國立東華大學

課　綱　Course Outline

企業管理學系博士班財務金融組

中文課程名稱 Course Name in Chinese	大數據統計分析				
英文課程名稱 Course Name in English	Statistical Analysis of Big Data				
科目代碼 Course Code	BM__85030	班　別 Degree	博士班 Ph.D.		
修別 Type	選修 Elective	學分數 Credit(s)	3.0	時　數 Hour(s)	3.0
先修課程 Prerequisite					

課程目標 Course Objectives

大數據有4種特性：（1）數據量巨大；（2）數據類型多樣；(3)數據快數累積；（4）數據價值密度低，因而無法應用傳統的統計方法來分析。本課程針對大數據特性所發展的統計方法做系統性的介紹，包含大數據計算平台，架構與統計軟體；大數據統計模型的建立與分析方法；大數據分析結果的呈現、說明與視覺化；及大數據實證應用，以提昇修課學生分析大數據的統計能力。

系教育目標 Dept.'s Education Objectives

1	培育具有財金議題研究能力之學者 Cultivation of research financial talents and academic researching potentials
2	培育具有管理能力之金融專業人才 Cultivation of manage finances talents for Financial professionals
3	培育具備當代財務金融理論與實證知識的教學人才 Good manner and professional ethics, understanding of the national financial theory and practice teaching capability.

系專業能力 Basic Learning Outcomes	課程目標與系專業能力相關性 Correlation between Course Objectives and Dept.'s Education Objectives
A	具備企業管理理論的核心專業知識	○
B	具備邏輯思考與企業經營管理問題解決的能力	●
C	具備高階財務金融的分析能力	●
D	具備高階企業財務管理專業能	○

E	具備英語閱讀溝通協調及教學之能力	
F	具備前瞻開創新興議題之能力	●
G	具備金融創新之能力	●

圖示說明Illustration ：● 高度相關 Highly correlated ○中度相關 Moderately correlated

課程大綱
Course Outline

包含4個主題：1.大數據計算平台，架構與統計軟體；2.模型建立與分析方法；3.結果呈現、說明與視覺化；4.實證應用。

Topics:

1. What is big data? How different is the statistical methods for big data different from conventional statistical methods?
2. Architecture for analyzing big data
3. Introduction Big data technologies: A/B testing, crowdsourcing, data fusion and integration, genetic algorithms, machine learning, natural language processing, signal processing, simulation, time series analysis, visualisation. tensors, multilinear subspace learning, massively parallel-processing (MPP) databases, search-based applications, data mining, distributed file systems, distributed databases, cloud-based infrastructure (applications, storage and computing resources) and the Internet.
4. Visuatisation
5. Data mining
6. Text mining
7. Pattern recognition
8. Split and conquer technology
9. Statistical learning
10. Time series data mining:
Indexing, clustering, classification, prediction, anomaly detection
11. Similarity search in times series data
12. Feature-based dimensionality reduction

資源需求評估（師資專長之聘任、儀器設備的配合···等）
Resources Required (e.g. qualifications and expertise, instrument and equipment, etc.)

多部多核心電腦，投影機
Hadoop系統，R統計軟體
大數據資料庫
Handouts

課程要求和教學方式之建議
Course Requirements and Suggested Teaching Methods

1‧由授課教師或邀請講者講解大數據統計理論與實例應用。
2‧作業包含以實際的大數據，運用大數據統計方法分析與結果的視覺化。
3‧專題實作：修課學生須選擇一個有趣及重要的議題，運用課堂上所學的方法與技術，完成一份大數據分析的報告。

其他
Miscellaneous

課程成績評定之依據和標準
Grading Criteria
Held the examinations and assign the home-works
Homework, class attendance and discussion 50%
Project 50%