

# » 大數據與商業分析

## Big Data and Business Analytics

楊立偉教授

台灣大學管理學院

2020

- 
- ◆ 課程介紹
  - ◆ 選課介紹

# 授課老師 楊立偉教授

## ◆ 研究領域為資料庫及語意分析技術、知識管理、數位行銷

|    |  |
|----|--|
| 現任 | 台大工管系暨商研所兼任助理教授 2006~                              |
|    | 台大資管系兼任助理教授 2017~                                  |
|    | 資訊及通信國家標準技術委員                                      |
|    | 意藍資訊 董事總經理 ( 創辦人 ) 1999~<br>國內規模最大的網路情報與社群口碑自動分析平台 |
|    | 龍捲風科技 董事總經理<br>國內企業搜尋引擎市佔率最高；國際檢索競賽第一名             |
| 經歷 | 獲選100 MVP最有價值經理人，擁有超過20項語意分析專利                     |
|    | 獲國家雲端創新獎、數位時代「創業之星」首獎、台大校級教學優良獎                    |
|    | 獲數位時代雜誌選為代表台灣軟體的封面人物之一                             |

# 課程介紹 (1)

- 為管理學院特色課程，並為創新創業學程認列之選修課程
- 「大數據」( Big Data ) 自 2011 年起逐漸進入商業大眾的視野，相關書籍與論述如雨後春筍般出現，各行各業也都對此一名詞充滿期待。現實商業情境中，數據分析是圍繞「大數據」各種想像的基礎條件。著眼於培養企業需求日般的數據分析人才，此一跨系課程針對台大管理學院的同學而設，透過講演與專案，概念與實作並重，有系統介紹數據分析的商管相關應用，為有志於成為商業應用領域數據科學家的不同背景修課同學，奠定未來發展的基礎。

# 課程介紹 (2)

- 本課程內容以金融及零售數據應用為主，涵蓋統計分析、文字探勘、機器學習入門等技巧，分為金融市場及電子商務二大專題實作。主題如下：

## Part I 文字及社群數據分析

期中專題：財經數據分析實作 (加密貨幣及股價預測)

## Part II 行銷及零售數據分析、個人化推薦

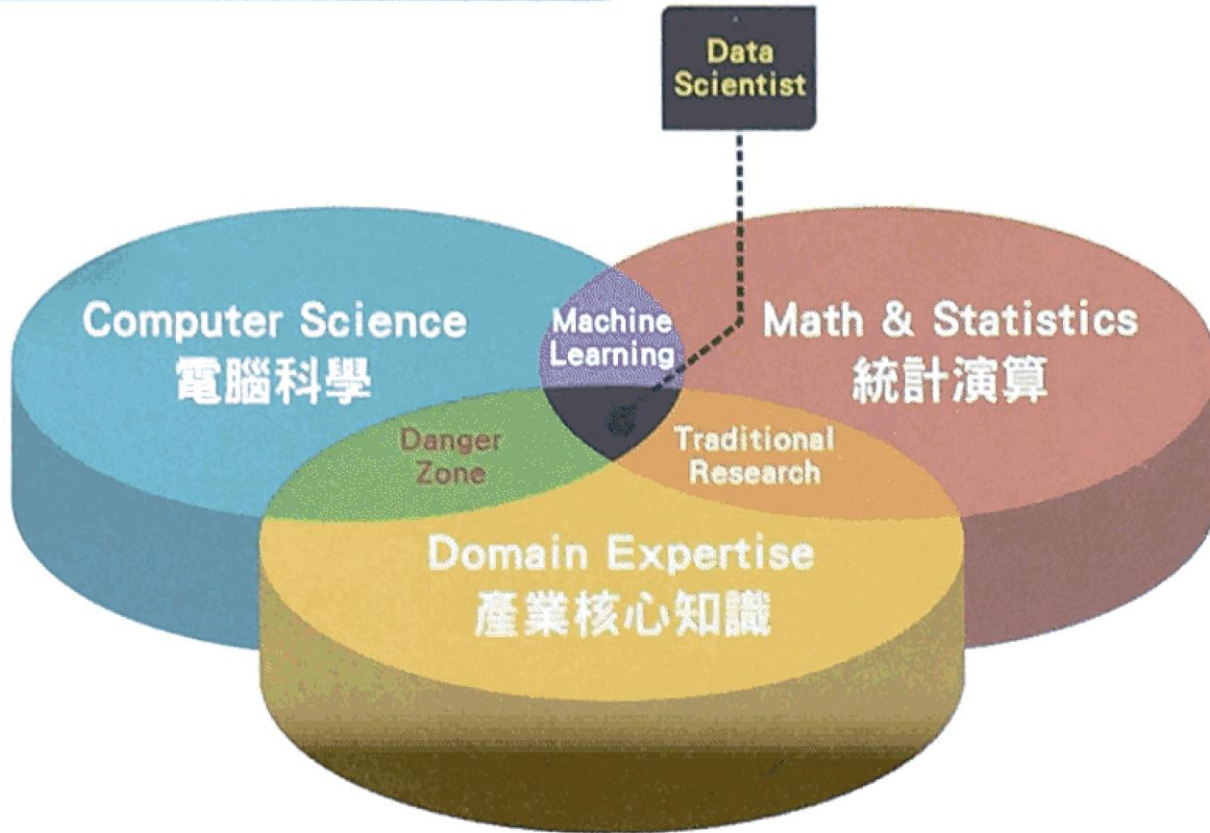
期末專題：電商數據分析實作 (客群分級及銷售預測)

# 課程介紹 (3)

- ◆ 兼具講課及實作，選修同學需可能需要投入時間、心力完成各項要求。
- ◆ 修課者最好具備管理學及統計基礎知識；需程式設計基礎或隨本課程自學相關程式語言或工具。
  - 程式語言建議採用Python、R、Java任一
  - 選配其他分析工具 (例如SAS, Weka, Tableau, 資料庫等)
- ◆ 其他請參考課程網頁

<http://homepage.ntu.edu.tw/~wyang/bda2020>

圖13：大數據需要跨界人才



數據科學家就像傳說中的獨角獸，得熟悉橫跨三大領域的知識，包含電腦科學、統計分析和產業核心知識。前兩者著重於技術層面，而產業核心知識即是大數據的哲學思考來源，愈理解產業核心，就愈不容易迷失在數據海。但要注意danger zone，擁有電腦科學和產業核心知識的人才，容易低估統計演算邏輯的重要性，而落入錯誤判斷的陷阱

Source: 大數據玩行銷 Big Data Marketing

# Introduction

## ◆ Big Data

- Data that exist in very large volumes and many different varieties (data types) and that need to be processed at a very high velocity (speed).

## ◆ Analytics

- Systematic analysis and interpretation of data—typically using mathematical, statistical, and computational tools—to improve our understanding of a real-world domain.



# Types of Analytics

## ◆ Descriptive analytics

- describes the past status of the domain of interest using a variety of tools through techniques such as reporting, data visualization, dashboards, and scorecards

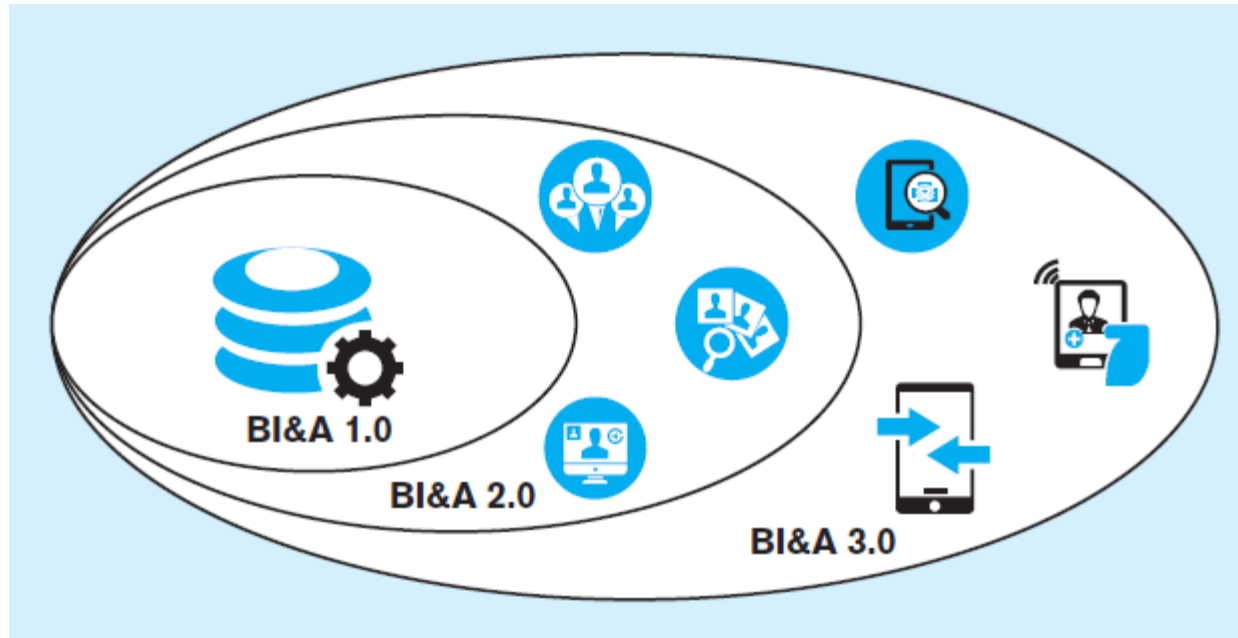
## ◆ Predictive analytics

- applies statistical and computational methods and models to data regarding past and current events to predict what might happen in the future

## ◆ Prescriptive analytics

- uses results of predictive analytics along with optimization and simulation tools to recommend actions that will lead to a desired outcome

# Figure - Generations of Business Intelligence and Analytics



Adapted from Chen et al., 2012

## **BI&A 1.0**

Focus on structured quantitative data largely from relational databases

## **BI&A 2.0**

Include data from the Web (web interaction logs, customer reviews, social media)

## **BI&A 3.0**

Include data from mobile devices, (location, sensors, etc.) as well as Internet of Things

# 2020選課介紹 (1)

- ◆ 因應防疫，本學期採遠距教學為主，進行方式如下
  - 預錄講義及解說，課前觀看 ([前往ntu cool](#))
  - 原課堂時間改為線上遠距討論問答 (約1小時)
  - 同學分成小組，自行討論及實作專題報告 (建議4~6人左右，自行分組並推選組長，成員不得全為同系所，分組報告均含互評成績)；3/18前向助教登記名單，超過期限則由助教隨機編組
  - 個人作業自行繳交，分組作業除繳交報告外，需錄製簡報影片，放上平台後彼此觀摩，並遠距討論問答

# 2020選課介紹

- ◆ 總人數上限140人，分為工管系所30人、資管系所30人、管院其他三系所(20人)、4.不限系所開放名額抽籤(60人)
- ◆ 欲加選同學請於3/6前線上登記系所及學號，由助教通知是否選上及加選授權碼
- ◆ 登記網址 <https://forms.gle/AKaQ1t1zsgoABTey9>

# » 問題討論