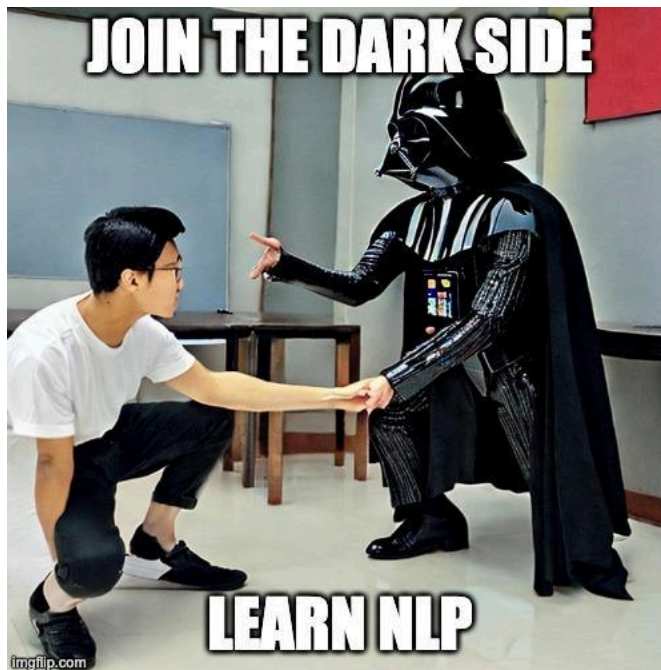# 2110572: Natural Language Processing Systems
Mondays 9:30-12:30
ตึก 100 ปี ห้อง 201A
ผู้สอน อ.เอกพล และ อ.พีรพล

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 advanced undergraduate or graduate students who have already taken some machine learning course.

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

Schedule (กำหนดการอาจมีการเปลี่ยนแปลง)

คาบเรียน	เนื้อหา	การบ้าน
9-Jan-2023	Intro; Traditional tokenization	HW1
16-Jan-2023	Tokenization (deep learning)	HW2
23-Jan-2023	Language Model	HW3
30-Jan-2023	Word Representation	HW4
6-Feb-2023	Holiday No class	
13-Feb-2023	PoS Tagging	HW5
20-Feb-2023	Text Categorization	HW6
27-Feb-2023	Parsing	HW7, take home announced
6-Mar-2023	Midterm Exam Week – No class	
13-Mar-2023	Attention mechanism & Transformer & BERT takehome due (14)	HW8 out
20-Mar-2023	In class midterm exam	HW7 due
27-Mar-2023	BERT++ MT & QA Project Announcement + Paper Announcement	HW8 due, HW9 & HW10 out
3-Apr-2023	Recent Research in NLP	
10-Apr-2023	Paper Presentation & Progress Report	HW9 due
17-Apr-2023	Songkran Holiday	HW10 due
24-Apr-2023	NLP Application (Guest)	Guest Report
1-May-2023	Project Presentation	

Note 23 Jan will be completely online

การส่งงานสาย
การส่งงานสายเกิน 5 นาทีหลังจากหมดกำหนดส่งจะไม่ได้คะแนนในทุกกรณี

เกณฑ์การวัดผล
Assignments 35% (4%x10 HW capped at 35%)
Midterm 35% (in class exam 25% + take home 10%)
Paper presentation 8%
Guest report 2%
Project 20%

การตัดเกรด
> 80% A
> 75% B+
> 70% B
> 65% C+
> 60% C

> 55% D+
> 50% D
< 50% F

<u>หนังสือเรียน</u>
ไม่มีหนังสือเรียนบังคับ
หนังสือแนะนำ
Dan Jurafsky and James H. Martin, *Speech and Language Processing (3rd ed. draft)*,
https://web.stanford.edu/~jurafsky/slp3/

<u>Course Github</u>
https://github.com/ekapolc/NLP_ ₂₀₂₃

<u>Discord</u>
https://discord.gg/tBTQUGkh

<u>MyCourseVille Registration</u>
Avatar2