CS4493: NATURAL LANGUAGE PROCESSING

New Syllabus Proposal

Effective Term
Semester B 2024/25

Part I Course Overview

Course Title
Natural Language Processing

Subject Code
CS - Computer Science
Course Number
4493

Academic Unit
Computer Science (CS)

College/School
College of Computing (CC)

Course Duration
One Semester

Credit Units
3

Level
B1, B2, B3, B4 - Bachelor's Degree

Medium of Instruction
English

Medium of Assessment
English

Prerequisites
CS4487 Machine Learning

Precursors
Nil

Equivalent Courses
Nil

Exclusive Courses
Nil

Part II Course Details

Abstract

This course introduces algorithms and techniques for natural language processing, from computational linguistics for text processing to information extraction for language understanding. The topics include statistical and neural based language modeling, word representation, and pretrained language models such as BERT and GPT. Basic and advanced natural language processing tasks, such as machine translation, dialog systems, question answering, text classification/labeling/tagging, and knowledge graph will also be introduced.

Course Intended Learning Outcomes (CILOs)

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Describe statistical and neural processing for syntactic analysis of text properties.			x	
2	Discuss information extraction for understanding of text semantics.			x	
3	Demonstrate the effectiveness of natural language processing and understanding for real-world problems.		x		
4	Apply techniques in natural language processing and understanding for innovative applications.		x	x	x

A1: Attitude
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: Ability
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to real-life problems.

A3: Accomplishments
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

Learning and Teaching Activities (LTAs)

	LTAs	Brief Description	CILO No.	Hours/week (if applicable)
1	Lecture	Students will focus on the algorithms and techniques for natural language processing and understanding. Related real-world applications such as machine translation and dialog systems will be introduced.	1, 2, 3	3 hours per week

2	Tutorial	Students will work on a different problem set each week during the tutorial sessions, through which they can discover the main characteristics of different natural language processing techniques and integrate them for real-world problems. They will also be invited to present their solutions, and the class will be encouraged to provide comments.	1, 2, 3	8 hours per semester
3	Assignments	The students will implement selected natural language processing and understanding approaches, apply these approaches to real-world problems, and interpret the results. In this way, students can analyse the performance of different approaches.	1, 2, 3	
4	Project	The students will create a new system design and implement appropriate natural language processing approaches for innovative applications. The students will apply the principles they have learnt from the course for their design.	4	

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks (e.g. Parameter for GenAI use)
1	Assignments	1, 2, 3	20	Expect to have two assignments with programming elements for algorithm implementation.
2	Project	4	20	Around 1-3 students in a group to finish the project.

Continuous Assessment (%)

40

Examination (%)

60

Examination Duration (Hours)
2

Minimum Examination Passing Requirement (%)
30

Additional Information for ATs
For a student to pass the course, at least 30% of the maximum mark for the examination must be obtained.

Assessment Rubrics (AR)

Assessment Task
1. Assignments

Criterion
The ability to implement and assess the effectiveness of different algorithms and techniques.

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal level

Assessment Task
2. Project

Criterion
The ability and creativity in designing and implementing appropriate algorithms and techniques for innovative applications.

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal level

Assessment Task

3. Examination

Criterion

The extent to which the students can understand the algorithms and techniques, apply them with appropriate modification or design new solutions for different applications, and evaluate their performances.

Excellent (A+, A, A-)

High

Good (B+, B, B-)

Significant

Fair (C+, C, C-)

Moderate

Marginal (D)

Basic

Failure (F)

Not even reaching marginal level

Part III Other Information

Keyword Syllabus

NLP basics: language models, word representations (distributional representation and contextualized representation), recurrent neural networks, attention mechanism, transformers, and pretrained language models (BERT, GPT, etc.).

NLP task examples: machine translation, dialog systems, question answering, text classification/tagging/ labeling, knowledge graph, and other advanced topics.

Reading List

Compulsory Readings

	Title
1	Daniel Jurafsky and James H. Martin, Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, 3rd edition (online), 2020.
2	Christopher D. Manning, Hinrich Schutze, Foundations of Statistical Natural Language Processing, MIT Press, 1999.
3	Jacob Eisenstein, Natural Language Processing, online, 2018.

Additional Readings

	Title
1	Nil