

Academic Org: Dept of Computer Sci & Engg – Subject: AI: Systems & Tech

Course: AIST3120 **Course ID:** 014538 **Eff Date:** 2024-07-01 **Crse Status:** Active **Apprv. Status:** Approved **[New Course]**
Introduction to Natural Language Processing 自然語言處理導論

This course is an introductory course in the field of natural language processing. This course focuses on introducing what natural language processing is, the main concepts in natural language processing research, and the algorithm design of major natural language processing tasks. The setting of this course includes three parts. The first part is an introduction to the basic tasks and concepts in the field of natural language processing, such as parsers, language modeling, etc. The second part is about the main research methods of natural language processing, such as statistical learning methods and deep learning algorithms. The third part is an introduction to the research directions of natural language processing, such as the development of technologies of large language models.

本科是自然語言處理研究領域的入門課程。本科著重於介紹什麼是自然語言處理，自然語言處理研究的主要問題，以及一些主要自然語言處理任務的算法設計。本科的設定包括了三個部分：第一部分，關於自然語言處理領域基本任務和概念的介紹，如分詞，語言建模等等；第二部分，關於自然語言處理的主要研究方法，如統計學習方法介紹和深度學習算法介紹；第三部分，關於自然語言處理研究方向的介紹，如大語言模型等技術的介紹。

Grade Descriptor: A

EXCELLENT – exceptionally good performance and far exceeding expectation in all or most of the course learning outcomes; demonstration of superior understanding of the subject matter, the ability to analyze problems and apply extensive knowledge, and skillful use of concepts and materials to derive proper solutions.

有關等級說明的資料，請參閱英文版本。

B

GOOD – good performance in all course learning outcomes and exceeding expectation in some of them; demonstration of good understanding of the subject matter and the ability to use proper concepts and materials to solve most of the problems encountered.

有關等級說明的資料，請參閱英文版本。

C

FAIR – adequate performance and meeting expectation in all course learning outcomes; demonstration of adequate understanding of the subject matter and the ability to solve simple problems.

有關等級說明的資料，請參閱英文版本。

D

MARGINAL – performance barely meets the expectation in the essential course learning outcomes; demonstration of partial understanding of the subject matter and the ability to solve simple problems.

有關等級說明的資料，請參閱英文版本。

F

FAILURE – performance does not meet the expectation in the essential course learning outcomes; demonstration of serious deficiencies and the need to retake the course.

有關等級說明的資料，請參閱英文版本。

Equivalent Offering:

Units: 3 (Min) / 3 (Max) / 3 (Acad Progress)

Grading Basis: Graded

Repeat for Credit: N

Multiple Enroll: N

Course Attributes:

Topics:

COURSE OUTCOMES

Learning Outcomes:

At the end of the course of studies, students will be able to:

1. grasp the knowledge and concepts of the NLP field;
2. solve real NLP tasks via programming;
3. have a better understanding of NLP directions, such as the development of large language models.

Course Syllabus:

Week 1: Introduction
Week 2: Regular Expressions, Text Normalization, Edit Distance
Week 3: N-gram models
Week 4: Parsers
Week 4: Neural Networks
Week 5: Classification and Sentiment Analysis
Week 6: Logistic Regression
Week 7: Neural Language Models
Week 8: Recurrent Neural Networks
Week 9: Transformers
Week 10: Pre-trained Language Models
Week 11: Supervised Learning and Unsupervised Learning in NLP
Week 12: How does ChatGPT work?
Week 13: The future directions of Large Language Models

Assessment Type:

Homework or assignment : 50%
Lab reports : 20%
Project : 30%

Feedback for Evaluation:

1. Quiz and examinations
2. Course evaluation and questionnaire
3. Question-and-answer sessions during class
4. Student consultation during office hours or online

Required Readings:

-

Recommended Readings:

Dan Jurafsky and James H. Martin. Speech and Language Processing (3rd ed.), 2023.
<https://web.stanford.edu/~jurafsky/slp3/>

OFFERINGS

1. AIST3120

Acad Organization=CSD; Acad Career=UG

COMPONENTS

LAB : Size=50; Final Exam=N; Contact=1
LEC : Size=50; Final Exam=Y; Contact=2
TUT : Size=50; Final Exam=N; Contact=1

ENROLMENT REQUIREMENTS

1. AIST3120

Enrollment Requirement Group:

Prerequisite: (AIST1110 or CSCI1040 or CSCI1550 or CSCI2040) AND (ENGG1120 or ESTR1005) AND (ENGG2760 or ESTR2018)

New Enrollment Requirement(s):

Pre-requisite = (AIST1110 or CSCI1040 or CSCI1550 or CSCI2040) AND (ENGG1120 or ESTR1005) AND (ENGG2760 or ESTR2018)

Additional Information

eLearning hrs for blended cls 0
VTL-Onsite face-to-face hrs 0
VTL-Online synch. hrs 0
VTL-Online asynch. hrs 0
No. of micro-modules 0
Research components (UG) 1% - 49%

< E N D O F R E P O R T >