*Syllabus*

# Natural Language Processing

CPE 352/372/641

Asst. Prof. Dr. Nuttanart Muansuwan

**Time** Regular: Tuesday: 13.30-16.30 Room 1121**,** Inter & Graduate: Thursday: 18.00-21.00 (on-line, on-site sometimes at Room 1115)

**Learning outcomes**

1. Understand that languages have structures in different levels and that they can be computable. Be able to make the computer process words, sentences, and at the semantic level.
2. Understand and apply modern NLP efficiently for modern-day problems.

**Assessment**

1. Students can conduct text preprocessing tasks: tokenization, stemming, comparison
2. Students understand vectors and can use them to find word similarities, translations, etc.
3. Students can understand and use NLP algorithms for various tasks.
4. Student can develop and fine-tune language models for modern NLP applications

## Course plan (modules)

Module 1: Traditional NLP: introduction, word processing pipeline, regular expressions, morphology, syntax, semantics

Module 2: Applications: Named-entity recognition, machine translation, information retrieval, question-answering, sentiment analysis, chatbots

Module 3: Modern NLP: Transformer, Neural-based, Large language models (LLMs), finetuning

## Textbook & tools

- Link for e-book: Speech and Language Processing (3rd ed. draft) by Dan Jurafsky and James H. Martin https://web.stanford.edu/~jurafsky/slp3/.
- You will use Python for programming assignments. However, Python will not be taught here so it means you have to know how to code in Python prior to this course.

- Zoho notebook to create the notebook of each module: https://www.zoho.com/notebook/notebook-for-windows.html

  Video about Zoho notebook: https://www.youtube.com/watch?v=d_SYZVyVJXw&t=105s

## Communication

FB group (for announcements, course materials and assignments can be found in LEB2) -- https://www.facebook.com/groups/581582287849177 **Natural Language Processing 2025**

## Grading

Notebook from 3 modules: 12%, Presentation of notebook (1 module only) 5%, Attitude 3%*

Midterm 15%

Final 15%

Lab & homework: 25% (20 questions, Code a joke, Minimum edit distance, instruct lab, etc.)

Project 15%

Summary from reading & quiz 10%

## AI Tools Policy

Students are required to independently submit their solutions for individual assignments. Collaboration with generative AI tools such as Co-Pilot and ChatGPT is allowed, treating them as collaborators in the problem-solving process. However, the direct solicitation of answers or copying solutions, whether from peers or external sources, is strictly prohibited. Employing AI tools to substantially complete assignments or exams is also prohibited. Violation may result in 0 score.

*Attitude scores and dress code policy

As the instructor, I reserve the right to give attitude scores for students who collaborate by following the dress code policy. Everyone is supposed to get full 3% of attitude scores except when you violate the dress code policy when coming to the classroom.

**Dress code policy**: University uniform is most encouraged but you can wear other polite outfits. What is considered not polite and you should not wear are listed below.

1. No sandals or flip-flops.
2. No shorts or pajamas, nightwear or lounge wear.
3. No elephant pants or other variations.

**Schedule**

| **Topic** | **Regular Date (Tue.)** | **Grad/inter Date (Th.) ONLINE** except mentioned otherwise | **Task** |
|---|---|---|---|
| Introduction | Jan. 14 | Jan. 16 (on-site) | |
| Word, text processing, Regular expression | Jan. 21 | Jan. 23 | HW 1 |
| Text processing Pipeline | Jan. 28 | Jan. 30 | |
| Syntax and semantics | Feb. 4 | Feb. 6 | HW 2 |
| Conclusion of Module 1 | Feb. 11 | Feb. 13 | Zoho notebook Module 1 present |
| | Feb. 18 (on-site) | | Midterm exam |
| Named Entity recognition | Feb. 25 | Feb. 27 | |
| IR, QA, Chatbot | Mar. 4 | Mar. 6 | HW 3 |
| Sentiment analysis | Mar. 11 | Mar. 13 | HW 4 |
| Machine translation | Mar. 18 | Mar. 20 | |
| Conclusion Module 2 | Mar. 25 | Mar. 27 (on-site) | Zoho notebook Module 2 present |
| Modern NLP, neural-based transformer | Apr. 22 | Apr. 24 | |
| LLMs | Apr. 29 | May 1 | HW 5 |
| Finetuning | May 6 | May 9 | |
| Conclusion, term project | May 13 | May 15 (on-site) | Zoho notebook Module 3 present |
| | May 20 (on-site) | | Final exam |