國立台灣科技大學 114學年 第2學期　課程大綱

Spring 2026 NTUST Course Outline

授課教師：楊朝龍

Instructor:Chao-Lung Yang

課程名稱：巨量資料分析及應用

Course Title：Big Data Analytics and Applications

2026/11/3

課程代號： IM5211701 Course Code 學分數： 3 Credits	必選修：選修/半學年 Required/Electve:Elective/Half Yr. 先修課程： Prerequisites
節次教室： Time/Location	R2(MA-215) R3(MA-215) R4(MA-215)
專業核心能力： Core Professional Competencies	
課程網址： Course Website	
課程宗旨： Course Objectives	本課程介紹巨量資料的特性及相關應用，並透過實作方式演練巨量資料分析工具，以學習如何轉換巨量資料至具有競爭價值之知識及決策。課程中以商業及工業4.0相關的資料介紹實務上的案例。 This course introduces the main characteristics of Big Data and the relevant applications. By offering the practicing exercises, students can learn how to analyze data by big data analytic tools such as R and Python tools and create the valuable decision. This course also introduces the case studies regarding business innovation and industry 4.0.
課程大綱： Outline of Lectures	

巨量資料指的是以極為快速（Velocity）的方式，急速累積具大容量（Volume）及多樣性（Variety）的資料。由於其資料特性，運用傳統的資料庫分析工具及方法將變得很困難，因此分析此類的資料需要特別的工具及方法。巨量資料分析將著重於找出隱藏於巨量資料中未知的特徵、相關性及其他有用的資訊，並轉化這些資訊為具競爭優勢的知識或決策。本課程將介紹巨量資料及其相關應用，並以實作方式演練巨量資料分析工具，以提供對巨量分析的了解。本課程將介紹下列課題：

1. 巨量資料介紹
2. 巨量資料架構
3. 巨量資料應用
4. 巨量資料處理
5. 巨量資料探勘
6. 商業大數據相關案例
7. Hadoop系統簡介
8. MongoDB 系統簡介
9. 大型語言模型介紹
10. 檢索增強生成 Retrieval-Augmented Generation (RAG) 介紹

Big Data is a collection of datasets that are accumulated in high Velocity with super large Volumes containing a variety of data types. It is very difficult to process or analyze Big Data by using traditional database management tools or data processing applications. Big Data Analytics is the process of examining large amounts of data of a variety of types (big data) to uncover hidden patterns, unknown correlations and other useful information. Transforming Big Data to the valuable information or decision can provide competitive advantages and result in business benefits. This course will introduce Big Data and its application. Students will work on the hand-on project to learn the Big Data concept and applications. This course will cover the following topics.

1. Introduction of Big Data
2. Big Data Structure
3. Big Data Application
4. Big Data Manipulation
5. Data Mining on Big Data
6. Cast studies regarding business innovation
7. Introduction of Hadoop System
8. Introduction of MongoDB System
9. Introduction to Large Language Model,
10. Introduction to Retrieval-Augmented Generation (RAG)

授課方式：
Method of Instruction

講授 Lecture：40%

分組討論 Group discussion：20%

案例研討 Case study：10%

操做練習 Practical exercises：30%

講授 Lecture：Lecture Notes/Assignments/Readings
Lecture pdft, homework, and paper reading assignments will be provided at the class website of Moodle system. You will be responsible for downloading them to prepare for class and homework.
Supplementary Readings
Extensive reading materials regarding the term project will be provided at the class website to develop a broad understanding of research.%

教科書：
Textbooks

Lecturer will provide the materials.

參考書目：
References

	Paul Deitel and Harvey Deitel, Intro to Python for Computer Science and Data Science – Learning to Program with AI, Big Data and the Cloud, Global Edition, Pearson Education Limited Viktor Mayer-Schonberger, and Kenneth Cukier, “Big Data: A Revolution That Will Transform How We Live, Work, and Think,” ISBN-10: 9780544002692, Eamon Dolan/Houghton Mifflin Harcourt; 1 edition (March 5, 2013) Wes McKinney, “Python for Data Analysis.”, ISBN: 978-1-449-31979-3, O’ REILLY
修課須知： Notice	Your course grade will be based on homework assignments, projects, and a final term project. ・ No credit is given for exam if empty answer is received. ・ Partial grades may be given if the structure of answer is completed. Instructor and grader reserve the right to determine if the partial grade is given.
評量方式： Grading	Category Percentage Homework and Paper Reading Assignments (individual) 30% Analytic Projects (teamwork) 30% Final Project – report and presentation (Teamwork with EMBA students) 40% Total 100%
備註說明： Notes	Prerequisite Programming skill Basic competency in computer programming is REQUIRED. You are expected to write some codes to analyze the data in this class. The programming skills such as Java, R, Python, or Perl are preferred but not required. Any of computer programming such as C#/C++/C, VB, PHP, Matlab, are MUST.