開課班級: 碩熱農一A **授課老師: 廖世義,江陸克** **學分數: 3**

課程大綱:

本課程主要在讓國際學生對大數據有基本應用的概念，說明大數據分析應用在農業、製造業、商業行銷、線上零售業、健康照顧和金融業等不同領域，來說明大數據實際的在各產業運用，以滿足國際學生多樣化的背景。因此本課程的目的是讓學生熟悉大數據的應用和相關分析工具，落實在不同的產業面。首先介紹大數據資料的蒐集方法、應用成果、相關的技術分析的概念、資訊倫理與安全，因大數據易面臨到道德倫理等相關議題。同時以真實的資料庫案例進行深度討論與說明大數據應用的優點及限制。第二部分為介紹幾種演算法應用在不同的實際產業，透過應用軟體STATISTICA
v13的資料探勘模組，如：購物籃分析、決策樹、分群技術等演算法，讓同學熟悉資料探索之應用。本課程期末將要求參與的同學分組進行實務上產業之調查與分析，針對同學感興趣的個案進行大數據的個案分析，撰寫一份期末專案報告，針對個案提出幾點經營管理的建議。

outline:

This course provides Big Data concept and applications of Big Data analytics in different fields such as agriculture, manufacturing, marketing, online retailing, health care and banking. The objective of this course is to familiarize student with Big Data analysis as a tool for addressing the application in different fields. The course begins with a basic introduction to big data, as well as associated technical, conceptual and ethical challenges. Strengths and limitations of big data research are discussed in depth using real-world examples. The next part is analysis implementation of actual cases by introducing and applying special algorithms in different fields, familiar with applications of data exploration software as STATISTICA (v13) and its data miner. These specific algorithms include association rules, decision tree, clustering, and classification. Students then engage in case study exercises in which small groups of students develop and present a big data concept for a specific real-world case. Attending this course, students will have an opportunity to access to real data from different industries and know how to analysis the data with problem based learning. The goal by the end of this semester is for student to have “ analytics portfolio ” consisting of data analytics skill that students can use for their future career.

教學型態:
課堂教學

成績考核方式:
平時成績:%
期中考:%
期末考:%
其它:%

本科目教學目標:

1.培育熱帶農業專業人才。 To provide professional knowledge in tropical Agriculture.
2.提升農業專業知識與獨立研究能力。 To conduct in independent research and to contribute knowledge in agriculture-related fields. **3.促進國際農業發展與技術移轉。** To train and educate senior level professionals to conduct research.

參考書目:

課程進度表：

週次	起訖月日	授課單元(內容)	備註
第1週	9.14~9.21	Introduction to Big Data * Describe the Big Data landscape including examples of real world big data problems including the three key sources of Big Data: people, organizations, and sensors. * Explain the V ’ s of Big Data (volume, velocity, variety, veracity, valence, and value) and why each impacts data collection, monitoring, storage, analysis and reporting. * Get value out of Big Data by using a 5-step process to structure your analysis.	8日正式上課。8~12日課程加退選，轉學(系)生、復學生及延修生選課，雙主修、輔系申請，12日申辦抵免學分截止日
第2週	9.21~9.28	Introduction to Big Data * Describe the Big Data landscape including examples of real world big data problems including the three key sources of Big Data: people, organizations, and sensors. * Explain the V ’ s of Big Data (volume, velocity, variety, veracity, valence, and value) and why each impacts data collection, monitoring, storage, analysis and reporting. * Get value out of Big Data by using a 5-step process to structure your analysis.	
第3週	9.28~10.05	Introduce basic about Statistica Software Introduce about Statistica v13 interface, data import, descriptive statistics and correlation… How to get open data, and import to Statistica	28日(日)孔子誕辰紀念日/教師節(放假),29日(一)補假
第4週	10.05~10.12	;Introduction real case application of big data analytics in different fields Big Data Analytics Application in agriculture, manufacturing, marketing, online retailing, health care and banking	29日成績優異提前畢業者提出申請截止日
第5週	10.12~10.19	Introduction real case application of big data analytics in different fields Big Data Analytics Application in agriculture, manufacturing, marketing, online retailing, health care and banking	6日(一)中秋節(放假)，10日(五)國慶日(放假)
第6週	10.19~10.26	Data Cleansing & Preparation; Data	14日學生宿舍安全輔導暨複

		Summarization & Visualization) Find out outlier, missing data, combing or separate data	合式防災疏散演練。18日多益測驗
第7週	10.26~11.02	Association Rules (Baskets Analysis) Detecting relationships or associations between;specific values of categorical values in large data sets. This is a common task in many data mining projects applied to databases containing records of customer transactions (e.g. Items purchased by each customer). Allow analysts and researchers to uncover hidden pattern in large data sets.;	24日(五)補假，25日(六)光復暨古寧頭大捷日(放假)。
第8週	11.02~11.09	Association Rules (Baskets Analysis) Detecting relationships or associations between specific values of categorical values in large data sets. This is a common task in many data mining projects applied to databases containing records of customer transactions (e.g. Items purchased by each customer). Allow analysts and researchers to uncover hidden pattern in large data sets.	30日校課程委員會
第9週	11.09~11.16	Midterm	3~9日期中考試
第10週	11.16~11.23	Classification and Decision Tree classification systems based on multiple covariates or for developing prediction algorithms for a target variable. This method classifies a population into branch-like segments that construct an inverted tree with a root node, internal nodes, and leaf nodes. It commonly used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal, but are also a popular tool in machine learning	13日教務會議,16日教師期中成績上網登錄截止日
第11週	11.23~11.30	Classification and Decision Tree classification systems based on multiple covariates or for developing prediction algorithms for a	

		target variable. This method classifies a population into branch-like segments that construct an inverted tree with a root node, internal nodes, and leaf nodes. It commonly used in operations research, specifically in decision analysis, to help identify a strategy most likely to reach a goal, but are also a popular tool in machine learning.	
第12週	11.30~12.07	Cluster Analysis Handling large data sets and enabling clustering of continuous and/or categorical variables, and providing the functionality for complete unsupervised learning (clustering) for pattern recognition, with all deployment options for predictive clustering.	24~28體育運動週。24日校園路跑。27日運動大會夜間開幕，28日運動大會活動，29日101週年校慶活動日，照常上班
第13週	12.07~12.14	Cluster Analysis;Handling large data sets and enabling clustering of continuous and/or categorical variables, and providing the functionality for complete unsupervised learning (clustering) for pattern recognition, with all deployment options for predictive clustering.;	
第14週	12.14~12.21	Logistics Regression To predict outcome of a categorical dependent variable on the basic of predictor variables. Logistic regression is used in various fields, including machine learning, most medical fields, and social sciences	12日申請停修課程截止日
第15週	12.21~12.28	Logistics Regression To predict outcome of a categorical dependent variable on the basic of predictor variables. Logistic regression is used in various fields, including machine learning, most medical fields, and social sciences.	
第16週	12.28~1.04	Discriminant Analysis to develop discriminant functions that are nothing but the linear combination of independent variables that will discriminate between the categories of the dependent variable in a perfect	22日校務會議。25日行憲紀念日(放假)

		manner.	
第17週	1.04~1.11	Discriminant Analysis to develop discriminant functions that are nothing but the linear combination of independent variables that will discriminate between the categories of the dependent variable in a perfect manner.	1日(四)開國紀念日(放假)
第18週	1.11~1.18	Final Exam	5~11日期末考試，10~11日學生退宿