Title (Units): **COMP7045 Natural Language Processing (3,2,1)**

Course Aims: To give students a solid background of natural language processing and a wide-range of downstream applications. To learn creative designs and implementations, including critical thinking of methodologies, problem analytics, practical techniques and tools for text processing. Students after taking this course will be able to: 1) identify and creatively apply advanced techniques of natural language processing to solve downstream problems; and 2) build modules to creatively design natural language systems, and critically analyze their effectiveness.

Prerequisite: Nil (but some knowledge in artificial intelligence is preferable)

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
2	Identify relevant textural information processing techniques to meet real-world needs
3	Explain the advantages and limitations of methods developed for different tasks
	Professional Skill
4	Apply specific methods and techniques in a wide-range of downstream applications
5	Design and evaluate the solutions to technical problems
	Attitude
6	Engage in life-long independent learning for problem-solving and advance state-of-the art solutions

Calendar Description: This course is to provide an in-depth knowledge of natural language processing and a wide-range of downstream applications. Topics include: fundamental concepts, conceptual modeling of a natural language, part-of-speech tagging, syntactic parsing, critical thinking of downstream applications, and deep learning techniques for focused problems. The students will have a thorough understanding of natural language processing technique, and be able to creatively design models and critically analysis their effectiveness after taking this course.

Teaching and Learning Activities (TLAs):

CILOs	Type of TLA
1-5	Students will learn essential knowledge of natural language processing through lectures and tutorials. There will be written assignments, quizzes and final examination to evaluate the students' level of understanding.
4-5	Students will learn critical algorithms and techniques of some traditional problems through lectures and tutorials. Laboratory sessions will also be designed so that students could apply what they have learnt in lectures. There will be laboratory exercises and quizzes
4-6	Students are required to conduct a project based on a selected NLP topic individually and give a formal presentation on their proposed method. Instructor(s), teaching assistant and other students would ask questions related to their project.

Assessment:

No.	Assessment Methods	Weighting	CILOs to be addressed	Description of Assessment Tasks
1	Continuous Assessment – assignments and quizzes/tests	25%	1-6	Assignments, quizzes and labs will be used to consolidate their knowledge and develop their skills in natural language processing.
2	Continuous Assessment – project	25%	1-6	Individual/group project and report will further strength their understanding and problem solving skills.

3	Examination	50%	1-6	Examination will be used to assess students' overall understanding in the concepts, and their ability in applying these concepts 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 textural information processing and techniques application • Able to develop correct solutions to problems, accompanied by critical thinking, analytical thinking and creative thinking • Demonstrate a thorough understanding and solid knowledge of textural analytics, concepts, methodologies, and techniques • Able to apply a variety of techniques and relevant knowledge for fulfilling the real-world needs
Good (B)	• Achieve the first 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 textural analytics, concepts, methodologies, and techniques • Ability to make use of appropriate techniques and knowledge and apply them to familiar situations and problems
Satisfactory (C)	• Achieve most of the first 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 • Ability to make use of some techniques and knowledge and apply them to familiar situations
Fail (F)	• Achieve less than three of the CILOs, with little understanding of the associated concepts and underlying methodologies • Unable to provide solutions to simple problems • Knowledge of natural language processing 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	Text Preprocessing Techniques	1,2
III	Language Modeling and Syntactic Analysis	1,2,3
IV	Part-Of-Speech Tagging	1,3,5
V	Syntactic Parsing	1,3,5
VI	Downstream Applications of NLP Techniques	4,5,6
VII	Natural Language Processing in the Era of Deep Neural Networks	5,6

References:

- Daniel Jurafsky and James H. Martin. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition. Prentice Hall. 2000.
- 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

Course Content:

Topic

I. Introduction to natural language processing (NLP) and Core Concepts

II. Text Preprocessing Techniques
   A. Text Preprocessing: Tokenizing, Segmentation, Lemmatization, stemming, etc
   B. Regular expression

III. Language Modeling and Syntactic Analysis
   A. N-grams
   B. Smoothing for Language model
   C. Term Weighting and Vector Space Model

IV. Part-Of-Speech Tagging

V. Syntactic Parsing

VI. Downstream Applications of NLP Techniques
   A. Sentiment classification
   B. Machine translation
   C. Question answering
   D. Summarization

VII. Natural Language Processing in the Era of Deep Neural Networks
   Selected NLP Tasks with Deep Learning
   A. Word embedding (Word2Vec)
   B. Sentiment analysis with Recurrent Neural Network
   C. Sentence representation with convolutional Neural Network
   D. Parsing with Recursive Neural Network
   E. Attention Mechanism for Machine Translation