learning (Predictive Analytics)

• Prediction (numerical Y)

• Classification (categorical Y)

Unsupervised Learning

• Segmentation/Clustering

• Dimension Reduction

• Relationship Mining

• Recommendation System

• Network Analysis

**Page-3**

**The Data Mining Process**

Define purpose > Obtain data > Explore & Clean data > Determine DM task > Choose DM Methods >

Apply methods, Select final model > Evaluate Performance > Deploy

Data Preparation & Exploration (2-3)

1) Sampling 2) cleaning 3) summaries 4) visualization 5) Partitioning 6) Dimension reduction

* Prediction

1. MLR 2) k-nearest neighbor 3) Regression tress

4) Neural nets

* Classification

1) K-nearest neighbor 2) Naïve Bayes 3) Logistic regression

4) Classification trees 5) Neural nets

6) Discrimination analysis

* Network analysis
* Segmentation clustering
* Affinity analysis/Association rules
* Model evaluation & selection
* Deriving insights