Course Syllabus

1.	Course Number:	2110531
2.	Course Credit:	3 (3-0-6)
3.	Course Title:	Data Science and Data Engineering Tools
4.	Department	Computer Engineering
5.	Semester	First Semester
6.	Academic Year	2023
7.	Instructor	Peerapon Vateekul, Ph.D. Natawut Nupairoj, Ph.D. Veera Muangsin, Ph.D.
8.	Condition	-
9.	Status	Elective
10.	Curriculum	Computer Engineering
11.	Degree	M.Eng., M.Sc.
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. 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

- Section1 on Tue 1PM-4PM, Building ENG3 Room 417
- Section5 on Sat 1PM-4PM, Building ENG4 Room 18-16

#	Tue 1PM-4PM	Sat 1PM-4PM	Contents	Instructor	Module
1	08-Aug-23	12-Aug-23 Online	Introduction, Pandas, Data Prep (12-Aug is Mother Day)	Aj.Peerapon	Data Science
2	15-Aug-23	19-Aug-23	Traditional ML (1)	Aj.Peerapon	Data Science
3	22-Aug-23	26-Aug-23	Traditional ML (2)	Aj.Peerapon	Data Science
4	29-Aug-23	02-Sep-23	Deep Learning (1); CNN, RNN (LSTM, GRU)	Aj.Peerapon	Data Science
5	05-Sep-23	09-Sep-23	Deep Learning (2); Transformer	Aj.Peerapon	Data Science
6	12-Sep-23	16-Sep-23	Advanced topics (Generative AI) + Model monitoring (Mlflows)	Aj.Peerapon	Data Science
7	19-Sep-23	23-Sep-23	Big data architecture + data storage	Aj.Natawut	Big Data Eng.
	26-Sep-23	30-Sep-23	Midterm Exam Week (25 - 29 Sep 2023)		
8	03-Oct-23	07-Oct-23 Online	Web scraping (7-Oct is Graduation Day)	Aj.Natawut	Big Data Eng.
9	10-Oct-23	14-Oct-23	Data ingestion	Aj.Natawut	Big Data Eng.
10	17-Oct-23	21-Oct-23	Big data processing (Spark)	Aj.Natawut	Big Data Eng.
11	24-Oct-23	28-Oct-23	MLOps: Orchestration (Airflow) and serving (FastAPI, Seldon Core)	Aj.Natawut	Big Data Eng.
12	31-Oct-23	04-Nov-23	Guest speakers (AWS)	Aj.Peerapon	Data Science
13	07-Nov-23	11-Nov-23	Data visualization	Aj.Veera	Data Viz
14	14-Nov-23	18-Nov-23	Python visualization	Aj.Veera	Data Viz
15	21-Nov-23	25-Nov-23	Graph analysis & spatial analysis	Aj.Veera	Data Viz
	28-Nov-23	02-Dec-23	Final Exam Week (27 Nov - 12 Dec 2023) *** Final Exam on Sat 2 Dec 2023		

* There will be up to two guest speakers in the class.

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)	10%
Midterm Exam (Kaggle)	15%
Project	15%
Final Exam	30% (Lab Test)

15. Reading List

15.1. Required Text: N/A

15.2. Electronic Media or Websites:

16. LMS

16.1. CourseVille: "GenerativeAI"

16.2. Discord: https://discord.gg/CgVwVtnuh2

16.3. Github: https://github.com/pvateekul/2110531_DSDE_2023s1