Technologies and Implementation of Natural Language / 自然語言技術與實作

CJ Wu/吳齊人

Spring 2023 Syllabus

This introductory course is about a variety of ways to represent human languages for Computer Science (CS) undergraduate students. This course will introduce students to the basics of Natural Language Processing (NLP). The students will gain the ability to exploit languages representations to write programs based on the modern data-driven techniques. This course will include some ideas, including Machine Learning and the recent Deep Learning. There will be an emphasis on rapid prototyping, a useful skill in many other areas.

Topics covered include:

- Web crawling and indexes
- Language modeling
- Representation learning
- Word Embeddings
- Text classification
- Sequence modeling
- Machine learning models
- Deep learning models

Textbook

1. Jurafsky and Martin,. Speech and Language Processing. 3 edition https://web.stanford.edu/~jurafsky/slp3/ed3book_dec302020.pdf

2. Raghavan, and Schutze. 2008. Introduction to Information Retrieval. Cambridge University Press. https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf

References

1. Jurafsky and Martin,. Speech and Language Processing. 3 edition
https://web.stanford.edu/~jurafsky/slp3/ed3book_dec302020.pdf

2. Raghavan, and Schutze. 2008. Introduction to Information Retrieval. Cambridge University Press.
https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf

3. Natural Language Processing With Python's NLTK Package.
https://realpython.com/nltk-nlp-python/

4. Tencent AI LAB (中文)
https://ai.tencent.com/ailab/zh/paper/?page=1

5. CKIP Lab (中文)

https://ckip.iis.sinica.edu.tw/

6. Kaggle data

https://www.kaggle.com/

Grading

Your course grade will be determined approximately as follows:
- 40%: Final project
- 25%: Quizzes
- 35%: Exercises

Technologies and Implementation of Natural Language / 自然語言技術與實作

CJ Wu/吳齊人

Spring 2023

教學進度 Course Progress Outline

週次 No.	上課日期/星期 Date / Weekday	時數 Hours	授課教師 Instructor	教學進度 Outline	訊息 Note
1	2/13 — (Mon)	3	資工系 吳齊人	Introduction to NLP	
2	2/20 — (Mon)	3	資工系 吳齊人	Language modeling & Representation learning	
3	2/27 — (Mon)	3		Peace Memorial Day Holiday	
4	3/6 — (Mon)	3	資工系 吳齊人	Language modeling & Representation learning	HW#1
5	3/13 — (Mon)	3	資工系 吳齊人	Web crawling and indexes	
6	3/20 — (Mon)	3	資工系 吳齊人	Web crawling and indexes	HW#2
7	3/27 — (Mon)	3	資工系 吳齊人	Text Classification and Sentiment Analysis	
8	4/3 — (Mon)	3		Public Holiday	
9	4/10 — (Mon)	3	資工系 吳齊人	Text Classification and Sentiment Analysis	HW#3
10	4/17 — (Mon)	3	資工系 吳齊人	Word Embeddings	
11	4/24 — (Mon)	3	資工系 吳齊人	Word Embeddings	

Take care of yourself.

12	5/1 — (Mon)	3	資工系 吳齊人	Machine learning models for NLP	HW#4
13	5/8 — (Mon)	3	資工系 吳齊人	Machine learning models for NLP	
14	5/15 — (Mon)	3	資工系 吳齊人	Deep learning models for NLP	HW#5
15	5/22 — (Mon)	3	資工系 吳齊人	Deep learning models for NLP	
16	5/29 — (Mon)	3	資工系 吳齊人	State of the Art (SOTA)	
17	6/5 — (Mon)	3	資工系 吳齊人	Final project demo	
18	6/12 — (Mon)	3	資工系 吳齊人	Final project demo	

Take care of yourself.