Title (Units): **COMP4136 Natural Language Processing and Large Language Models (3,2,1)**

Course Aims: To introduce some essential knowledge of natural language processing and Large Language Models (LLMs), such as fundamental concepts, critical thinking of methodologies, application scenarios, practical techniques and tools. Students after taking this course will be able to: 1) identify and apply advanced techniques of natural language processing to process textural information; 2) build modules to design, implement and evaluate effective natural language systems in real-world contexts; and 3) gain insights into Large Language Models, train and use LLMs to solve practical problems.

Prerequisite: COMP3057 Introduction to AI and ML OR
COMP3117 Applied Artificial Intelligence OR
COMP3115 Exploratory Data Analysis and Visualization

Course Intended Learning Outcomes (CILOs):
Upon successful completion of this course, students should be able to:

No.	Course Intended Learning Outcomes (CILOs)
	Knowledge
1	Describe the fundamental concepts and methodologies of natural language processing and large language models
2	Explain the advantages and limitations of methods developed for different scenarios
3	Identify relevant textural information processing techniques to meet real-world needs
	Professional Skill
4	Apply specific methods and techniques in a number of natural language applications
5	Evaluate the solutions designed to technical problems

Calendar Description: This course introduces some essential knowledge of natural language processing and Large Language Models, such as fundamental concepts, critical thinking of methodologies, application scenarios, practical techniques and tools. Students will be given the opportunities to appreciate the needs and impacts of technical problem-solving with data and to develop real-world applications.

Teaching and Learning Activities (TLAs):

CILOs	Type of TLA
1 - 2	Students will learn essential concepts and principles of natural language processing and large language models through lectures and tutorials. Besides, written assignments, quizzes and final examination will be designed to evaluate the students' level of understanding.
2 - 3	Students will learn critical algorithms and techniques of some downstream tasks and applications through lectures and tutorials. Laboratory sessions will also be designed so that students could apply what they have learnt in lectures. There will include laboratory exercises and quizzes.
4 - 5	Students are required to conduct a project based on a selected topic about NLP and large language models and give a formal presentation on their proposed method. Peer evaluation by Instructor(s) and teaching assistant(s) will be involved.

Assessment:

No.	Assessment Methods	Weighting	CILOs to be addressed	Description of Assessment Tasks
1	Assessments, Quizzes and Labs	30%	1 - 4	Assignments, quizzes and labs will be used to consolidate their knowledge and develop their skills in natural language processing and large language models.

2	Project	30%	3 - 5	Individual/group project will further strengthen their understanding and problem-solving skills. Peer evaluation will be involved.
3	Examination	40%	1 - 5	Examination will be used to assess students' overall understanding in the concepts, algorithms, methodologies and their ability in applying these knowledge to solve problems.

Assessment Rubrics:

Excellent (A)	• Achieve all CILOs, demonstrating a good mastery of both the theoretical and practical aspects of the knowledge and skills associated with natural language processing and large language models • Able to develop correct solutions to problems, accompanied by critical thinking, analytical thinking and creative thinking • Demonstrate a thorough understanding and solid knowledge of natural language processing and large language models • Able to apply a variety of techniques and relevant knowledge for fulfilling the real-world needs
Good (B)	• Achieve most of the five CILOs, demonstrating a good understanding of the concepts and underlying methodologies • Able to develop correct solutions to problems, accompanied by adequate explanations • Demonstrate a competent level of knowledge of natural language processing and large language models • Able to make use of appropriate techniques and knowledge and apply them to familiar situations and problems
Satisfactory (C)	• Achieve some of the five CILOs, demonstrating a basic level of understanding of the concepts and underlying methodologies • Able to provide acceptable solutions to problems • Demonstrate an adequate level of knowledge of natural language processing and large language models • Able to make use of some techniques and knowledge and apply them to familiar situations
Marginal Pass (D)	• Achieve few of the five CILOs, with minimal understanding of the associated concepts and underlying methodologies • Able to provide solutions to simple problems • Demonstrate a basic level of knowledge of natural language processing and large language models • Able to make use of limited knowledge or techniques and apply them to some simple cases
Fail (F)	• Achieve none of the five CILOs, with little understanding of the associated concepts and underlying methodologies • Unable to provide solutions to simple problems • Knowledge of natural language processing and large language models falling below the basic minimum level • Unable to apply techniques and knowledge to situations or problems

Course Content and CILOs Mapping:

Content		**CILO No.**
I	Introduction to Natural Language Processing (NLP) and Core Concepts	1
II	NLP Models and Techniques	1, 2
III	Deep Neural Networks and Transformers for NLP	1, 2
IV	Large Language Models (LLMs) and Fundamental Algorithms for NLP	1, 2, 3
V	NLP Applications	3, 4, 5

References:

• Daniel Jurafsky and James H. Martin. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition with Language Models, 3rd edition. Prentice Hall. 2025.
• Christopher D. Manning and Hinrich Schütze. Foundations of Statistical Natural Language Processing. The MIT Press. 1999.
• Jason Brownlee. Deep Learning for Natural Language Processing. Machine Learning Mastery, 2018.
• Steven Bird, Ewan Klein, and Edward Loper. Natural Language Processing with Python. 1st edition, O'Reilly Media; 2009.
• Lane H and Dyshel M. Natural Language Processing in Action. Simon and Schuster. 2025.
• Raschka S. Build a Large Language Model (From Scratch). Simon and Schuster. 2024.

Course Content:

Topic

I. Introduction to Natural Language Processing (NLP) and Core Concepts

II. NLP Models and Techniques
1. Segmentation, Word-level Analysis, N-gram Language Models
2. POS Tagging, Syntactic Parsing
3. Data Mining in NLP
4. NLP Tools

III. Deep Neural Networks and Transformers for NLP
1. Word Embedding
2. Neural Architectures for Sequential Text Processing
3. Hierarchical and Structured Models for Language
4. Attention Mechanisms and Transformer Models for Text Representation, Understanding and Generation

IV. Large Language Models (LLMs) and Fundamental Algorithms for NLP
1. Neural Language Models
2. Pre-training and Fine-tuning
3. Model Alignment and Task-Specific Adaptation in NLP
4. Overview of Foundation Models (GPT, LLaMA, DeepSeek and others)

V. NLP Applications
1. Sentiment Classification
2. Machine Translation
3. Question Answering and Chatbots
4. Text Summarization