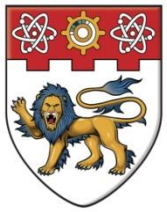


NANYANG TECHNOLOGICAL UNIVERSITY

BC2406 Business Analytics I: Predictive Techniques

Seminar 3
Basics in Predictive Analytics

Instructor: Prof. Lee Gun-woong

Nanyang Business School

Learning Objectives

- View business problems from a data perspective
- Understand the key techniques in Predictive Analytics
- Understand the key procedures in Predictive Analytics

Business Problems from a Data Perspective

Example: Hurricane Frances

- Consider an example from a New York Time story from 2004:
  - Hurricane Frances was on its way, barreling across the Caribbean, threatening a direct hit on Florida's Atlantic coast. Residents made for higher ground, but far away, in Bentonville, Ark., executives at Wal-Mart Stores decided that the situation offered a great opportunity for one of their newest data-driven weapons … predictive technology.

  - A week ahead of the storm's landfall, Linda M. Dillman, Wal-Mart's chief information officer, pressed her staff to come up with forecasts based on what had happened when Hurricane Charley struck several weeks earlier. Backed by the trillions of bytes' worth of shopper history that is stored in Wal-Mart's data warehouse, she felt that the company could 'start predicting what's going to happen, instead of waiting for it to happen,' as she put it. (Hays, 2004)

Example: Hurricane Frances

- The New York Times (Hays, 2004) reported that:"…the experts mined the data and found that the stores would indeed need certain products-and not just the usual flashlights. "We didn't know in the past that strawberry PopTarts increase in sales, like seven times their normal sales rate, ahead of a hurricane', Ms. Dillman said in a recent interview. "And the pre-hurricane top-selling item was beer."
- From such patterns, the company might be able to anticipate unusual demand for products and rush stock to the stores ahead of the hurricane's landfall. Indeed, that is what happened.

Questions

- What is the business problem?
- What kinds of data do you plan to use in this prediction?
- How will be the prediction useful?

Example: Predicting Customer Churn

- Assume you just landed a great analytical job with MegaTelCo, one of the largest telecommunication firms in the United States.
- They are having major problem with customer retention in their wireless business. In the mid-Atlantic region, 20% of cell phone customers leave when their contracts expire, and it is getting increasingly difficult to acquire new customers.
- Since the cell phone market is now saturated, the huge growth in the wireless market has tapered off.

Questions

- What is the business problem?
- How to prevent “churn”?
- If you know how to prevent “churn”, why we discuss “churn” again in BA??

Typical Business Problems and Tasks

- Telecommunication churn problem
  - A subtask that will likely to be part of the solution to any churn problem: estimate from historical data the probability of a customer terminating her contract shortly after it has expired.
- Direct marketing
- Online advertising
- Credit scoring
- Fraud detection
- Product recommendation

Summary

- Each data-driven business decision-making problem is unique
  - Comprising its own combination of goals, desires, constraints, and even personalities.
  - The overall problem needs to be decomposed into subtasks
- Some of these subtasks are unique to the particular business problem, but others are common tasks
  - Walmart case
  - Target case
  - Customer churn case

Data Mining
Tasks / Techniques

Data Mining

- **Data mining** is the **extraction of meaningful information from large databases via technologies.**
- Data mining has come to be referenced by a few similar terms
  - Descriptive/Predictive Analytics
  - Exploratory Data Analysis
  - Data-driven Discovery
  - Knowledge Discovery in Databases (KDD)

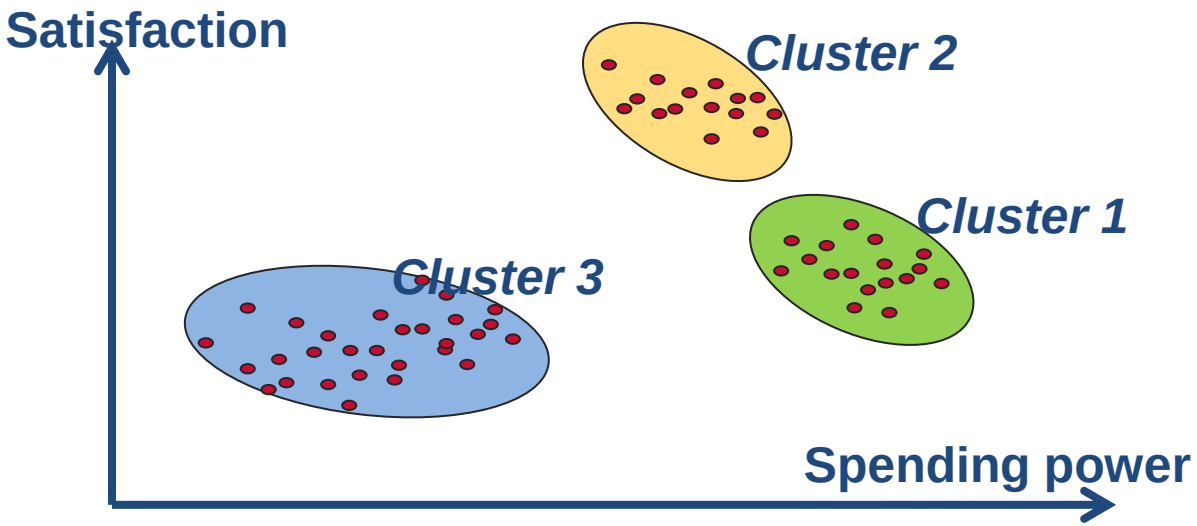
Data Mining

- Clustering
- Association
- Classification
- Prediction
- Data Reduction
- Data Visualization
- Text-mining

Clustering

- Clustering attempts to group individuals in a population together by their similarity, but not driven by any specific purpose.
- Group a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense or another) to each other than to those in other groups (clusters)
- Example question: “ Do our customers form natural groups or segments?”

Business Problem: Identify customers with similar buying behavior



Cluster 1 - Gold customers

- High credit card spending, medium customer satisfaction

Cluster 2 - Silver customers

- Medium credit card spending, high customer satisfaction

Cluster 3 - Bronze customers

- Low credit card spending, low customer satisfaction

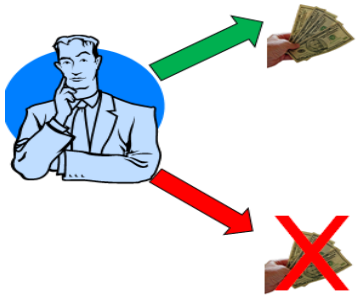
Association

- Association (Co-occurrence grouping) attempts to find association (usually co-occurrence) between entities based on transactions involving them.
  - "what goes with what"
  - " What items are commonly purchased together?"
  - If X was purchased, Y was also purchased
  - Analyzing purchase records from a supermarket.
- Recommendation systems

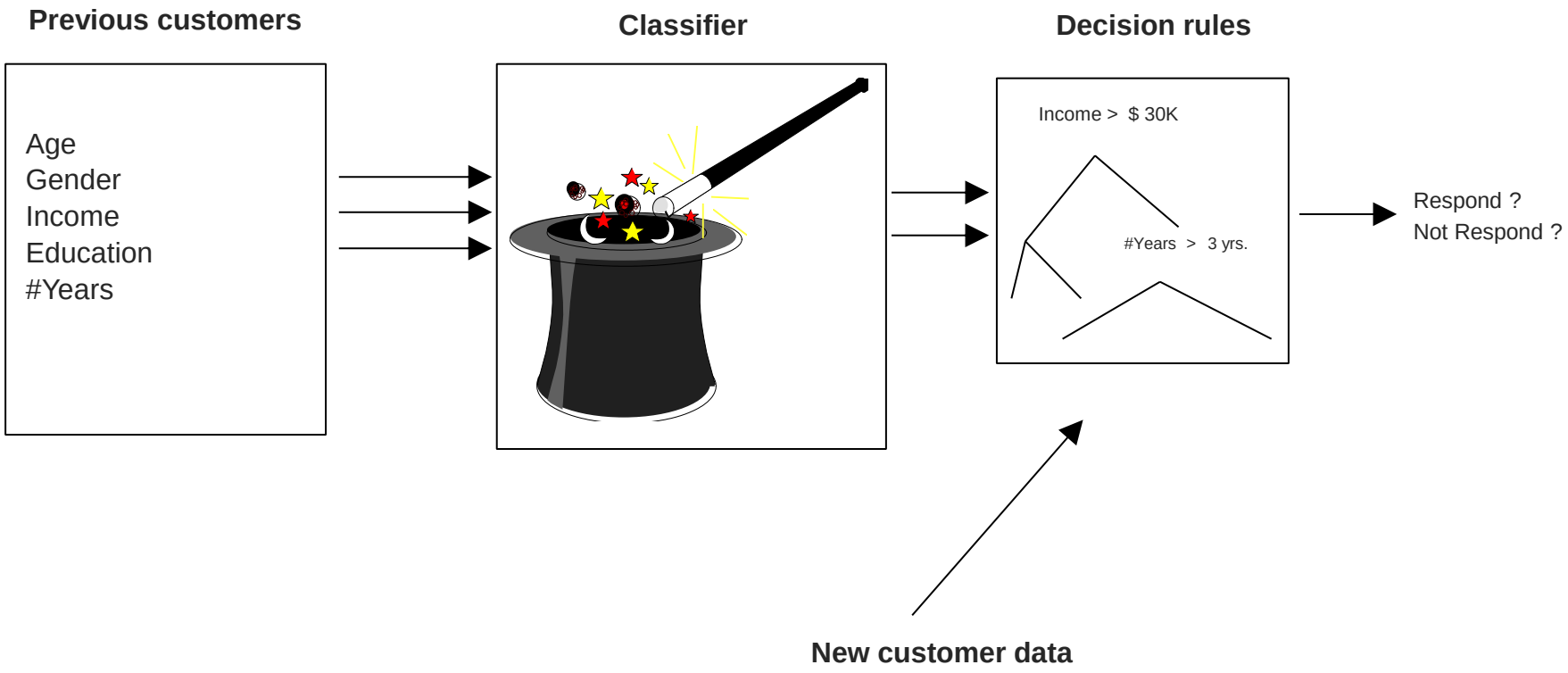
NANYANG TECHNOLOGICAL UNIVERSITY

Classification

- Classification attempt to predict, for each individual in a population, which of a (small) set of classes this individual belongs to.
- Main goal is to build a model that can be applied to unclassified data
- Example question: "Among all customers of MegaTelCo, which are likely to respond to a given offer?"
  - Two classes: (1) will respond and (2) will not respond.



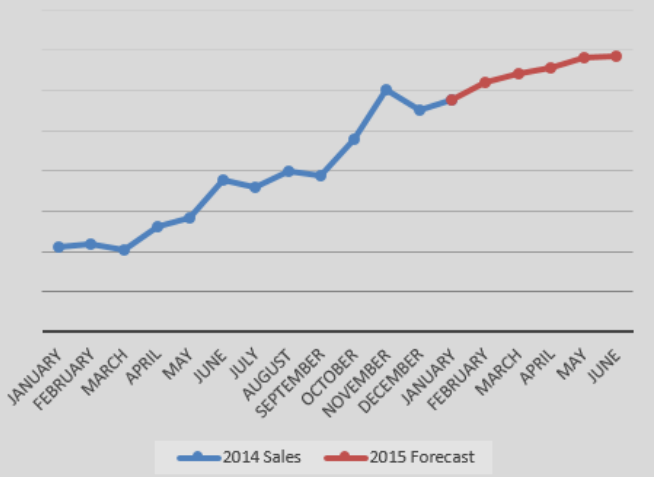
Business Problem: Among all customers of MegaTelCo, which are likely to respond to a given offer?"



Prediction

- Prediction attempts to estimate or predict, for each individual, the numerical value of some variable for that individual.
- Prediction is similar to classification, except that we are trying to predict the value of a numerical variables (e.g., amount of purchase) rather than a class (e.g., purchaser or non-purchaser).
- Example question: " How much will a given customer use the service?"

	2014 Sales	2015 Forecast
January	420	
February	440	
March	410	
April	524	
May	567	
June	755	
July	720	
August	800	
September	780	
October	960	
November	1 200	
December	1 100	
January	1 150	1 150
February		1 240
March		1 280
April		1 310
May		1 360
June		1 370



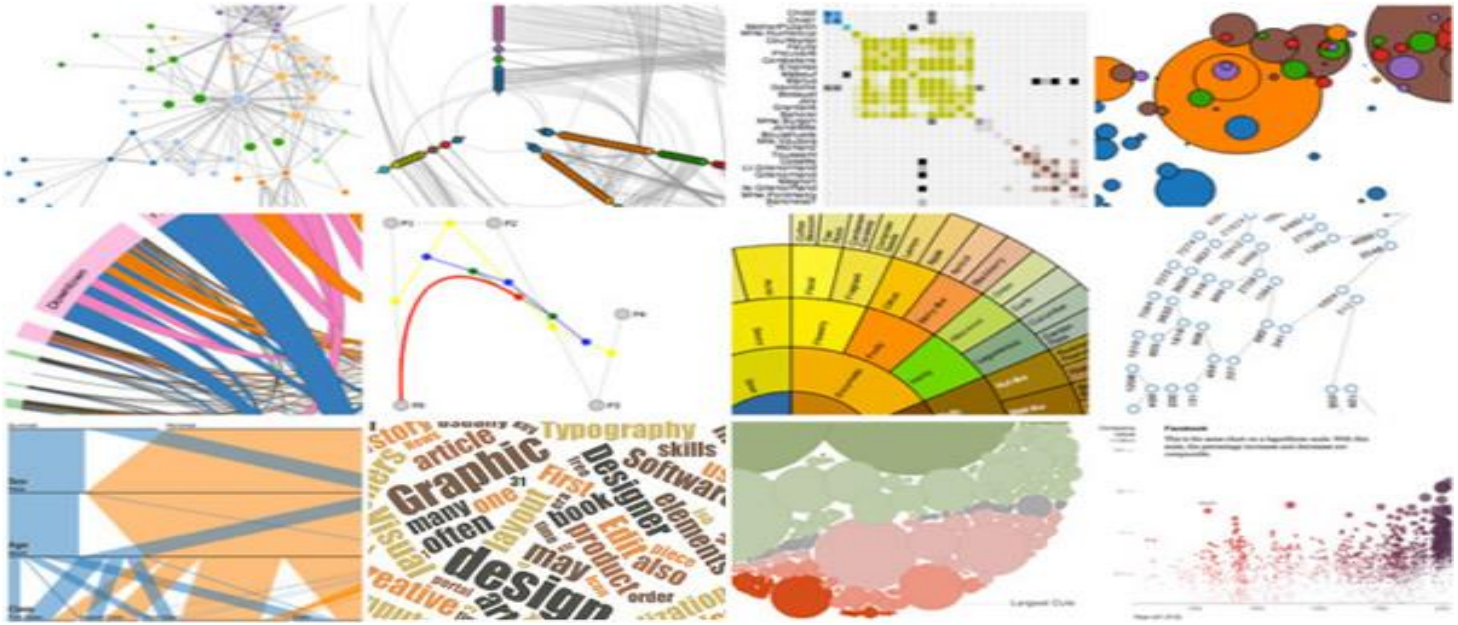
NANYANG TECHNOLOGICAL UNIVERSITY

Data Reduction

- Data reduction attempts to take a large set of data and replace it with a smaller set of data that contains much of the important information in the larger set.
- Reducing the number of variables/columns (e.g., principal components)
  - From {Gender, Income, Age, Education, etc.} To {Income, Education}
- Reducing the number of records/rows (e.g., clustering)
  - From {Client1, Client2, Client3, …} To {Gold, Silver, Bronze}

Data Visualization

- Graphs and plots of data
- Histograms, boxplots, bar charts, scatterplots
- Especially useful to examine relationships between pairs of variables

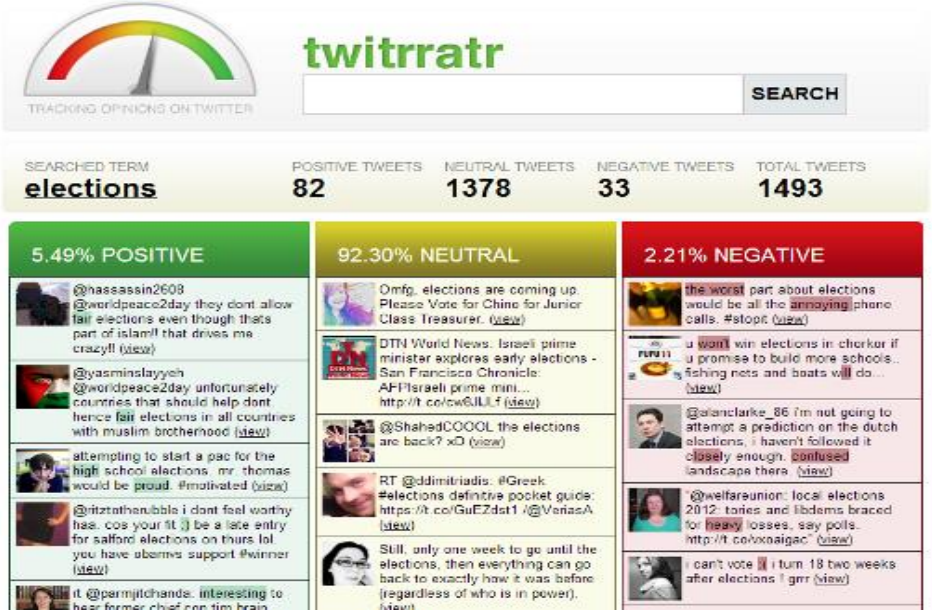


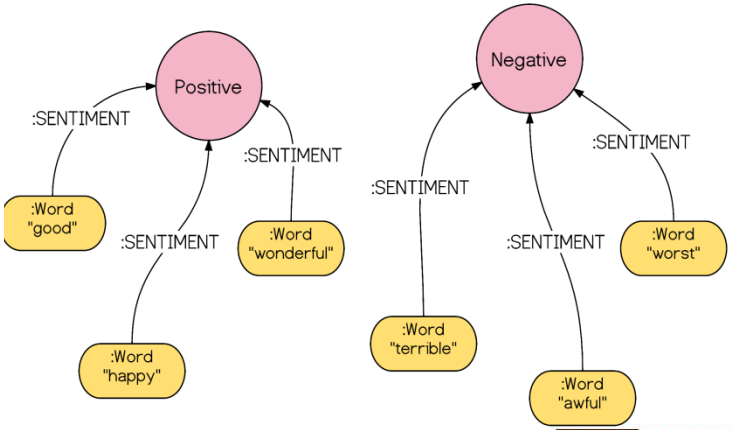
Text Mining

- Discover meaningful patterns from textual data

A word cloud of terms used in mobile App descriptions

- Sentiment analysis





NANYANG TECHNOLOGICAL UNIVERSITY

Supervised Tasks

- Supervised: specific target defined
  - “ Can we find groups of customers who have particularly high likelihoods of canceling their service soon after their contracts expires?”
- So, if a specific target can be provided, the problem can be phrased as a supervised one.
- Predict a single “target” or “outcome” variable
- In data, target value is known

Unsupervised Methods

- Unsupervised: no specific purpose or target specified
  - "Do our customers naturally fall into different group?"
- There is no target (outcome) variable to predict or classify

Exercises

- A pharmacy wants to know whether some patients will switch to other pharmacies?
- An insurance firm wants to identify groups of policy holders with high average claim cost
- A doctor wants to judge whether this patient has the disease or not
- A government wants to identify groups of houses by house type, value, and location

Group Discussion

Is each of the following activities a data mining task? If yes, which kind of data mining task it is, classification, prediction or clustering?

- Dividing the customers of a bank according to their loyalty scores. (2 marks)
- Identify the salesman with the highest sales. (2 marks)
- Predicting the future box office of movies using online reviews of these movies. (2 marks)
- Predicting the outcomes of tossing a (fair) pair of dice. (2 marks)
- Predicting a visitor to the website to be a hacker based on the visitor's past online activities. (2 marks)

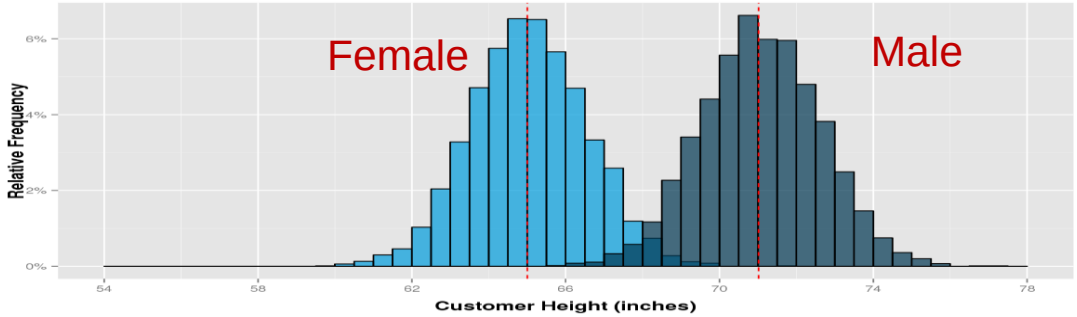
(Exam Question, 2015)

Other Analytics Techniques to Deal with Various Tasks

- Statistics
- Data Querying
- Optimization

Statistics

- Different uses in business analytics
  - For the computation of particular numeric values of interest from data. These values often include sums, averages, rates, and so on, which could be called "summary statistics".
  - Helps us understand different data distributions.



  - Helps us understand how to use data to test hypothesis.
    - Product A will be more purchased by Female customers than Male customers.

Data Querying

- A specific request for a subset of data, formulated in a technical language and posed to a database system.
- Differs fundamentally from data mining in that there is no discovery of patterns or models.
- e.g.,: select * from customers

  where age >45

  and churn = ‘y’
- In firms, what kinds of questions can be answered by query?

Optimizing

- An objective function with some constraints
- STEP1 : Define an objective function that represents our goal
- STEP2 : The function can be calculated for a particular set of weights and a particular set of data.
- STEP3 : Find the optimal value for the weights by maximizing or minimizing the objective function.

Optimization Problem

- The regular air fare between Boston and San Francisco is $500. An airline using planes with a capacity of 300 passengers on this route observes that they fly with an average of 180 passengers. Market research tells the airlines' managers that each $ 5 fare reduction would attract, on average, 3 more passengers for each flight. How should they set the fare to maximize their revenue?

- Step 1: Revenue = Price * Quantity
  - Price = $500 - $5N
  - Quantity = 180 + 3N, where N is #passengers

- Step 2: Maximize Revenue(N) = (500 – 5N)*(180 + 3N)

  $= 90{,}000 + 600N + 15N^2$

- Step 3: Revenue(N)` = 600 – 30N = 0,

  By the first derivative, 20 is the only critical number

  Revenue(N)`` = - 30

  By the second derivative, the Revenue has a maximum at N= 20

The best fare to maximize the Revenue is then: $400 (=$500 -5*20) with 240 (= 180 + 3*20) passengers of a revenue of Revenue(20) = $96,000

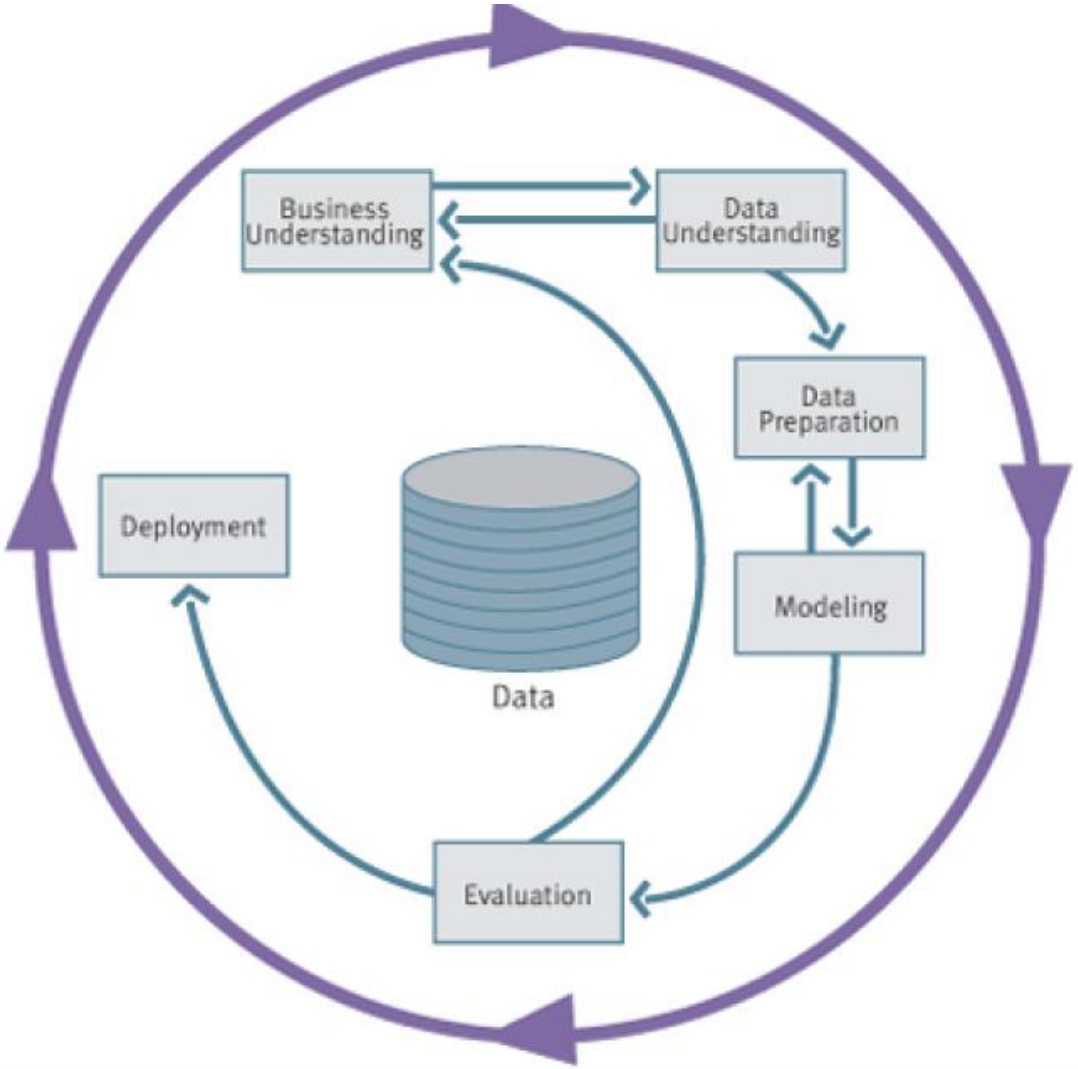
NANYANG TECHNOLOGICAL UNIVERSITY

Which techniques to use?

- Who are the most profitable customers? (Database Query)
- Is there really a difference between the profitable customers and the average customer? (Statistics, hypothesis testing)
- Will some particular new customer be profitable? How much revenue should I expect this customer to generate? (Data Mining)

Procedure in Predictive Analytics

Procedure in Predictive Analytics



Cross-Industry Standard Process for Data Mining (Shearer, 2000)

Data Preparation

- Often proceeds along with data understanding.
- Explore, clean, pre-process data
  - Converting data to tabular format.
  - Removing or inferring missing values.
  - Converting data to different types.

Modeling

- Specify task (classification, clustering, etc.)
- Choose the techniques (regression, neural networks, etc.)
- Output of modeling is some sort of rules or pattern capturing regularities in the data.

Evaluation

- Assess the data mining results rigorously and to gain confidence that they are valid and reliable before moving on.
  - Validate results
  - Compare models
  - Iterative implementation
- Consider the comprehensibility of the model to the stakeholders.
- Stakeholders should check and see whether the model is going to do more good than harm.

Deployment

- Put the best model into real use in order to realize some return on investment.
- The clearest cases of deployment involve implementing a predictive model in some information system or business process.
- Churn example: A model for predicting the likelihood of churn could be integrated with the business procedures for churn management --- for example, by sending special offers to customers who are predicted to be particularly at risk.

What's Next…

- Course Topic
  - Seminar 4: Basics in Predictive Analyses
  - Readings
    - "Seminar 4_Lecture.pdf" lecture slides
- Assignment
  - Read "Seminar 4_R.pdf" tutorial, and Do the exercises in the document