



FACULTY OF APPLIED SCIENCES
BACHELOR OF SCIENCE IN ARTIFICIAL INTELLIGENCE
LEARNING MODULE OUTLINE

Academic Year	2025/2026	Semester	1
Module Code	CSAI3122		
Learning Module	Natural Language Processing		
Pre-requisite(s)	CSAI2121 Probability and Statistics		
Medium of Instruction	English		
Credits	3	Contact Hours	45 hrs
Instructor	Dr Patrick Pang	Email	patrickpang@mpu.edu.mo
Office	N46B, Wui Chi Building, Main Campus	Office Phone	8599-6886

MODULE DESCRIPTION

A substantial amount of data is stored in the form of text today. Examples include web pages, social media posts, instant messaging, legal documents, etc. Such unstructured text creates many challenges in understanding and harnessing knowledge within. In this module, students will learn basic knowledge of natural languages and computational approaches for working with text. Students will also develop an understanding of the main algorithms of natural language processing (NLP) and various NLP applications, such as sentiment analysis, text mining and topic modelling.

MODULE INTENDED LEARNING OUTCOMES (ILOS)

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

M1.	Apply basic NLP concepts into some basic tasks of textual data processing; (C1)
M2.	Identify basic challenges of the computational modelling of natural language; (C2)
M3.	Analyse and explain the mathematical and/or algorithmic basis of common techniques used in natural language processing; (C2, C3)
M4.	Design and construct programmes to processing natural languages with existing libraries. (C3, C5, C6, C12)

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	1. Introduction to Natural Language Processing	3
2-3	2. Basic Text Processing	6
	2.1. Words and Corpora	
	2.2. Tokenization	
	2.3. Stemming and Lemmatization	
4-5	3. Regular Expressions (RE)	6
	3.1. Regular Expressions	
	3.2. RE and Finite State Automata	
6	4. N-gram Language Modeling	3
7	5. Naïve Bayes and Text Classification	3
8	6. Parts-Of-Speech (POS) Tagging	3
9-10	7. Word Vectors	6
	7.1. Vector Semantics and Embeddings	
	7.2. TF-IDF	
	7.3. word2vec	



11-12	8. Neural Networks and Sentiment Analysis	6
13	9. Transformers	3
14	10. Large Language Models	3
15	11. Natural Language Processing for Non-English Text	3

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/Classwork	20%	C1, C5, C12	M1, M4
A2. Project	20%	C1, C2, C5, C6, C12	M1, M2, M3, M4
A3. Test	20%	C1, C2, C3	M1, M2, M3
A4. Examination	40%	C1, C2, C3	M1, M2, M3

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

1. Daniel Jurafsky and James H. Martin. (2025). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models*, 3rd edition. <https://web.stanford.edu/~jurafsky/slp3>.

REFERENCES

1. John Atkinson-Abutridy. (2022). *Text Analytics: An Introduction to the Science and Applications of Unstructured Information Analysis*. Chapman & Hall.
2. Manika Lamba and Margam Madhusudhan. (2022). *Text Mining for Information Professionals: An Uncharted Territory*. Springer.
3. Steven Bird, Ewan Klein, Edward Loper. (2009). *Natural Language Processing with Python – Analyzing Text with the Natural Language Toolkit*. O'Reilly.

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/.