Lecture 1
Week 1 (March 10)

33459-01 Principles of Knowledge Discovery in Data

Introduction to the course and Introduction to KDD

Lecture by: Dr. Osmar R. Zaïane

Who Am I?

Database Laboratory

오스마 자이안

蔡頤安

UNIVERSITY OF
ALBERTA

Osmar R. Zaïane, Ph.D.
Associate Professor
Department of Computing Science

221 Athabasca Hall
Edmonton, Alberta
Canada T6G 2E8

Telephone: Office +1 (780) 492 2860
Fax +1 (780) 492 1071
E-mail: zaiane@cs.ualberta.ca
http://www.cs.ualberta.ca/~zaiane/

PhD on *Web Mining & Multimedia Mining* With **Dr. Jiawei Han** at Simon Fraser University, Canada

Research Interests:
Data Mining,
Web Mining,
Multimedia Mining,
Data Visualization,
Information Retrieval.

Applications:
Analytic Tools,
Adaptive Systems,
Intelligent Systems,
Diagnostic and Categorization,
Recommender Systems

Achievements:
(in last 6 years):
1 PhD and 16 MSc,
79 publications,
WEBKDD and MDM/KDD co-chair (2000 to 2003)
Currently: 5 PhD and 2 MSc students

Principles of Knowledge Discovery in Data

Class and Office Hours

Class:
Fridays from 13:00 to 16:00

Office Hours:
Thursdays from 14:00 to 15:00

Course Requirements

- Understand the basic concepts of database systems
- Understand the basic concepts of artificial intelligence and machine learning
- Be able to develop applications in C/C^{++} or Java

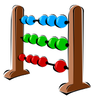
Course Objectives

To provide an introduction to knowledge discovery in databases and complex data repositories, and to present basic concepts relevant to real data mining applications, as well as reveal important research issues germane to the knowledge discovery domain and advanced mining applications.

Students will understand the fundamental concepts underlying knowledge discovery in databases and gain hands-on experience with implementation of some data mining algorithms applied to real world cases.

Evaluation and Grading

There is a midterm and a final exam for this course. There are also assignments and a small project.
I will be evaluating all these activities out of 100% and give a final grade based on the evaluation of the activities.



Assignments (2)	10%
Midterm	25%
Final Exam	40%
Project	15%

Tentative

- Project
  - Quality of presentation + quality of report + quality of demos
  - Project demo (June 16th)

- A+ will be given only for outstanding achievement.

More About Evaluation

Re-examination.

None, except as per regulation.

Collaboration.

Collaborate on assignments and projects, etc; do not merely copy.

Plagiarism.

Work submitted by a student that is the work of another student or any other person is considered plagiarism. Cases of plagiarism are immediately referred to the Dean, who determines what course of action is appropriate.

Projects

	Choice	Deliverables
	Implement data mining project	final demo + short project report Implementations: C/C++ or Java, OS: Linux, Window XP/2000 , or other systems.

Project itself will be determined at a later date

Assignments

1- Competition in one algorithm implementation
2- Use of educational DM tool to evaluate algorithms

Course Schedule (Tentative, subject to changes)

There are 15 weeks from March 10th to June 16th.

Week 1:	March 10: Introduction to Data Mining
Week 2:	March 17: Association Rules
Week 3:	March 24: Association Rules (advanced topics)
Week 4:	March 31: Sequential Pattern Analysis
Week 5:	April 7 : Classification (Neural Networks)
Week 6:	April 14 : Classification (Decision Trees and others)
Week 7:	April 21 : Midterm
Week 8:	April 28 : Data Clustering
Week 9:	May 5 : No class (Public Holiday – Children's day)
Week 10:	May 12 : Clustering Analysis and Outlier Detection
Week 11:	May 19 : Contrast sets
Week 12:	May 26 : Web Mining
Week 13:	June 2 : Web Mining
Week 14:	June 9 : Final Exam
Week 15:	June 16 : Project Demos

Away (out of town)
To be confirmed

Week 6 and week 12

Assignment Due dates
-Assignment 1
week 8
-Assignment 2
week 11

Course Content

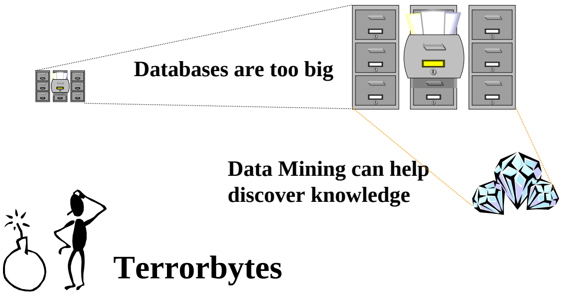
- Introduction to Data Mining
- Association analysis
- Sequential Pattern Analysis
- Classification and prediction
- Contrast Sets
- Data Clustering
- Outlier Detection
- Web Mining

Lecture 1 Objectives

Get a rough initial idea what knowledge discovery in data and data mining are.

Get an overview about the functionalities and the issues in data mining.

We Are Data Rich but Information Poor



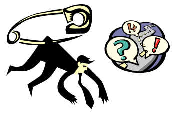
Databases are too big

Data Mining can help discover knowledge

Terrorbytes

What Should We Do?

We are not trying to find the needle in the haystack because DBMSs know how to do that.



We are merely trying to understand the consequences of the presence of the needle, if it exists.

What Led Us To This?

Necessity is the Mother of Invention

- Technology is available to help us collect data
  - Bar code, scanners, satellites, cameras, etc.
- Technology is available to help us store data
  - Databases, data warehouses, variety of repositories…
- We are starving for knowledge (competitive edge, research, etc.)

We are swamped by data that continuously pours on us.

1. We do not know what to do with this data
2. We need to interpret this data in search for new knowledge

Evolution of Database Technology

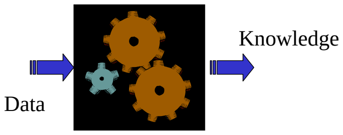
- **1950s**: First computers, use of computers for census
- **1960s**: Data collection, database creation (hierarchical and network models)
- **1970s**: Relational data model, relational DBMS implementation.
- **1980s**: Ubiquitous RDBMS, advanced data models (extended-relational, OO, deductive, etc.) and application-oriented DBMS (spatial, scientific, engineering, etc.).
- **1990s**: Data mining and data warehousing, massive media digitization, multimedia databases, and Web technology.

Notice that storage prices have consistently decreased in the last decades

What Is Our Need?

Extract interesting knowledge

(rules, regularities, patterns, constraints) from data in large collections.



A Brief History of Data Mining Research

- 1989 IJCAI Workshop on Knowledge Discovery in Databases (Piatetsky-Shapiro)

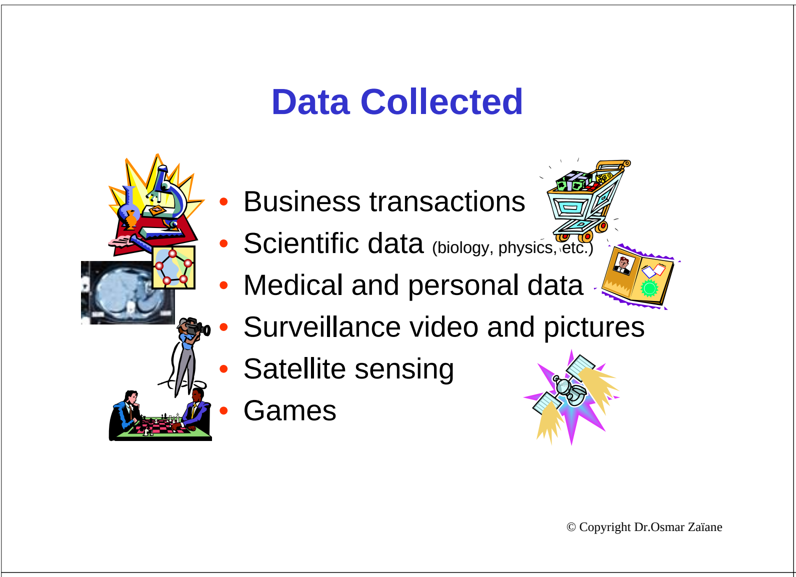
  Knowledge Discovery in Databases (G. Piatetsky-Shapiro and W. Frawley, 1991)
- 1991-1994 Workshops on Knowledge Discovery in Databases

  Advances in Knowledge Discovery and Data Mining (U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, 1996)
- 1995-1998 International Conferences on Knowledge Discovery in Databases and Data Mining (KDD'95-98)

  Journal of Data Mining and Knowledge Discovery (1997)
- 1998-2005 ACM SIGKDD conferences

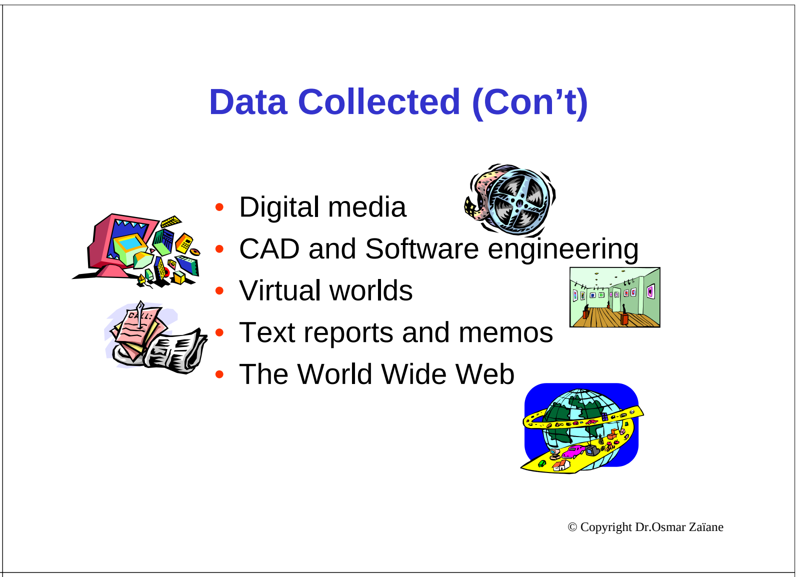
Introduction - Outline

- What kind of information are we collecting?
- What are Data Mining and Knowledge Discovery?
- What kind of data can be mined?
- What can be discovered?
- Is all that is discovered interesting and useful?
- How do we categorize data mining systems?
- What are the issues in Data Mining?
- Are there application examples?



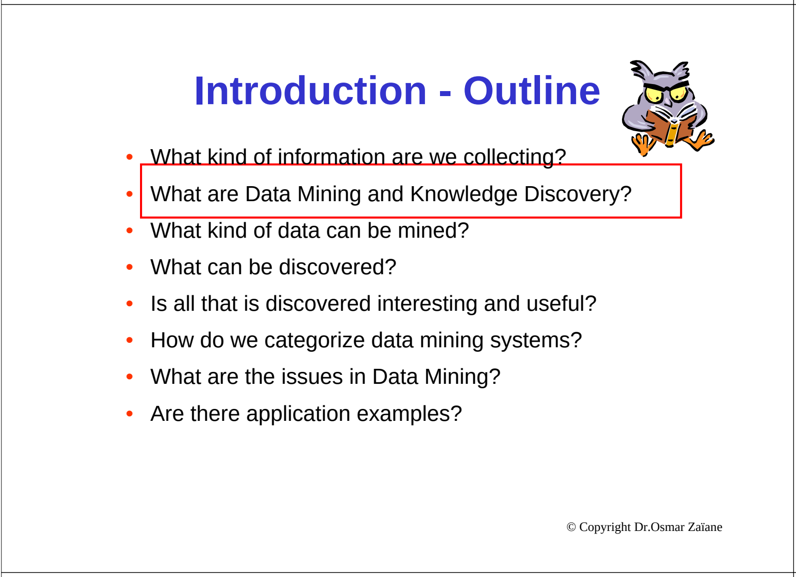
Data Collected

- Business transactions
- Scientific data (biology, physics, etc.)
- Medical and personal data
- Surveillance video and pictures
- Satellite sensing
- Games



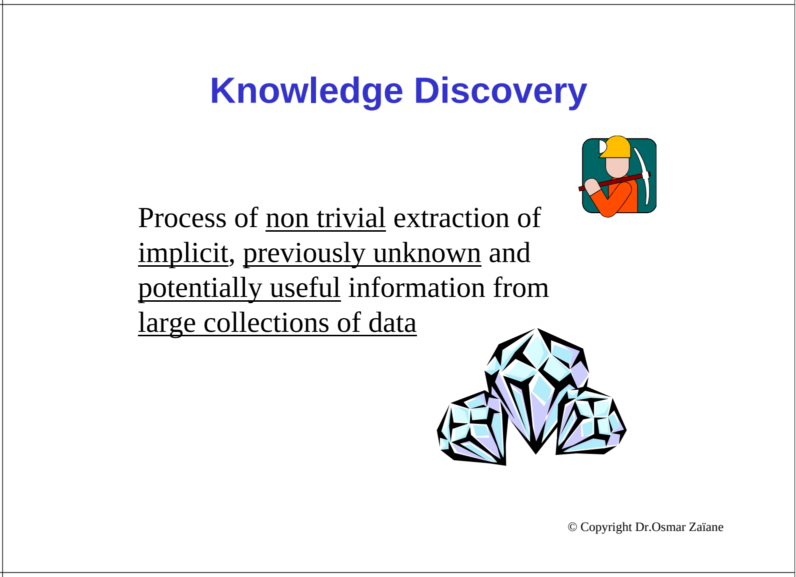
Data Collected (Con't)

- Digital media
- CAD and Software engineering
- Virtual worlds
- Text reports and memos
- The World Wide Web



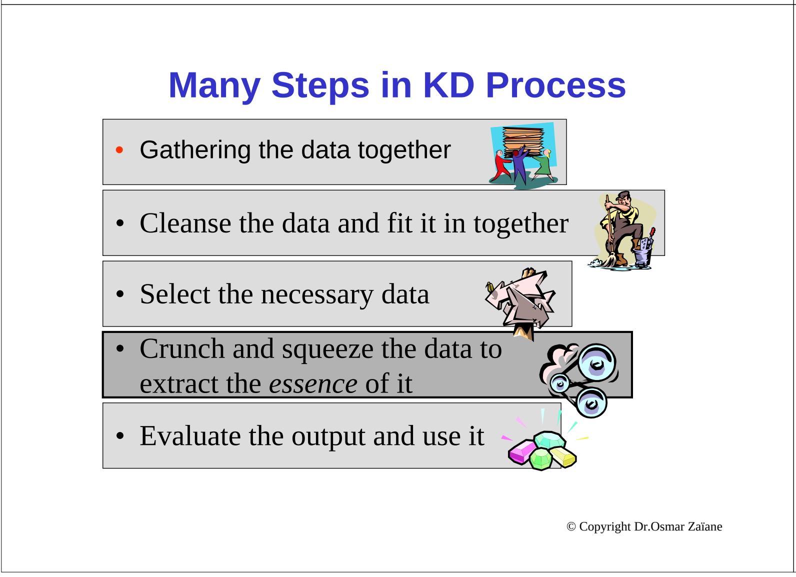
Introduction - Outline

- What kind of information are we collecting?
- What are Data Mining and Knowledge Discovery?
- What kind of data can be mined?
- What can be discovered?
- Is all that is discovered interesting and useful?
- How do we categorize data mining systems?
- What are the issues in Data Mining?
- Are there application examples?



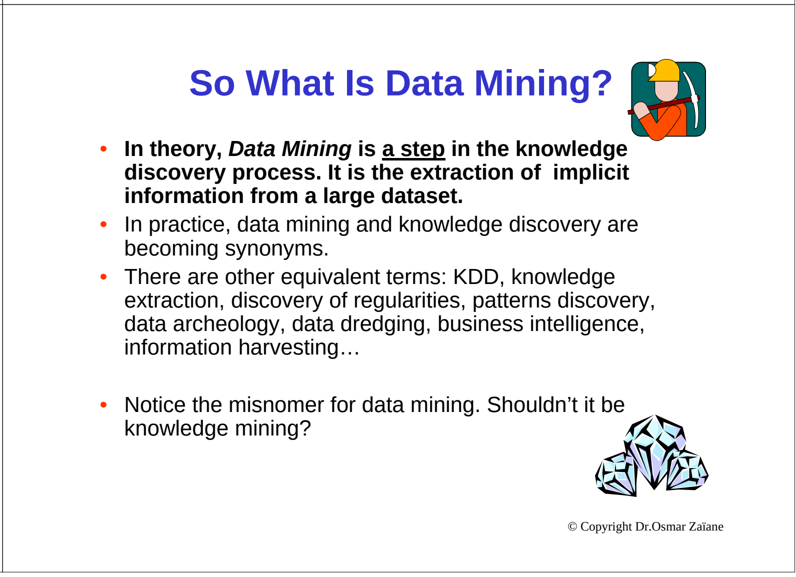
Knowledge Discovery

Process of non trivial extraction of implicit, previously unknown and potentially useful information from large collections of data



Many Steps in KD Process

- Gathering the data together
- Cleanse the data and fit it in together
- Select the necessary data
- Crunch and squeeze the data to extract the *essence* of it
- Evaluate the output and use it

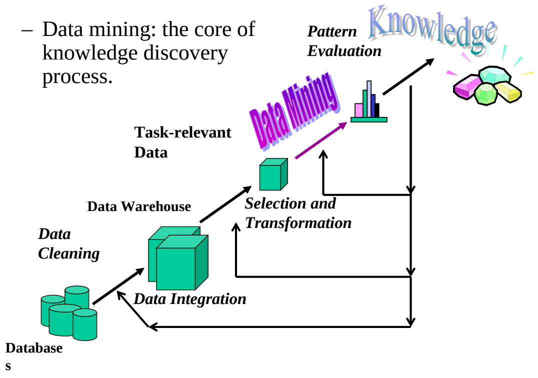


So What Is Data Mining?

- **In theory, *Data Mining* is a step in the knowledge discovery process. It is the extraction of implicit information from a large dataset.**
- In practice, data mining and knowledge discovery are becoming synonyms.
- There are other equivalent terms: KDD, knowledge extraction, discovery of regularities, patterns discovery, data archeology, data dredging, business intelligence, information harvesting…
- Notice the misnomer for data mining. Shouldn't it be knowledge mining?

Data Mining: A KDD Process

– Data mining: the core of knowledge discovery process.



Pattern Evaluation

Knowledge

Data Mining

Task-relevant Data

Selection and Transformation

Data Warehouse

Data Cleaning

Data Integration

Databases

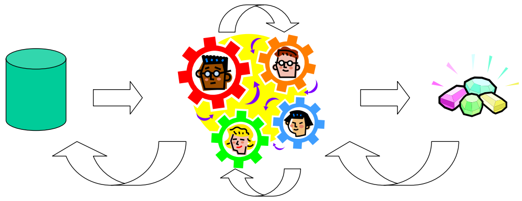
Steps of a KDD Process

- ❑ Learning the application domain
  (relevant prior knowledge and goals of application)
- ❑ Gathering and integrating of data
- ❑ Cleaning and preprocessing data (may take 60% of effort!)
- ❑ Reducing and projecting data
  (Find useful features, dimensionality/variable reduction,…)
- ❑ Choosing functions of data mining
  (summarization, classification, regression, association, clustering,…)
- ❑ Choosing the mining algorithm(s)
- ❑ Data mining: search for patterns of interest
- ❑ Evaluating results
- ❑ Interpretation: analysis of results.
  (visualization, alteration, removing redundant patterns, …)
- ❑ Use of discovered knowledge

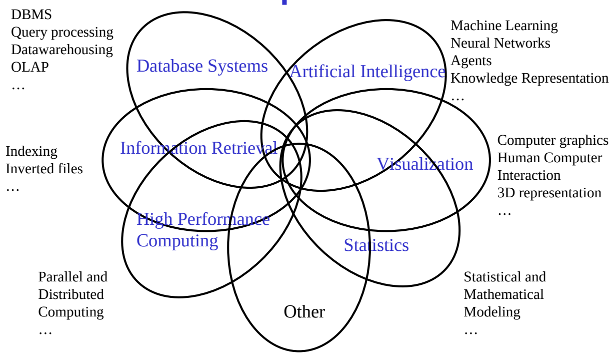
KDD Steps can be Merged

Data cleaning + data integration = data pre-processing
Data selection + data transformation = data consolidation

KDD Is an Iterative Process



KDD at the Confluence of Many Disciplines



DBMS
Query processing
Datawarehousing
OLAP
…

Database Systems

Artificial Intelligence

Machine Learning
Neural Networks
Agents
Knowledge Representation
…

Indexing
Inverted files
…

Information Retrieval

Visualization

Computer graphics
Human Computer Interaction
3D representation
…

High Performance Computing

Statistics

Parallel and Distributed Computing
…

Other

Statistical and Mathematical Modeling
…

Introduction - Outline

- What kind of information are we collecting?
- What are Data Mining and Knowledge Discovery?
- What kind of data can be mined?
- What can be discovered?
- Is all that is discovered interesting and useful?
- How do we categorize data mining systems?
- What are the issues in Data Mining?
- Are there application examples?

Data Mining: On What Kind of Data?

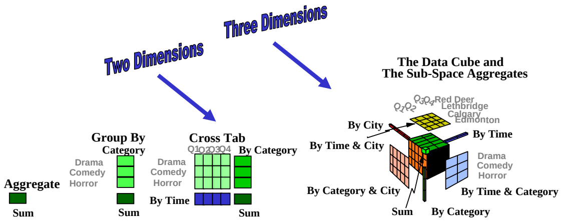
- Flat Files
- Heterogeneous and legacy databases
- Relational databases

  and other DB: Object-oriented and object-relational databases
- Transactional databases

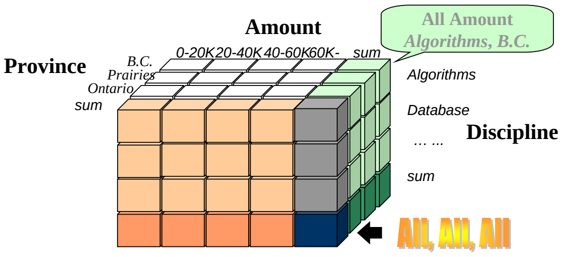
  Transaction(<u>TID</u>, Timestamp, UID, {item1, item2,…})

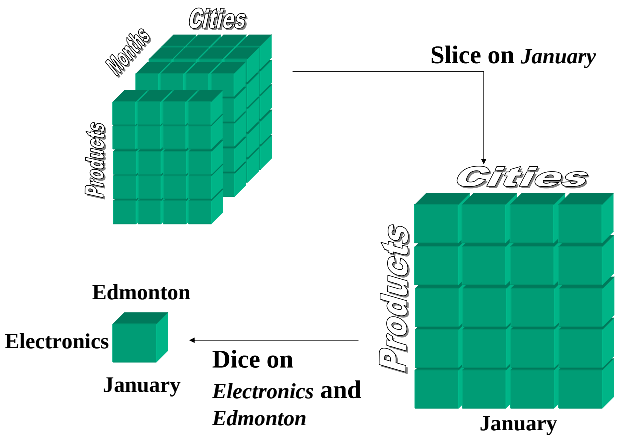
Data Mining: On What Kind of Data?

- Data warehouses



Construction of Multi-dimensional Data Cube

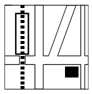




Data Mining: On What Kind of Data?

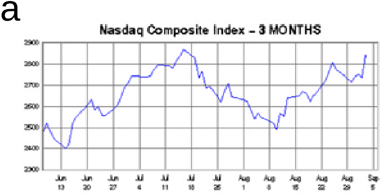
- Multimedia databases

- Spatial Databases



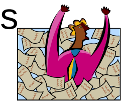
Data Mining: On What Kind of Data?

- Time Series Data and Temporal Data



Data Mining: On What Kind of Data?

- Text Documents



- The World Wide Web
    - ➢The content of the Web
    - ➢The structure of the Web
    - ➢The usage of the Web

Introduction - Outline

- What kind of information are we collecting?
- What are Data Mining and Knowledge Discovery?
- What kind of data can be mined?
- What can be discovered?
- Is all that is discovered interesting and useful?
- How do we categorize data mining systems?
- What are the issues in Data Mining?
- Are there application examples?

What Can Be Discovered?

What can be discovered depends
upon the data mining task employed.

- Descriptive DM tasks
  Describe general properties
- Predictive DM tasks
  Infer on available data

Data Mining Functionality

- Characterization:

Summarization of general features of objects in a target class. (Concept description)

Ex: Characterize grad students in Science

- Discrimination:

Comparison of general features of objects between a target class and a contrasting class. (Concept comparison)

Ex: Compare students in Science and students in Arts

Data Mining Functionality (Con't)

- Association:

Studies the frequency of items occurring together in transactional databases.

Ex: buys(x, bread) → buys(x, milk).

- Prediction:

Predicts some unknown or missing attribute values based on other information.

Ex: Forecast the sale value for next week based on available data.

Data Mining Functionality (Con't)

- Classification:

Organizes data in given classes based on attribute values. (supervised classification)

Ex: classify students based on final result.

- Clustering:

Organizes data in classes based on attribute values. (unsupervised classification)

Ex: group crime locations to find distribution patterns.

Minimize inter-class similarity and maximize intra-class similarity

Data Mining Functionality (Con't)

- Outlier analysis:

Identifies and explains exceptions (surprises)

- Time-series analysis:

Analyzes trends and deviations; regression, sequential pattern, similar sequences…

Introduction - Outline

- What kind of information are we collecting?
- What are Data Mining and Knowledge Discovery?
- What kind of data can be mined?
- What can be discovered?
- Is all that is discovered interesting and useful?
- How do we categorize data mining systems?
- What are the issues in Data Mining?
- Are there application examples?

Is all that is Discovered Interesting?

A data mining operation may generate thousands of patterns, not all of them are interesting.

- Suggested approach: Human-centered, query-based, focused mining

Data Mining results are sometimes so large that we may need to mine it too (Meta-Mining?)

How to measure? ➔ *Interestingness*

Interestingness

- Objective vs. subjective interestingness measures:
  - Objective: based on statistics and structures of patterns, e.g., support, confidence, lift, correlation coefficient etc.
  - Subjective: based on user's beliefs in the data, e.g., unexpectedness, novelty, etc.

Interestingness measures: A pattern is interesting if it is

- ➢easily understood by humans
- ➢valid on new or test data with some degree of certainty.
- ➢potentially useful
- ➢novel, or validates some hypothesis that a user seeks to confirm

Can we Find All and Only the Interesting Patterns?

- Find all the interesting patterns: Completeness.
  - Can a data mining system find all the interesting patterns?
- Search for only interesting patterns: Optimization.
  - Can a data mining system find only the interesting patterns?
  - Approaches
    - First find all the patterns and then filter out the uninteresting ones.
    - Generate only the interesting patterns --- mining query optimization

Like the concept of *precision* and *recall* in information retrieval

Introduction - Outline

- What kind of information are we collecting?
- What are Data Mining and Knowledge Discovery?
- What kind of data can be mined?
- What can be discovered?
- Is all that is discovered interesting and useful?
- How do we categorize data mining systems?
- What are the issues in Data Mining?
- Are there application examples?

Data Mining: Classification Schemes

- There are many data mining systems.

  Some are specialized and some are comprehensive

- Different views, different classifications:
  - Kinds of knowledge to be discovered,
  - Kinds of databases to be mined, and
  - Kinds of techniques adopted.

Four Schemes in Classification

- **Knowledge to be mined**:
  - Summarization (characterization), comparison, association, classification, clustering, trend, deviation and pattern analysis, etc.
  - Mining knowledge at different abstraction levels:
    primitive level, high level, multiple-level, etc.
- **Techniques adopted**:
  - Database-oriented, data warehouse (OLAP), machine learning, statistics, visualization, neural network, etc.

Four Schemes in Classification (con't)

- **Data source to be mined**: (application data)
  - Transaction data, time-series data, spatial data, multimedia data, text data, legacy data, heterogeneous/distributed data, World Wide Web, etc.
- **Data model on which the data to be mined is drawn**:
  - Relational database, extended/object-relational database, object-oriented database, deductive database, data warehouse, flat files, etc.

Designations for Mining Complex Types of Data

- **Text Mining:**
  - Library database, e-mails, book stores, Web pages.
- **Spatial Mining:**
  - Geographic information systems, medical image database.
- **Multimedia Mining:**
  - Image and video/audio databases.
- **Web Mining:**
  - Unstructured and semi-structured data
  - Web access pattern analysis

Introduction - Outline

- What kind of information are we collecting?
- What are Data Mining and Knowledge Discovery?
- What kind of data can be mined?
- What can be discovered?
- Is all that is discovered interesting and useful?
- How do we categorize data mining systems?
- What are the issues in Data Mining?
- Are there application examples?

Requirements and Challenges in Data Mining

- Security and social issues
- User interface issues
- Mining methodology issues
- Performance issues
- Data source issues

Requirements/Challenges in Data Mining (Con't)

- Security and social issues:
  - ❖ Social impact
    - Private and sensitive data is gathered and mined without individual's knowledge and/or consent.
    - New implicit knowledge is disclosed (confidentiality, integrity)
    - Appropriate use and distribution of discovered knowledge (sharing)
  - ❖ Regulations
    - Need for privacy and DM policies

Requirements/Challenges in Data Mining (Con't)

- User Interface Issues:
  - ❖ Data visualization.
    - Understandability and interpretation of results
    - Information representation and rendering
    - Screen real-estate
  - ❖ Interactivity
    - Manipulation of mined knowledge
    - Focus and refine mining tasks
    - Focus and refine mining results

Requirements/Challenges in Data Mining (Con't)

- Mining methodology issues
  - – Mining different kinds of knowledge in databases.
  - – Interactive mining of knowledge at multiple levels of abstraction.
  - – Incorporation of background knowledge
  - – Data mining query languages and ad-hoc data mining.
  - – Expression and visualization of data mining results.
  - – Handling noise and incomplete data
  - – Pattern evaluation: the interestingness problem.

(Source JH)

Requirements/Challenges in Data Mining (Con't)

- Performance issues:
  - ❖ Efficiency and scalability of data mining algorithms.
    - Linear algorithms are needed: no medium-order polynomial complexity, and certainly no exponential algorithms.
    - Sampling
  - ❖ Parallel and distributed methods
    - Incremental mining
    - Can we divide and conquer?

Requirements/Challenges in Data Mining (Con't)

- Data source issues:
  - ❖ Diversity of data types
    - Handling complex types of data
    - Mining information from heterogeneous databases and global information systems.
    - Is it possible to expect a DM system to perform well on all kinds of data? (distinct algorithms for distinct data sources)
  - ❖ Data glut
    - Are we collecting the right data with the right amount?
    - Distinguish between the data that is important and the data that is not.

Requirements/Challenges in Data Mining (Con't)

- Other issues
  - – Integration of the discovered knowledge with existing knowledge: A knowledge fusion problem.

Introduction - Outline

- What kind of information are we collecting?
- What are Data Mining and Knowledge Discovery?
- What kind of data can be mined?
- What can be discovered?
- Is all that is discovered interesting and useful?
- How do we categorize data mining systems?
- What are the issues in Data Mining?
- Are there application examples?

Potential and/or Successful Applications

- Business data analysis and decision support
  - Marketing focalization
    - Recognizing specific market segments that respond to particular characteristics
    - Return on mailing campaign (target marketing)
  - Customer Profiling
    - Segmentation of customer for marketing strategies and/or product offerings
    - Customer behaviour understanding
    - Customer retention and loyalty

Potential and/or Successful Applications (con't)

- Business data analysis and decision support (con't)
  - Market analysis and management
    - Provide summary information for decision-making
    - Market basket analysis, cross selling, market segmentation.
    - Resource planning
  - Risk analysis and management
    - "What if" analysis
    - Forecasting
    - Pricing analysis, competitive analysis.
    - Time-series analysis (Ex. stock market)

Potential and/or Successful Applications (con't)

- Fraud detection
  - Detecting telephone fraud:
    - Telephone call model: destination of the call, duration, time of day or week. Analyze patterns that deviate from an expected norm.

      British Telecom identified discrete groups of callers with frequent intra-group calls, especially mobile phones, and broke a multimillion dollar fraud.
  - Detecting automotive and health insurance fraud
  - Detection of credit-card fraud
  - Detecting suspicious money transactions (money laundering)

Potential and/or Successful Applications (con't)

- Text mining:
  - Message filtering (e-mail, newsgroups, etc.)
  - Newspaper articles analysis
- Medicine
  - Association pathology - symptoms
  - DNA
  - Medical imaging

Potential and/or Successful Applications (con't)

- Sports
  - IBM Advanced Scout analyzed NBA game statistics (shots blocked, assists, and fouls) to gain competitive advantage.

    Spin-off ➔ VirtualGold Inc. for NBA, NHL, etc.
- Astronomy
  - JPL and the Palomar Observatory discovered 22 quasars with the help of data mining.
  - Identifying volcanoes on Jupiter.

Potential and/or Successful Applications (con't)

- Surveillance cameras
  - Use of stereo cameras and outlier analysis to detect suspicious activities or individuals.
- Web surfing and mining
  - IBM Surf-Aid applies data mining algorithms to Web access logs for market-related pages to discover customer preference and behavior pages (e-commerce)
  - Adaptive web sites / improving Web site organization, etc.
  - Pre-fetching and caching web pages
  - Jungo: discovering best sales

Warning: Data Mining Should Not be Used Blindly!

- Data mining approaches find regularities from history, but history is not the same as the future.
- Association does not dictate trend nor causality!?
  - Drinking diet drinks leads to obesity!
  - David Heckerman’s counter-example (1997):
    - buy **hamburgers** 33% of the time, buy **hot dogs** 33% of the time, and buy both **hamburgers** and **hot dogs** 33% of the time; moreover, they buy **barbecue sauce** if and only if they buy **hamburgers**.
    - **hot dogs ➔ barbecue-sauce** has both high support and confidence.(Of course, the rule **hamburgers➔ barbecue-sauce** even higher confidence, but that is an obvious association.)
    - A manager who has a deal on **hot dogs** may choose to sell them at a large discount, hoping to increase profit by simultaneously raising the price of **barbecue**
    - **HOT**-**DOGS** causes **BARBECUE**-**SAUCE** is not part of any possible causal model, could avoid a pricing fiasco.

Quick Overview of some Data Mining Operations

Association Rules
Clustering
Classification

What Is Association Mining?

- **Association rule mining searches for relationships between items in a dataset**:
  - Finding association, correlation, or causal structures among sets of items or objects in transaction databases, relational databases, and other information repositories.
  - Rule form: “**Body → Head [support, confidence]**”.
- **Examples:**
  - buys(x, “bread”) → buys(x, “milk”) [0.6%, 65%]
  - major(x, “CS”) ^ takes(x, “DB”) → grade(x, “A”) [1%, 75%]

Basic Concepts

A transaction is a set of items: $T=\{i_a, i_b, \ldots i_t\}$

$T \subset I$, where I is the set of all possible items $\{i_1, i_2, \ldots i_n\}$

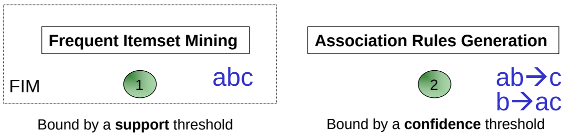
D, the task relevant data, is a set of transactions.

An association rule is of the form:
$P \rightarrow Q$, where $P \subset I$, $Q \subset I$, and $P \cap Q = \varnothing$

P➔Q holds in *D* with <u>support</u> s
and
P➔Q has a <u>confidence</u> c in the transaction set *D*.

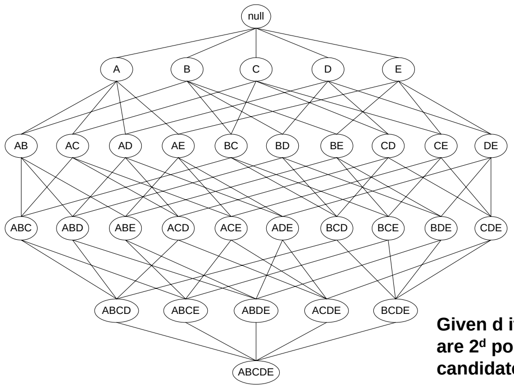
Support(P➔Q) = Probability(P∪Q)
Confidence(P➔Q)=Probability(Q/P)

Association Rule Mining


Frequent Itemset Mining
FIM
1
abc
Association Rules Generation
2
ab→c
b→ac
Bound by a **support** threshold
Bound by a **confidence** threshold

- Frequent itemset generation is still computationally expensive

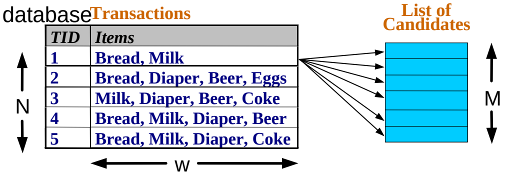
Frequent Itemset Generation



Given d items, there are 2^d possible candidate itemsets

Frequent Itemset Generation

- Brute-force approach (Basic approach):
  - Each itemset in the lattice is a candidate frequent itemset
  - Count the support of each candidate by scanning the database

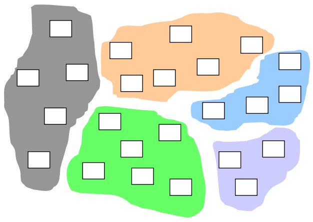


TID	Items
1	Bread, Milk
2	Bread, Diaper, Beer, Eggs
3	Milk, Diaper, Beer, Coke
4	Bread, Milk, Diaper, Beer
5	Bread, Milk, Diaper, Coke

  - Match each transaction against every candidate
  - Complexity ~ O(NMw) => **Expensive since M = 2^d !!!**

Obviously not the right way to do it.

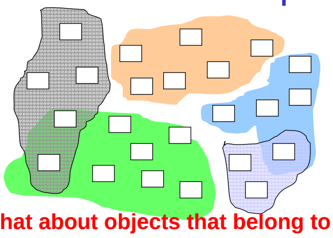
Grouping



Grouping
Clustering
Partitioning

- **We need a notion of similarity or closeness** (what features?)
- **Should we know apriori how many clusters exist?**
- **How do we characterize members of groups?**
- **How do we label groups?**

Grouping

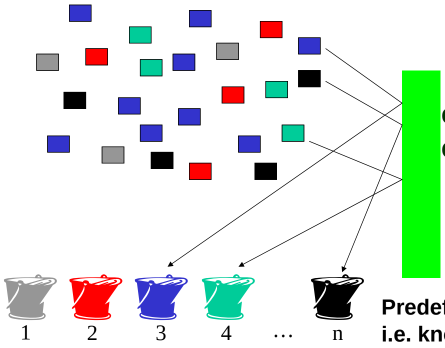


Grouping
Clustering
Partitioning

What about objects that belong to different groups?

- **We need a notion of similarity or closeness** (what features?)
- **Should we know apriori how many clusters exist?**
- **How do we characterize members of groups?**
- **How do we label groups?**

Classification



Classification
Categorization

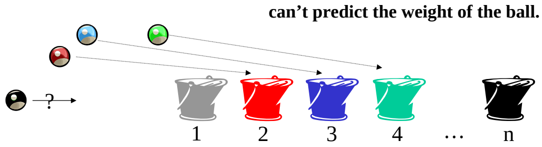
Predefined buckets
i.e. known labels

What is Classification?

The goal of data classification is to organize and categorize data in distinct classes.

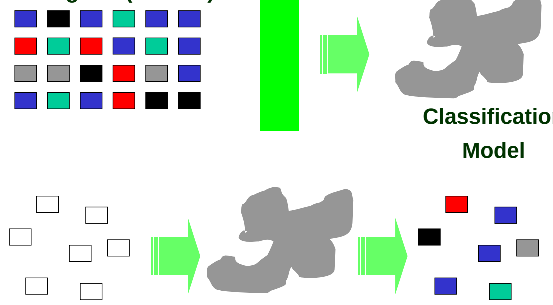
- A model is first created based on the data distribution.
- The model is then used to classify new data.
- Given the model, a class can be predicted for new data.

With classification, I can predict in which bucket to put the ball, but I can't predict the weight of the ball.

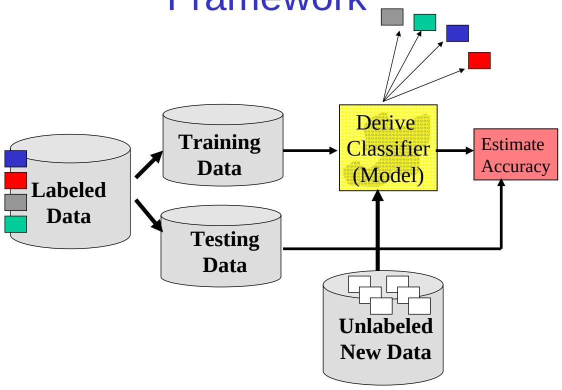


Classification = Learning a Model

Training Set (labeled)



Framework



Classification Methods

- ❖ Decision Tree Induction
- ❖ Neural Networks
- ❖ Bayesian Classification
- ❖ K-Nearest Neighbour
- ❖ Support Vector Machines
- ❖ Associative Classifiers
- ❖ Case-Based Reasoning
- ❖ Genetic Algorithms
- ❖ Rough Set Theory
- ❖ Fuzzy Sets
- ❖ Etc.