Course Syllabus

1. Course Number: 2110446
2. Course Credit: 3 (3-0-6)
3. Course Title: Data Science and Data Engineering
4. Department Computer Engineering
5. Semester Second Semester
6. Academic Year 2024
7. Instructor Assoc. Prof. Peerapon Vateekul, Ph.D.
   Asst. Prof. Natawut Nupairoj, Ph.D.
   Assoc. Prof. Veera Muangsin, Ph.D.
8. Condition -
9. Status Approved Elective
10. Curriculum Computer Engineering
11. Degree B.Eng.
12. Hours/Week 3-Hour Lecture & Lab
13. Course Description

Data Science is the study of the discovery of knowledge from data. Being a data scientist requires an integrated skill set spanning mathematics, statistics, machine learning, databases, and other branches of computer science along with a good understanding of the craft of problem solving. Data Engineering is the study of how to engineer or process data, i.e., data cleansing, data storing, etc. There are three main parts in this course:

- Data engineering: Data exploration & preparation
- Data analysis: Machine Learning techniques
- Data visualization: Storytelling via data

14. Course Outline

14.1. Learning Objectives

- Describe what Data Science and Data Engineering are and the skill sets needed
- To be able to explore and understand collected data
- To be able to analyze data by apply traditional machine learning techniques
- To be able to visualize data in relation to spatial and temporal points of views

14.2. Learning Contents

- ENG100 Room 201A
- Prerequisites: Python programming, Statistics
- *** Please bring your laptop to the class ***
- Students need to attend the class on-site at least 80% (at least 12 weeks) as a mandatory criterion to “pass” this course.

#	Tue (1PM-4PM)	Contents	Instructor	Module
1	7-Jan-25	Introduction & data preparation	Aj.Peerapon	Data Science
2	14-Jan-25	Traditional ML (1)	Aj.Peerapon	Data Science
3	21-Jan-25	Traditional ML (2)	Aj.Peerapon	Data Science
4	28-Jan-25	DL (1)	Aj.Peerapon	Data Science
5	4-Feb-25	DL (2)	Aj.Peerapon	Data Science
6	11-Feb-25	Advanced topics & Guest speakers * Modern ML techniques (Transformer) * Sharing data scientist experience (in Thailand & Abroad) [2-3 guest speakers]	Aj.Peerapon	Data Science
7	18-Feb-25	Big data architecture + data storage	Aj.Natawut	Big Data Eng.
8	25-Feb-25	Web scraping	Aj.Natawut	Big Data Eng.
	5-Mar-25	Midterm Exam Week (3-7 Mar)		
9	11-Mar-25	Data ingestion	Aj.Natawut	Big Data Eng.
10	18-Mar-25	Big data processing (Spark)	Aj.Natawut	Big Data Eng.
11	25-Mar-25	MLOps: Orchestration (Airflow) and serving (FastAPI, Seldon Core)	Aj.Natawut	Big Data Eng.
12	1-Apr-25	Data visualization	Aj.Veera	Data Viz.
13	8-Apr-25	Python visualization, Streamlit, and Dash	Aj.Veera	Data Viz.
	15-Apr-25	Songkran Holiday		
14	22-Apr-25	Graph visualization	Aj.Veera	Data Viz.
	30-Apr-25	Final Exam Week (28 Apr -14 May) TBD: Tue 6 May 2025 at 1PM - 4PM		

14.3. Method: Lecture and Lab

14.4. Learning Media: PowerPoint presentation, Zoom

14.5. Evaluation

- Module1 Assignment (data analytics) 15%
- Module2 Assignment (data engineering) 15%
- Module3 Assignment (data visualization) 5%
- Midterm Exam (Kaggle) 15%
- Project 15%
- Attendance 5%
- Final Exam 30% (Lab Test)

15. Reading List
15.1. Required Text: N/A
15.2. Electronic Media or Websites:

16. LMS
- 16.1. myCourseVille: “Sagemaker”
- 16.2. Discord: https://discord.gg/hDAjRQwQ3J
- 16.3. Github: https://github.com/pvateekul/2110446_DSDE_2024s2