

LT3233: COMPUTATIONAL LINGUISTICS

Effective Term

Semester B 2023/24

Part I Course Overview

Course Title

Computational Linguistics

Subject Code

LT - Linguistics and Translation

Course Number

3233

Academic Unit

Linguistics and Translation (LT)

College/School

College of Liberal Arts and Social Sciences (CH)

Course Duration

One Semester

Credit Units

3

Level

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

Medium of Instruction

English

Medium of Assessment

English

Prerequisites

(i) LT2231 Introduction to Language Technology (applicable to students of BA in Linguistics and Language Applications), OR (ii) CS2311 Computer Programming (applicable to students of BSc. in Data Science and students of BEng in Data and Systems Analytics), OR (iii) MS3111 Quantitative Business Analysis with Visual Basic for Applications or CS2360 Java Programming or IS2240 Python Programming for Business (applicable to students of BBA in Business Analysis)

Precursors

LT2229 Fundamentals of Linguistics or LT2290 Introduction to Language Studies

Equivalent Courses

CTL3233 Computational Linguistics

Exclusive Courses

Nil

Part II Course Details

Abstract

This course aims at introducing students to some of the major issues and solutions in natural language processing. The underlying computational properties of natural languages are considered at the lexical, syntactic, and semantic level from linguistic and statistical perspectives. Modern corpus-based deep learning techniques will be discussed. Selected natural language applications will also be introduced. Concepts taught in class will be reinforced by hands-on practical exercises.

Course Intended Learning Outcomes (CILOs)

CILOs		Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Identify the major areas of study in computational linguistics and natural language processing (NLP)	30	x	x	
2	Explain the major issues in NLP and discuss, competently and critically, computer programming for different approaches to their solution in general and with particular reference to English and Chinese	40	x	x	x
3	Write computer programs to compile and use lexical, syntactic and semantic resources to tackle various NLP subtasks	30		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	Lectures to explain the major issues in NLP and introduce computer programming for different approaches to their solution. Interaction between teacher and students is expected.	1, 2	3 hours
2	Demonstration of computer programming for handling various NLP subtasks to students in lectures and/or tutorials.	2	

3		Teacher-facilitated class/ group discussions on the technical issues and the strengths and weaknesses of different approaches to NLP subtasks in lectures and/or tutorials.	2	
4		In-class hands-on exercises on computer programming to handle various NLP subtasks, which might involve the design and preparation of various linguistic resources (e.g. writing context-free rules for parsing) and/or simple program fragments. (We assume that the students' main working programming language is Python.)	3	

Assessment Tasks / Activities (ATs)

ATs	CILO No.	Weighting (%)	Remarks (e.g. Parameter for GenAI use)	
1	Homework assignments on the concepts of computer programming and on the major issues in natural language processing.	1, 2, 3	50	

Continuous Assessment (%)

50

Examination (%)

50

Examination Duration (Hours)

2

Assessment Rubrics (AR)**Assessment Task**

1. Homework Assignments

Criterion

Knowledge, attitude creativity and performance in presenting and completing demons/assignments

Excellent (A+, A, A-)

Excellent knowledge of major issues in language processing and various approaches to their solution. Excellent, creative application of computing and programming knowledge to basic language processing subtasks.

Good (B+, B, B-)

Good knowledge of major issues in language processing and various approaches to their solution. Good application of computing and programming knowledge to basic language processing subtasks.

Fair (C+, C, C-)

Adequate knowledge of major issues in language processing and various approaches to their solution. Fair application of computing and programming knowledge to basic language processing subtasks.

Marginal (D)

Basic familiarity with the subject matter. Marginal ability to apply basic computing and programming knowledge to basic language processing subtasks.

Failure (F)

Poor familiarity with the subject matter. Poor ability or fail to apply computing and programming knowledge to basic language processing subtasks.

Assessment Task

2. Examination

Criterion

Knowledge, attitude creativity and performance in presenting and completing demons/assignments

Excellent (A+, A, A-)

Excellent knowledge of major issues in language processing and various approaches to their solution. Excellent, creative application of computing and programming knowledge to basic language processing subtasks.

Good (B+, B, B-)

Good knowledge of major issues in language processing and various approaches to their solution. Good application of computing and programming knowledge to basic language processing subtasks.

Fair (C+, C, C-)

Adequate knowledge of major issues in language processing and various approaches to their solution. Fair application of computing and programming knowledge to basic language processing subtasks.

Marginal (D)

Basic familiarity with the subject matter. Marginal ability to apply basic computing and programming knowledge to basic language processing subtasks.

Failure (F)

Poor familiarity with the subject matter. Poor ability or fail to apply computing and programming knowledge to basic language processing subtasks.

Part III Other Information

Keyword Syllabus

Tokenisation, Part-of-speech tagging, N-gram models, Context-free grammars, Parsing, Linear classifiers, Feedforward neural networks, Computational graph and backpropagation, Word embeddings, Recurrent neural networks, LSTMs and GRUs, Attention and transformers, Transfer learning

Reading List

Compulsory Readings

Title	
1	Lecture notes for the course
2	Jurafsky, D. and Martin, J.H. (2021) Speech and Language Processing (3rd Edition). : https://web.stanford.edu/~jurafsky/slp3/
3	Bird, S., Klein, E. and Loper, E. Natural Language Processing with Python. https://www.nltk.org/book/

Additional Readings

Title	
1	Manning, C.D. and Schutze, H. (1999) Foundations of Statistical Natural Language Processing. Cambridge, MA: The MIT Press.
2	Eisenstein, J. Introduction to Natural Language Processing, MIT Press, 2019.
3	Rao, D. and McMahan, B., Natural Language Processing with PyTorch: Build Intelligent Language Applications Using Deep Learning. O' Relly, 2019.
4	Stanford NLP course: Natural Language Processing with Deep Learning: http://web.stanford.edu/class/cs224n/