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

SFACULTY OF APPLIED SCIENCES

BACHELOR OF SCIENCE IN ARTIFICIAL INTELLIGENCE

LEARNING MODULE OUTLINE

Academic Year	2024/2025	Semester	2
Module Code	CSAI2123		
Learning Module	Introduction to Data Science		
Pre-requisite(s)	-		
Medium of Instruction	English		
Credits	3	Contact Hours	45 hrs
Instructor	Dr. Jie Zhang	Email	jpeter.zhang@mpu.edu.mo
Office	Room N56, 5/F, WUI CHI Building, Main Campus	Office Phone	8599-6837

MODULE DESCRIPTION

This module offers a comprehensive introduction to the dynamic field of Data Science, encompassing foundational concepts, practical applications, and essential tools for data analysis. Students will embark on a journey through the core elements of data science, exploring key architectures, applications, and the pivotal role of Python in data science workflows. Subsequently, they will delve into the realm of data manipulation and computation using NumPy and Pandas, gaining proficiency in handling and processing data efficiently. The module further immerses students in the art of data visualization, showcasing techniques with Matplotlib, Seaborn, and Plotnine to effectively communicate insights derived from data. Finally, students will unravel the intricacies of various data analysis algorithms including clustering, dimensionality reduction, association rule mining, classification, and regression techniques, empowering them to extract valuable patterns and make data-driven decisions with confidence. Through this module, students will acquire a solid foundation in data science, enabling them to navigate complex data landscapes and derive meaningful insights to drive business success.

MODULE INTENDED LEARNING OUTCOMES (ILOS)

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

M1.	Analyse domain-specific problems and formulate them into data science tasks; (C1, C2, C5)
M2.	Evaluate and select suitable techniques and tools for data science tasks; (C3)
M3.	Use practical skills to apply key concepts, techniques and tools for discovering, analysing, visualizing, and presenting data; (C12, C13)
M4.	Complete a final project to showcase their data analysis skills. (C2, C3, C4, C5, C6)

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

PILOs		M1	M2	M3	M4
P1.	Select and apply proven methods, tools and techniques to the effective and efficient implementation of information systems on common platforms, including the Internet platform;		✓	✓	✓
P2.	Acquire essential knowledge in specific fields of artificial intelligence, including machine learning, computer vision and natural language processing;	✓	✓	✓	
P3.	Apply necessary mathematical techniques to model, analyse and devise solutions to complex problems;	✓			
P4.	Work independently to develop an understanding of, and the knowledge and skills associated with the general support and mitigation of security risks of computer systems and networks;			✓	✓
P5.	Design and implement both relational and non-relational data stores, with an emphasis on how to organise, maintain, retrieve and analyse information;				
P6.	Distinguish the fundamental and operational issues of computer systems and artificial intelligence applications, with considerations of user, business, ethical, societal and environmental needs;	✓			
P7.	Evaluate, prepare and communicate effectively on technical information to both technical and non-technical audience;				✓
P8.	Work as an effective member of a team in the analysis, design and development of software systems, with recognition of requirement to support equality, diversity and inclusion;				
P9.	Use project planning, risk management and quality management techniques in solutions to complex problems;				
P10.	Build the capacity and desire for lifelong learning and to learn advanced and emerging technologies on one's own.				✓

MODULE SCHEDULE, COVERAGE AND STUDY LOAD

Week	Content Coverage	Contact Hours
1-2	**1. Overview of Data Science**	6
	1.1. Overview of Data Science	
	1.2 Architectures	
	1.3. Applications	
	1.4. Introduction to Python for Data Science	
3-6	**3. Data Manipulation and Computation**	12
	3.1. Introduction to NumPy	
	3.2. Introduction to Pandas	
	3.3 Data Manipulation and Computation with NumPy and Pandas	

7-10	**2. Data Visualization**	12
	2.1. Overview of Data Visualization	
	2.2. Data Visualization with Matplotlib	
	2.3. Data Visualization with Seaborn	
	2.4. Data Visualization with Plotnine	
11-15	**4. Data Analysis Algorithms**	15
	4.1. Clustering Algorithms	
	4.2. Dimensionality Reduction Algorithms	
	4.3. Association Rule Mining Algorithms	
	4.4 Classification Algorithms	
	4.5 Regression Algorithms	

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. In-class tutorials and exercises		✓	✓	✓

ATTENDANCE

Attendance requirements are governed by the Academic Regulations Governing Bachelor'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	20	C12, C13	M3
A2. Project	20	C2, C3, C4, C5, C6	M4
A3. Test	20	C1, C2	M1, M2
A4. Examination	40	C1, C2	M1, M2

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 must take the re-sit examination even if the overall score for the module is 50 or above.

Students with a score of less than 35 in the final examination must take the re-sit examination even if the overall score for the module is 50 or above.

Students with an overall final grade of less than 35 are NOT allowed to take the re-sit examination.

REQUIRED READINGS

There is no official text for this module. Module notes are distributed in the module.

Textbook(s)

1. Jake VanderPlas. (2017) Python Data Science Handbook: Essential Tools for Working with Data. M O'Reilly Media. https://github.com/jakevdp/PythonDataScienceHandbook
2. Wes McKinney (2017) Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython 2nd Edition. O'Reilly Media. https://github.com/wesm/pydata-book

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/.