2110594: Natural Language Processing
Mondays 9:30-12:30
ผู้สอน อ.ฟีรพล และ อ.เอกพล

JOIN THE DARK SIDE

DO NLP WITH US
memegenerator.net

Course Outline
The course will give an overview of the Natural Language Processing techniques in the transition period between traditional techniques and deep learning. The course will touch upon various standard NLP tasks, such as tokenization, language modeling, semantics, part-of-speech tagging, and parsing. Each topic will discuss both conventional and deep learning techniques. The second part of the course will go into application domains such as document classification, question answering, and chatbots. Many of the assignments will be in the context of Thai language. This course is not meant to be an entry-level machine learning course, and will not cover the basics of machine learning. This course is recommended for 4th year students who have already taken some machine learning course, and 3rd year students who are interested in doing senior projects that are related to NLP.

Tools: Python, keras, ntlk, scikit-learn
Prerequisites: some background in machine learning/data science

<u>Schedule</u> (กำหนดการอาจมีการเปลี่ยนแปลง)

คาบเรียนที่	เนื้อหา	การบ้านและควิซ
1 - 7/1	Intro to NLP and basic tokenization (P & E) Longest matching and maximal matching	HW1-Hello GCloud & basic tokenization
2 - 21/1	Deep learning DNN, CNN, LSTM with Keras. (E)	HW2-TUMKUDv2
3 - 28/1	PoS tagging and information extraction (E) CRF and beam search	HW3-CRF
4 - 4/2	Language modeling (P) N-gram, perplexity, and discounting	HW4-LM
5 - 11/2	Representation (P) TF-IDF, word and sentence embeddings, adaptation	HW5-Embeddings
6 - 18/2	Document/sentiment classification (E) LDA, Naive Bayes, EM, LDA2Vec	HW6-Intent classification
7 - 25/2	In class midterm	
	Midterm exam week	
8 - 11/3	Parsing (E) PCFG, Recursive neural networks	HW7-Parsing
9 - 18/3	Attention & MT1 (P)	HW8-Text Generation
10 - 25/3	MT2 & QA (E)	HW9-Project scoping
11 - 1/4	Recent research in NLP (P)	HW10-Paper assignment
12 - 8/4	Industry talks (Guest speakers)	
13 - 22/4	Misc topics	Paper Presentation & Progress Report
14 - 29/4	Project presentation	

<u>การส่งงานสาย</u>
การส่งงานสายเกิน5นาทีหลังจากหมดกำหนดส่งจะไม่ได้คะแนนในทุกกรณี

<u>เกณฑ์การวัดผล</u>

Assignments 35% (4% each with 5% extra)
Midterm 35%
Project 30%

<u>การตัดเกรด</u>
> 80% A
> 75% B+
> 70% B
> 65% C+
> 60% C
> 55% D+
> 50% D
< 50% F

<u>หนังสือเรียน</u>
ไม่มีหนังสือเรียนบังคับ