CPE 232 Data Models

Department of Computer Engineering, KMUTT, 2/2025
Section: 1, 2

Instructor:	Dr. Sansiri Tarnpradab	**Email:**	sansiri.tarn@mail.kmutt.ac.th
TA:	Paweekorn Soratyathorn	**Email:**	punsoratyathorn@gmail.com
	Sawit Koseeyaumporn		folk.sawit@gmail.com
	Warapol Pratumta		topwarapol8850@gmail.com
	Nitis Jadipang		nitis.jadipang@gmail.com
	Pheeraphat Siriphatthanakhun		lukmeepook@gmail.com
Class Time:	Fri 13.30 - 15.30 (Lecture) 15.30 - 17.30 (Lab)	**Room:**	CB 2506

COURSE OBJECTIVES:

This course aims to introduce fundamental concepts of working with data, focusing on how data can be understood, represented, explored, and analyzed to gain meaningful insights. The course covers essential aspects of data handling, including identifying data types, preparing and cleaning data, conducting exploratory data analysis, and creating data visualizations. It also introduces analytical and selected machine learning techniques to apply to real-world data for further analysis and interpretation.

COURSE LEARNING OUTCOMES:

- Students are able to evaluate and apply suitable data modeling techniques to analyze real-world data.
- Students are able to create meaningful visualization to address relevant problems.
- Students are able to understand data science process along with roles of data scientists.

COURSE PLATFORM:

- LEB2 is used for course announcements, distribution of course materials, and assignment submission.
- Discord is used for communication with teaching assistants and discussion among students.

COURSE GRADING:

Quiz	15%
Assignments	25%
Midterm Exam	20%
Final Exam	20%
Project & Presentation	20%

HONESTY POLICY:

The Computer Engineering Department's honesty policy will be strictly enforced. Any assigned work including lab work, if copied with permission, all persons involved will receive a negative score equivalent to the full score of the assigned work for first violation; a second violation will result in F for the course.

LATE POLICY:

Student are given up to one week to complete the assignments. For each day late, 10% will be deducted.

RESOURCES:

- Wes Mckinney (2022). Python for Data Analysis. O'Reilly Media. (Link)
- Claus O. Wilke (2019). Fundamental of Data Visualization. O'Reilly Media. (Link)
- Aurelien Geron (2019). Hands-On Machine Learning with Scikit-Learn & TensorFlow. O'Reilly Media. (Link)

Note: The listed resources are not exhaustive. Additional references may be provided throughout the course.

COURSE SCHEDULE:

The following is a *tentative* course schedule. The instructor may revise parts of the outline to better align with the background, knowledge, and interests of the students.

No.	Topics	Remarks
1	Introduction to Data Models	
	Lab 1: Let's review Python!	
2	Types of Data	
	Lab 2: Let's see types of Data!	
3	Data Preparation	
	Lab 3: Let's play with data!	
4	Data Visualization	
	Lab 4: Let's visualize data!	
5	Exploratory Data Analysis (EDA)	
	Lab 5: Let's explore data!	
6	**Midterm Exam** *(Theory & Coding)*	
7	Machine Learning: Classification	
	Lab 6: Let's classify data!	Proposal due
8	Model Parameter Tuning	
	Lab 7: Let's play with parameters!	

9	Machine Learning: Regression	
	Lab 8: Let's predict from data!	
10	Machine Learning: Clustering	
	Lab 9: Let's group data!	
11	Project Presentation	
12	**Final Exam** *(Theory)*	

Note: Any additional modifications to the syllabus will be announced in class and reflected on the LEB2 course page. Students are advised to regularly check LEB2 for the most up-to-date schedule.