

Course Syllabus

Natural Language Processing (NLP)

Super AI Engineer Course by AI Association of Thailand

Course	:	Natural Language Processing (NLP)
Credit	:	3 credits
Semester	:	January 2022 – April 2022
Course Outline	:	Efficiently encoding of spelling rules. Tagging words according to their part of speech. Parsing English sentences. Language translation. Modeling language semantics. Organization, representation, and access to information. Techniques for filtering junk email. Techniques for discovering the different meanings of a word. Categorization, indexing, and content analysis. Information retrieval application. Web Scraping. Design and maintenance of dictionary and thesauri. Use of codes, formats, and standards. On-hand workshops. Introduction to Automatic Speech Recognition, Speech recognition components, ASR Technique and Partii Demo, Basic structure ASR System, KALDI – Introduction, Example script – common-voice, Data preparation: Thai syllable system, Language model preparation.
Instructor	:	Prof. Dr. Thanaruk Theeramunkong (thanaruk@siit.tu.ac.th) (SIIT) Dr. Thepchai Supnithi (thepchai@nectec.or.th) (NECTEC) Dr. Peerachet Porkaew (peerachet.porkaew@nectec.or.th) (NECTEC and CAS) Dr. Prachya Boonkwan (prachya.boonkwan@nectec.or.th) (NECTEC) Dr. Kobkrit Viriyayudhakorn (kobkrit@iapp.co.th) (iAPP Technology) Dr. Nuttapon Sanglerdsinlapachai (nuttapon.sanglerdsinlapachai@nectec.or.th) (NECTEC) Dr. Kwanchiva Thangthai (kwanchiva.thangthai@nectec.or.th) (NECTEC) A. Nattapon Kurpukdee (nattapon.kurpukdee@nectec.or.th) (NECTEC)
Grading	:	Attendance / Quiz 20% Examination 40% On-hand Project 40% Top 20% → ‘A’. Bottom 20% and/or students whose score < 30% → ‘F’
Quiz	:	Quizzes are randomly conducted in the classes
Projects	:	The project aims to give you experience of natural language processing. The project will be classified into individual hackathon projects, small group projects, and big bang group projects.
Course Material	:	http://mooc.aiat.or.th/

Schedule:

No.	Topics	Hours
1	Introduction to Natural Language Processing	3
2	Syntactic Processing: Grammar and Parsing	3
3	Semantic Processing: Logic, Knowledge Representation, Semantic Composition	3
4	Statistical NLP Models: Tagging and Statistical Parsing, Context-sensitive and Advance Language Models	3
5	Statistical Machine Translation (SMT) and Neural Machine Translation (NMT)	3
6	Machine Translation: Data Preparation and Evaluation	3
7	Encoder and Decoder: LSTM, BERT, Transformer	3
8	Pretrained Models and their Applications: Text Classification	3
9	Web Scraping: HTML, Regular Expression, Document Classification	3
10	Topic Modeling, Phishing Web Detection and Sentiment Analysis	3
11	Automatic Speech Recognition, ASR Technique and Partii Demo Basic structure ASR System, KALDI	3
12	Speech script – common-voice, Data preparation: Thai syllable system Language model preparation	3
13	Project Workshop 1	10
14	Project Workshop 2	10
15	Project Workshop 3	10
16	Examination	
	Lecture	36
	Workshop	30