Data Mining and Data Warehouse

Student Name : Tan Chong Raen
Student ID : 1121115725

Multimedia University
(Cyberjaya Campus)

August 27, 2015

Introduction and Problem statement

Introduction

- *Data Mining* : a technique used to discover the meaningful and valuable information from a massive of data and covert it into the knowledge of someone or something behavior.

Problem statement

- Amount of time and cost used for discovery of patterns of data.
- Need to study a massive of data to predict the response of customer.

Research method

- Theoretical work
- Approaches(Use in data mining):
  i. Verification Driven (most common)
  ii. Discovery Driven (most complicated)
- Six phases(Build a data mining project):
  i. Problem Definition
  ii. Data Exploration
  iii. Date Preparation
  iv. Modeling
  v. Evaluation
  vi. Deployment

Critical literature review

Operations associated with Data Mining

i. *Verification Driven*
   a) Multidimensional analysis
   b) Query and reporting
   c) Statistical analysis

ii. *Discovery Driven*
   a) Predictive modeling
   b) Deviation detection
   c) Link analysis
   d) Database Segmentation

Conclusion

- *Data mining* is more accuracy and efficiency.
- *Data mining* may provide the wrong information.
- *Internet* and *Data Warehousing* is more important.

What did we learn? Possible extension or Future work?

What did we learn?

In this research assignment, we learn that the process of a completed data mining project and the advantages of data mining.

Possible extension or Future work?

i. Social media (Facebook, Twitter and so on)
ii. Bioinformatics
iii. Education