



Data Mining

資料探勘

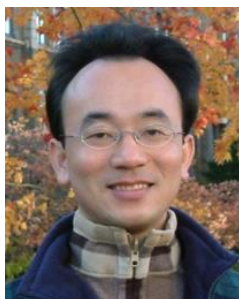
Introduction to Data Mining

(資料探勘導論)

1032DM01

MI4

Wed, 7,8 (14:10-16:00) (B130)



Min-Yuh Day

戴敏育

Assistant Professor

專任助理教授

Dept. of Information Management, Tamkang University

淡江大學 資訊管理學系

<http://mail.tku.edu.tw/myday/>

2015-02-25



淡江大學103學年度第2學期 課程教學計畫表

Spring 2015 (2015.02 - 2015.06)

- 課程名稱：資料探勘 (Data Mining)
- 授課教師：戴敏育 (Min-Yuh Day)
- 開課系級：資管四P (TLMXB4P)
- 開課資料：選修 單學期 2 學分 (2 Credits, Elective)
- 上課時間：週三 7,8 (Wed 14:10-16:00)
- 上課教室：B130

課程簡介

- 本課程介紹資料探勘 (Data Mining) 的基礎概念及應用技術。
- 課程內容包括
 - 資料探勘導論
 - 關連分析 (Association Analysis)
 - 分類與預測 (Classification and Prediction)
 - 分群分析 (Cluster Analysis)
 - 海量資料分析 (Big Data Analytics)
 - 文字探勘與網頁探勘 (Text and Web Mining)
 - SAS企業資料採礦實務 (SAS EM)
 - 資料探勘個案分析與實作

Course Introduction

- This course introduces the **fundamental concepts** and **applications technology** of data mining.
- Topics include
 - Introduction to Data Mining
 - Association Analysis
 - Classification and Prediction
 - Cluster Analysis
 - Big Data Analytics
 - Text and Web Mining
 - **Data Mining Using SAS Enterprise Miner (SAS EM)**
 - **Case Study and Implementation of Data Mining**

課程目標 (Objective)

- 瞭解及應用 資料探勘基本概念與技術。
- Understand and apply the fundamental concepts and technology of data mining

課程大綱 (Syllabus)

週次 (Week)	日期 (Date)	內容 (Subject/Topics)
1	2015/02/25	資料探勘導論 (Introduction to Data Mining)
2	2015/03/04	關連分析 (Association Analysis)
3	2015/03/11	分類與預測 (Classification and Prediction)
4	2015/03/18	分群分析 (Cluster Analysis)
5	2015/03/25	個案分析與實作一 (SAS EM 分群分析) : Case Study 1 (Cluster Analysis – K-Means using SAS EM)
6	2015/04/01	教學行政觀摩日 (Off-campus study)
7	2015/04/08	個案分析與實作二 (SAS EM 關連分析) : Case Study 2 (Association Analysis using SAS EM)
8	2015/04/15	個案分析與實作三 (SAS EM 決策樹、模型評估) : Case Study 3 (Decision Tree, Model Evaluation using SAS EM)

課程大綱 (Syllabus)

週次 (Week)	日期 (Date)	內容 (Subject/Topics)
9	2015/04/22	期中報告 (Midterm Project Presentation)
10	2015/04/29	期中考試週 (Midterm Exam)
11	2015/05/06	個案分析與實作四 (SAS EM 迴歸分析、類神經網路) : Case Study 4 (Regression Analysis, Artificial Neural Network using SAS EM)
12	2015/05/13	海量資料分析 (Big Data Analytics)
13	2015/05/20	文字探勘與網頁探勘 (Text and Web Mining)
14	2015/05/27	期末報告 (Final Project Presentation)
15	2015/06/03	畢業考試週 (Final Exam)

教學方法與評量方法

- 教學方法
 - 講述、討論、賞析、實作、問題解決
- 評量方法
 - 紙筆測驗、實作、報告、上課表現

教材課本

- 講義 (Slides)
- 參考書籍
 - Applied Analytics Using SAS Enterprise Mining, Jim Georges, Jeff Thompson and Chip Wells, SAS, 2010
 - Big Data, Data Mining, and Machine Learning: Value Creation for Business Leaders and Practitioners, Jared Dean, Wiley, 2014
 - Data Science for Business: What you need to know about data mining and data-analytic thinking, Foster Provost and Tom Fawcett, O'Reilly, 2013

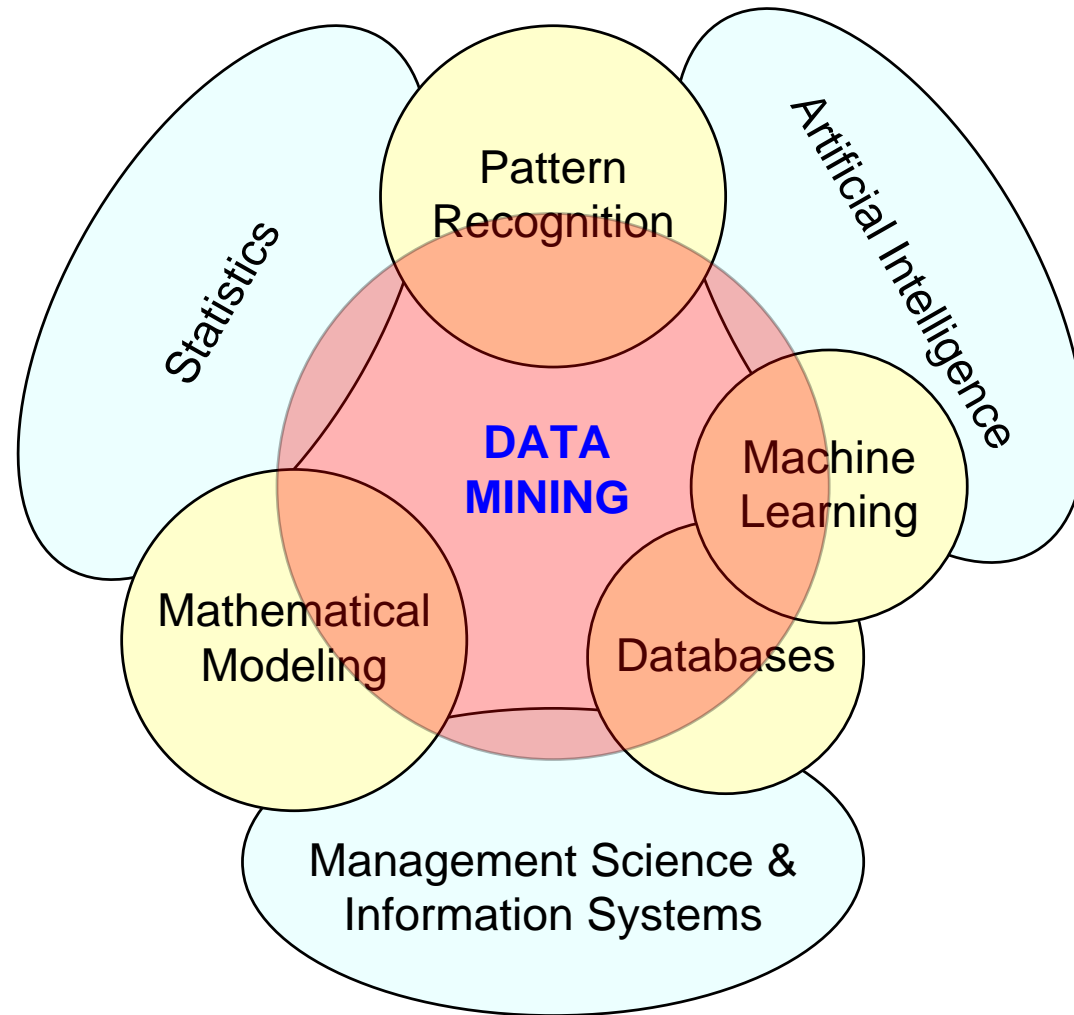
作業與學期成績計算方式

- 作業篇數
 - 3篇
- 學期成績計算方式
 - 期中評量：30 %
 - 期末評量：30 %
 - 其他（課堂參與及報告討論表現）：40 %

Team Term Project

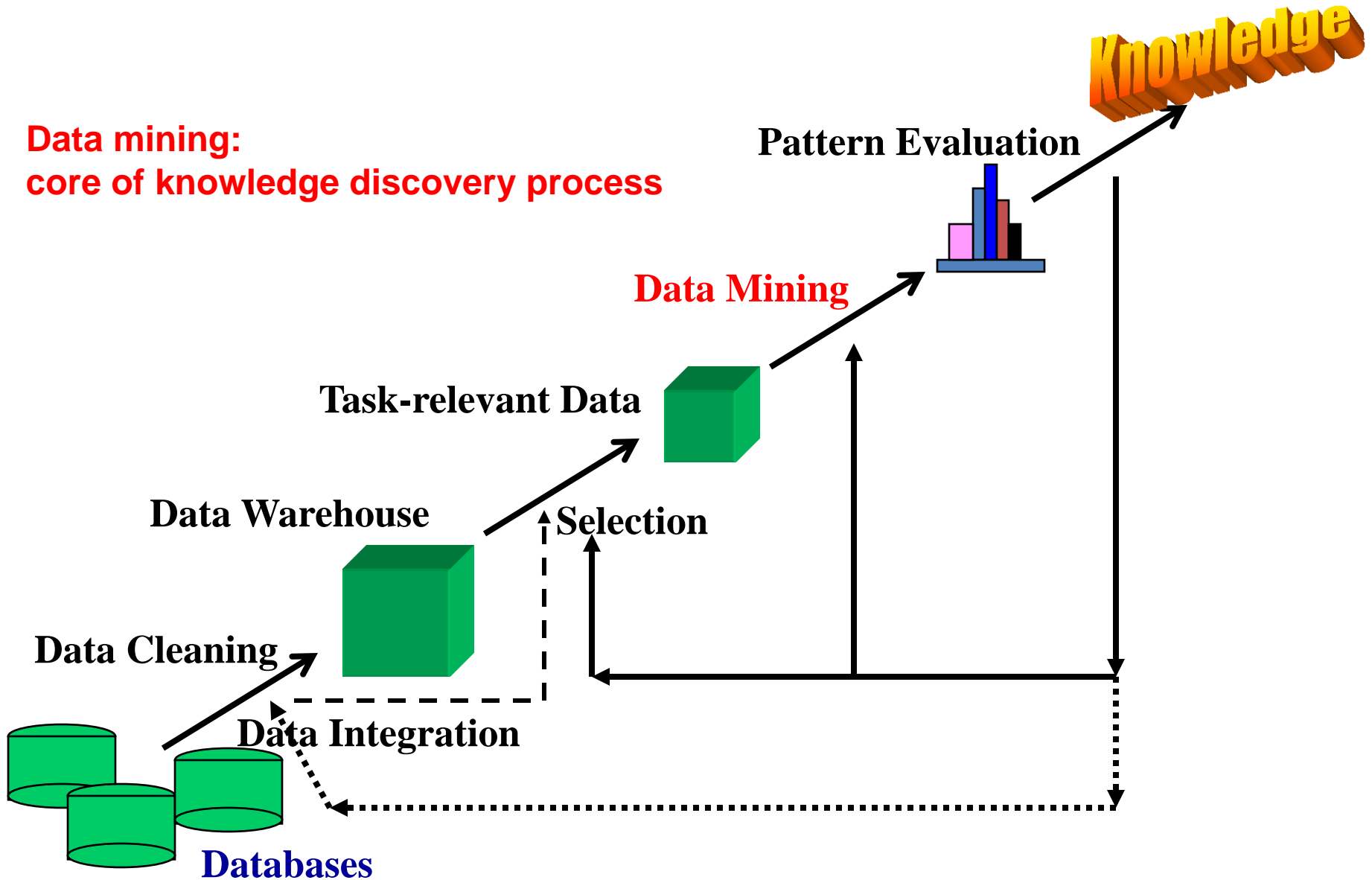
- Term Project Topics
 - Data mining
 - Web mining
 - Business Intelligence
 - Big Data Analytics
- 3-4 人為一組
 - 分組名單於 2015/03/04 (三) 課程下課時繳交
 - 由班代統一收集協調分組名單

Data Mining at the Intersection of Many Disciplines



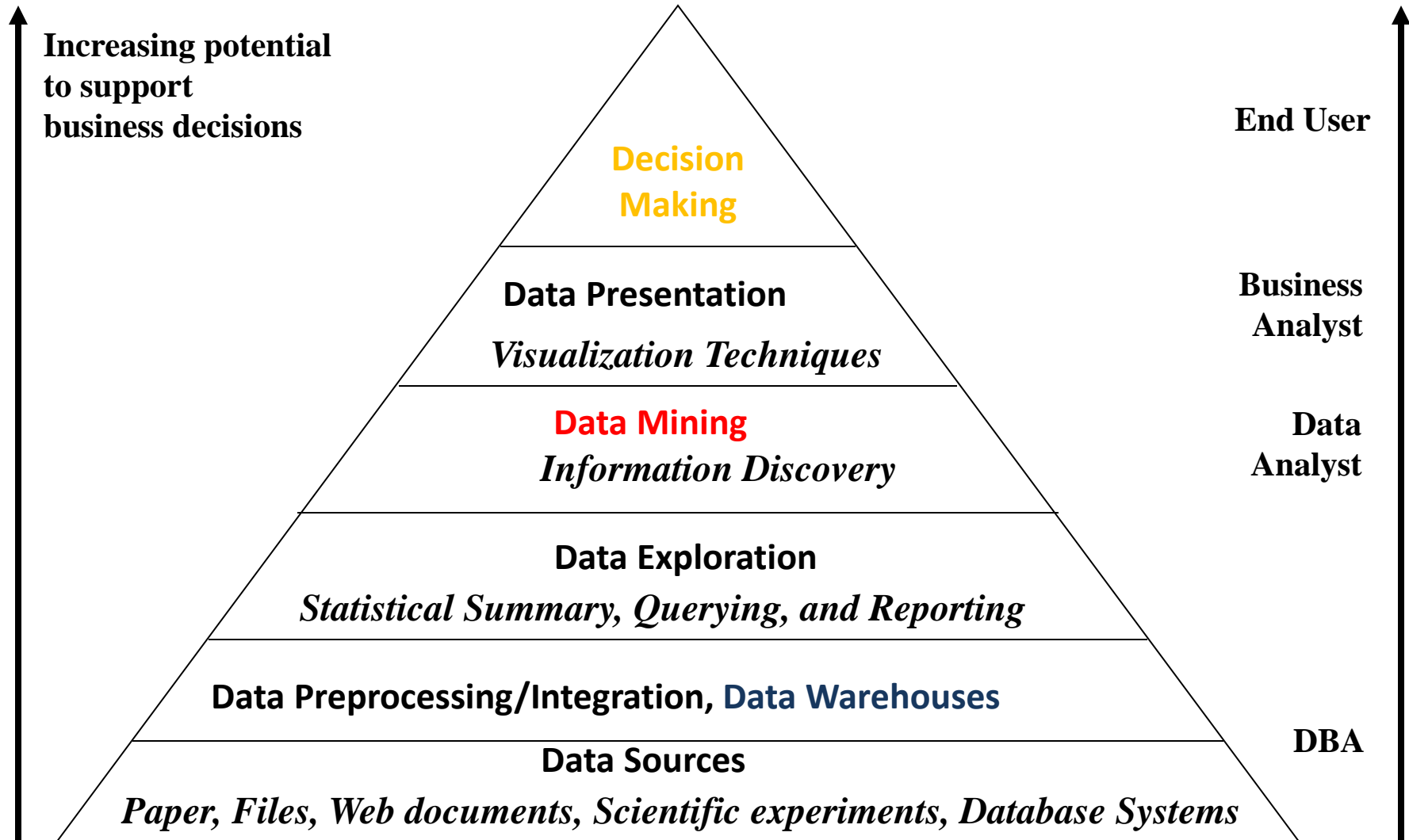
Knowledge Discovery (KDD) Process

Data mining:
core of knowledge discovery process

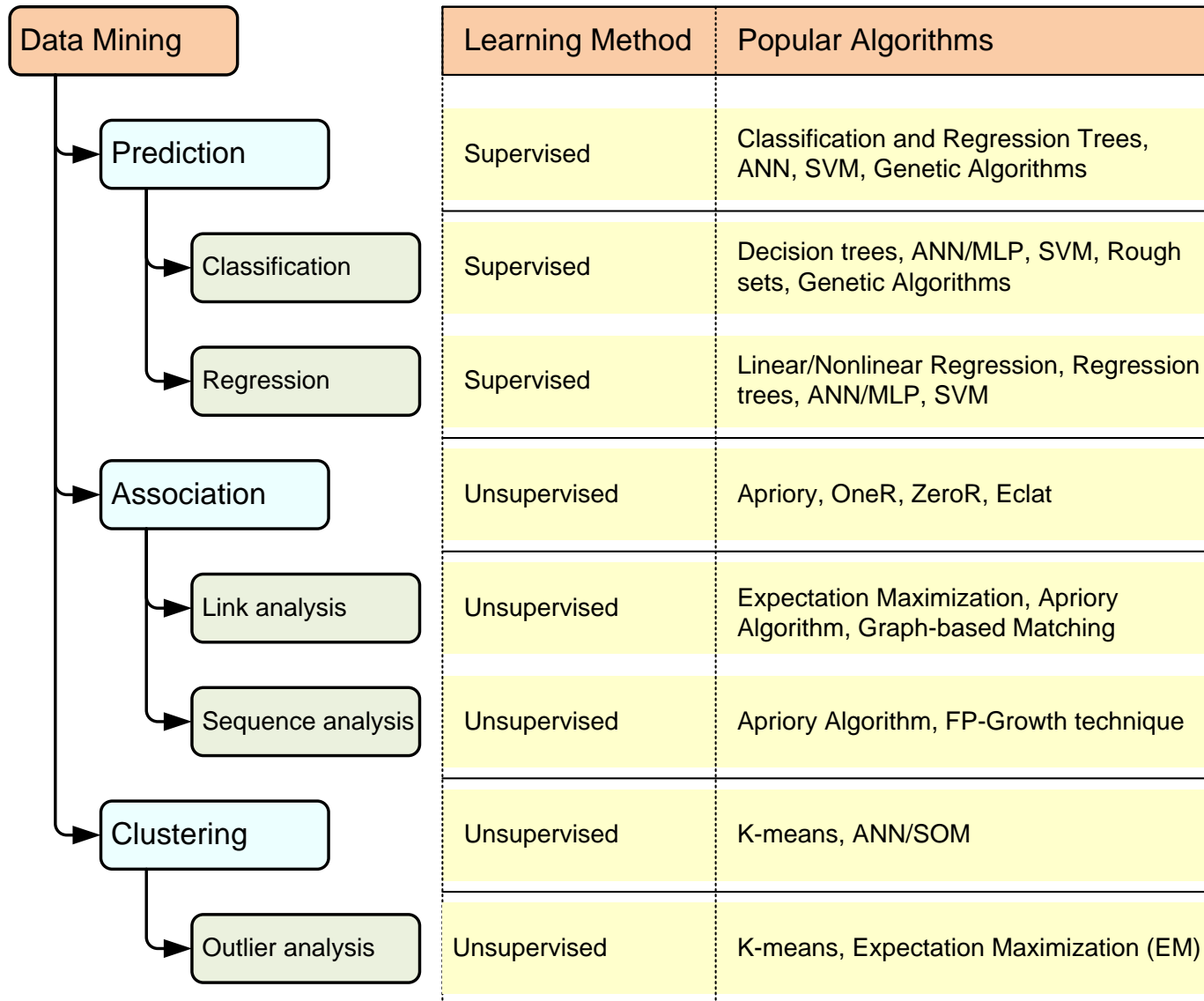


Data Warehouse

Data Mining and Business Intelligence



A Taxonomy for Data Mining Tasks

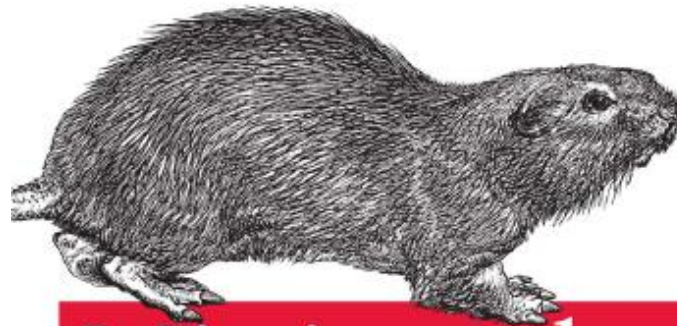


The Evolution of BI Capabilities



Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites

*Analyzing Data from Facebook, Twitter, LinkedIn,
and Other Social Media Sites*



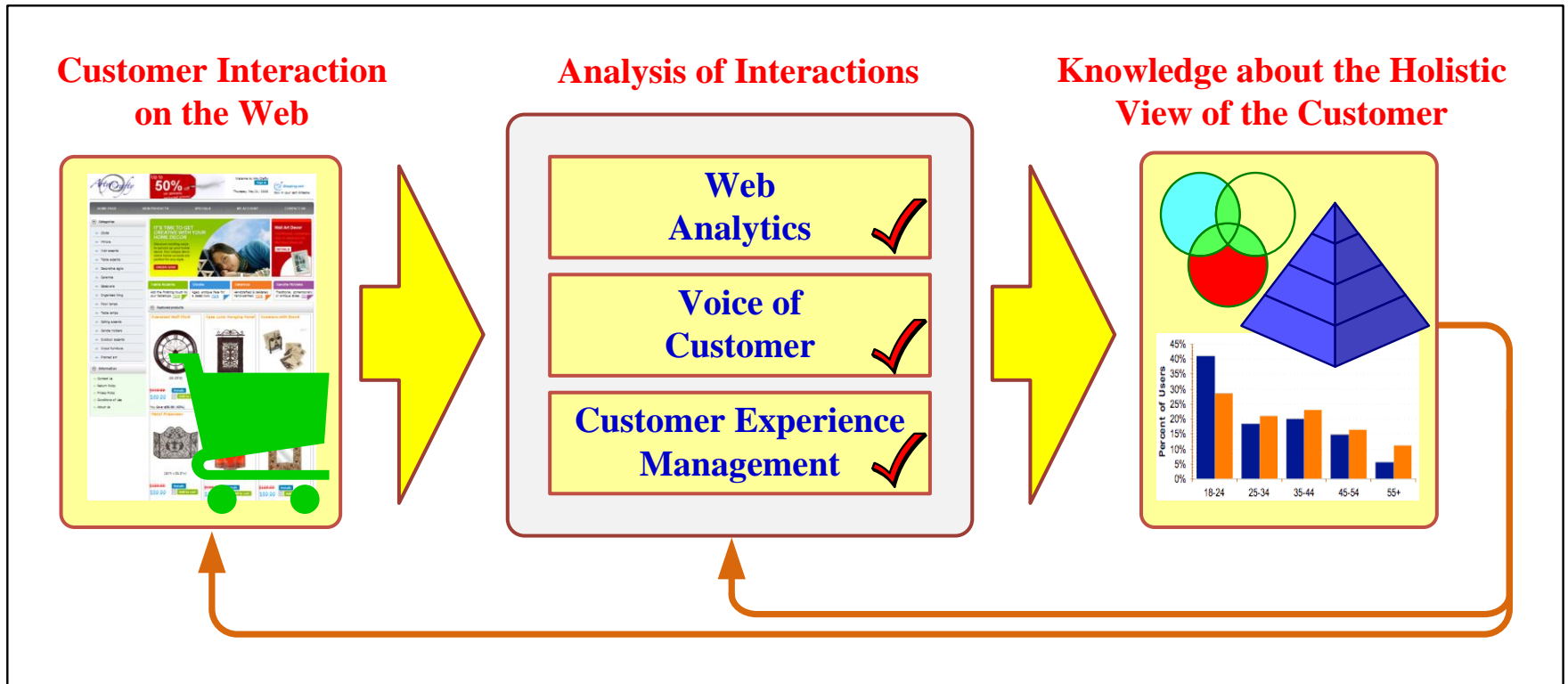
Mining the
Social Web

O'REILLY®

Matthew A. Russell

Web Mining Success Stories

- Amazon.com, Ask.com, Scholastic.com, ...
- Website Optimization Ecosystem



Business Intelligence Trends

1. **Agile** Information Management (IM)
2. **Cloud** Business Intelligence (BI)
3. **Mobile** Business Intelligence (BI)
4. **Analytics**
5. **Big Data**

Business Intelligence Trends: Computing and Service

- Cloud Computing and Service
- Mobile Computing and Service
- Social Computing and Service

Business Intelligence and Analytics

- Business Intelligence 2.0 (BI 2.0)
 - Web Intelligence
 - Web Analytics
 - Web 2.0
 - Social Networking and Microblogging sites
- Data Trends
 - Big Data
- Platform Technology Trends
 - Cloud computing platform

Business Intelligence and Analytics: Research Directions

1. Big Data Analytics

- Data analytics using Hadoop / MapReduce framework

2. Text Analytics

- From Information Extraction to Question Answering
- From Sentiment Analysis to Opinion Mining

3. Network Analysis

- Link mining
- Community Detection
- Social Recommendation

Data Scientist:

The Sexiest Job of the 21st Century

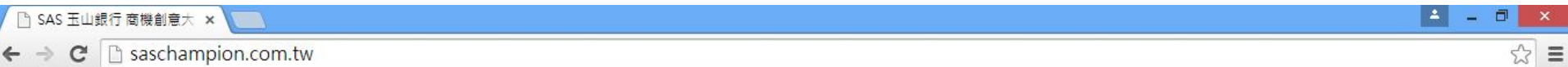
**Meet the people who
can coax treasure out of
messy, unstructured data.**

*by Thomas H. Davenport
and D.J. Patil*

When Jonathan Goldman arrived for work in June 2006 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their friends and colleagues to join. But users weren't seeking out connections with the people who were already on the site at the rate executives had expected. Something was apparently missing in the social experience. As one LinkedIn manager put it, "It was like arriving at a conference reception and realizing you don't know anyone. So you just stand in the corner sipping your drink—and you probably leave early."

SAS與玉山銀行

【大數據數位行銷應用大賽】



最新消息

大賽起源

活動辦法

我要報名

常見問題



第四屆資料科學家系列

大數據 數位行銷 應用大賽



贏在求職起跑點 提升職場競爭力
想要躋身進入資料科學家的行列
想要證明自己是資料分析的高手

挑戰
\$300,000
高額總獎金

SAS與
玉山銀行優先
招募與面試

我要報名

資料科學家系列

大數據數位行銷應用大賽

在這個巨量資料的時代，懂得巨量分析的專業人才「資料科學家」(Data Scientist)將成為未來炙手可熱的明日之星。

<http://saschampion.com.tw/>

SAS與玉山銀行

【大數據數位行銷應用大賽】



第四屆資料科學家系列大

www.accupass.com/go/saschampion

ACCUPASS 找活動 辦活動 申請贊助 登入 註冊

活動介紹 報名表下載 票券資訊 討論區

討論 分享 6

活動時間與地點報名辦法

- (一) 報名日期：2015年2月24日(二)至3月18日(三)，報名限量150組，額滿為止！
- (二) 起跑說明會：2015年3月26日(四) - 玉山銀行人力發展中心 1樓登峰廳 (參賽團隊每組皆須指派2名隊員出席)
- (三) 初賽訓練課程：2015年3月28日(六)至4月1日(三) - 臺北大學資訊中心及逢甲大學電腦教室，每梯次為期一天。
- (四) EG個人能力檢測：2015年5月3日(日) - 臺北大學資訊中心、淡江大學電腦教室、逢甲大學電腦教室，共9場次，考試時間為14:00-16:00。
- (五) 複賽訓練課程：2015年5月8日(五)至5月9日(六) - 臺北大學資訊中心，共2梯次，每梯次為期一天。*備註*複賽入圍之隊伍方可參加
- (六) 複賽繳件日期：2015年5月29日(五)
- (七) 決賽簡報日期：2015年6月12日(五)
- (八) 頒獎典禮日期：2015年6月30日(二)

報名即將在**2015年2月24日**正式開跑！
早鳥報名優惠每組只要**\$600元**！
(含每人初賽SAS教育訓練課程與教材、EG能力檢測)

**挑戰巨量分析、迎接數位行銷新浪潮，
快召集身邊分析好手，一同組隊來踢館！**

<http://www.accupass.com/go/saschampion>

Summary

- This course introduces the **fundamental concepts** and **applications technology** of data mining.
- Topics include
 - Introduction to Data Mining
 - Association Analysis
 - Classification and Prediction
 - Cluster Analysis
 - Big Data Analytics
 - Text and Web Mining
 - **Data Mining Using SAS Enterprise Miner (SAS EM)**
 - **Case Study and Implementation of Data Mining**

Contact Information

戴敏育 博士 (Min-Yuh Day, Ph.D.)

專任助理教授

淡江大學 資訊管理學系

電話：02-26215656 #2846

傳真：02-26209737

研究室：B929

地址：25137 新北市淡水區英專路151號

Email：myday@mail.tku.edu.tw

網址：<http://mail.tku.edu.tw/myday/>

