澳門理工大學
Universidade Politécnica de Macau
Macao Polytechnic University

FACULTY OF APPLIED SCIENCES

MASTER OF SCIENCE IN BIG DATA AND INTERNET OF THINGS

LEARNING MODULE OUTLINE

Academic Year	2024/2025	Semester	1
Module Code	COMP6132		
Learning Module	Introduction to Big Data		
Pre-requisite(s)	Nil		
Medium of Instruction	English		
Credits	3	Contact Hours	45 hrs
Instructor	Wuman Luo	Email	luowuman@mpu.edu.mo
Office	Rm. A323, Chi Un Building	Office Phone	8599-6321

MODULE DESCRIPTION

This learning module covers the characteristics of Big Data, the sources of massive data in enterprises and sensor networks, and the challenges in data preparation, data storage and analytic processing. The students will acquire skills and working knowledge of the Big Data tools and technologies. This module focuses on the planning, designing and implementing Big Data solutions. Examples and exercises of Big Data systems are used to provide hands-on experiences in the workings of major components in Big Data solutions. The students will also be able to integrate the Big Data tools to form coherent solutions for business problems. Finally, other topics relevant to the area of Big Data are presented, such as alternative large-scale processing platforms, non-relational data stores, and Cloud Computing execution infrastructure.

MODULE INTENDED LEARNING OUTCOMES (ILOS)

On completion of this learning module, students will be able to:

M1.	Identify the characteristics and challenges of Big Data in a range of complex application domains; (AHEP4-M2)
M2.	Model and implement efficient Big Data solutions by integrating various advanced technologies; (AHEP4-M3, AHEP4-M5)
M3.	Analyze the trade-offs in Big Data processing system design in complex infrastructures to handle a wide range of application problems; (AHEP4-M2, AHEP4-M3)
M4.	Demonstrate an understanding of non-relational databases in Big Data analytical pipelines. (AHEP4-M2, AHEP4-M3)

These ILOs aims to enable students to attain the following Programme Intended Learning Outcomes (PILOs):

PILOs		**M1**	**M2**	**M3**	**M4**
P1.	Master the principles of system engineering and relevant enabling technologies for building of IoT solutions				
P2.	Critically evaluate scientific methodologies and mathematical models for Big Data and its applications			✓	
P3.	Master the advanced software and programming tools and techniques for IoT solutions and Big Data		✓		
P4.	Explain the processes involved in IoT solutions and Big Data analytics in a typical business setting			✓	✓
P5.	Explain different application domains and analyze their requirements for IoT and Big Data	✓			✓
P6.	Apply knowledge in advanced communication and multimedia technologies for the design and implementation of IoT solutions				
P7.	Apply knowledge in applied statistics, machine learning, leading-edge technologies and programming techniques for Big Data		✓	✓	
P8.	Design and carry out an advanced project following an ethical and professional methodology				
P9.	To demonstrate advanced knowledge and R&D techniques in Big Data and IoT				
P10.	To investigate and develop new, emerging ICT technology for Big Data and IoT				
P11.	To develop a global vision on the critical development and new application of Big Data and IoT				
P12.	To communicate technically and effectively in both speaking and writing				
P13.	To have a positive attitude towards society and the environment.				
P14.	To adhere to high moral standards and commit to excellence in life-long learning.				

MODULE SCHEDULE, COVERAGE AND STUDY LOAD

Week	Content Coverage	Contact Hours
1-3	1. Overview	9 hours
	1.1 Sources of Big Data	
	1.2 Characteristics of Big Data	
	1.3 Data Science and Data Analytics	
	1.4 Data Preprocessing	
4-7	2. Scalable Computing Systems	12 hours
	2.1 Distributed File Systems and Commodity Clusters	
	2.2 Parallel Programming Models	
	2.3 HDFS: Blocks and Replication, Read / write operations	
	2.4 Apache Hadoop System	
	2.5 Big Data Architecture and Data Pipeline	
8-11	3. Large-scale Data Processing with Spark	10 hours
	3.1 The Scala Programming Language	
	3.2 The Apache Spark Analytics Engine	
	3.3 Resilient Distributed Datasets (RDD)	
	3.4 RDD Operations: Transformations and Actions	
11-13	4. Structured Data Analytics	7 hours
	4.1 Handling Structured and Semi-Structured Data	
	4.2 Big Data File Formats and View Materialization	
13-15	5. NoSQL Databases	7 hours
	5.1 Features and Design Goals	
	5.2 Data Modelling and Querying	

TEACHING AND LEARNING ACTIVITIES

In this learning module, students will work towards attaining the ILOs through the following teaching and learning activities:

Teaching and Learning Activities	M1	M2	M3	M4
T1. Lectures	✓	✓	✓	✓
T2. Labs/Practices		✓	✓	

ATTENDANCE

Attendance requirements are governed by the Academic Regulations Governing Master's Degree Programmes of the Macao Polytechnic University. Students who do not meet the attendance requirements for the learning module shall be awarded an 'F' grade.

ASSESSMENT

In this learning module, students are required to complete the following assessment activities:

Assessment Activities	Weighting (%)	AHEP4 LOs	ILOs to be Assessed
A1. Assignment(s)	40%	AHEP4-M2, AHEP4-M5	M2, M3
A2. Test(s)	20%	AHEP4-M2, AHEP4-M3	M1, M3, M4
A3. Examination	40%	AHEP4-M2, AHEP4-M3, AHEP4-M5	M1, M3, M4

The assessment will be conducted following the University's Assessment Strategy (see www.mpu.edu.mo/teaching_learning/en/assessment_strategy.php). Passing this learning module indicates that students will have attained the ILOs of this learning module and thus acquired its credits.

Students with an overall score of less than 35 in the coursework will fail the module even if the overall score for the module is 50 or above.

Students with a score of less than 35 in the final examination will fail the module even if the overall score for the module is 50 or above.

REQUIRED READINGS

There is no official required readings for this module. Module notes are distributed in the class.

REFERENCES

1. T. Erl, W Khattak (2016). Big Data Fundamentals: Concepts, Drivers & Techniques, ServiceTech.
2. T. White (2015). Hadoop: The Definitive Guide: Storage and Analysis at Internet Scale, O'Reilly.
3. M. Guller (2015). Big Data Analytics with Spark, Apress.

STUDENT FEEDBACK

At the end of every semester, students are invited to provide feedback on the learning module and the teaching arrangement through questionnaires. Your feedback is valuable for instructors to enhance the module and its delivery for future students. The instructor and programme coordinators will consider all feedback and respond with actions formally in the annual programme review.

ACADEMIC INTEGRITY

The Macao Polytechnic University requires students to have full commitment to academic integrity when engaging in research and academic activities. Violations of academic integrity, which include but are not limited to plagiarism, collusion, fabrication or falsification, repeated use of assignments and cheating in examinations, are considered as serious academic offenses and may lead to disciplinary actions. Students should read the relevant regulations and guidelines in the Student Handbook which is distributed upon the admission into the University, a copy of which can also be found at www.mpu.edu.mo/student_handbook/.