ITD3019-02, Data Mining

Fall 2020

Class Hours: Tue 1:00pm - 2:40pm / Thu 11:am - 12:40pm

Professor: Dr. Young-Rae Cho [redacted]

Office: Chang-jo-kwan 328

TA: Won-Doo Seo [redacted]

Tutor: Jihee Han [redacted]

Online office hour: Mon 6-7pm / Wed 12-1pm

Course Web Page: https://it.yonsei.ac.kr/adslab/faculty/data_mining/

Description

Introduction to the concepts, techniques and applications of data mining. Topics include (1) data mining concepts and methods such as association rule mining, pattern mining, classification and clustering, and (2) applications of data mining techniques to complex types of data in various fields.

Objectives

- To understand the basic concepts and techniques of Data Mining.
- To develop computational skills of implementing data mining algorithms to solve practical problems.
- To gain experience of performing a research project on Data Mining.

Textbook

- *Data Mining: Concepts and Techniques*, 3rd Edition, by Jiawei Han, et al., Morgan Kaufmann

Prerequisites

- Algorithms, Databases, Statistics
- Proficiency in Python programming

Assignments

- 7 programming assignments using Python
- Implementing data mining algorithms and analyzing results
- Submission of source codes and results via YSCEC

Exams

- 1 midterm exam (online)
- 1 final exam (online)

Project

- A team (or an individual) project
- Developing a new algorithm and comparing its performance with other methods

- Submission of a report and source codes (if requested) by email
- Oral presentation of project results (optional)

Grading

- Assignments (45%), Quiz (5%), Final exam (30%), Project (20%)

Policies

- According to the university policy, absences more than 30% will cause getting an F as the final grade of the course no matter what scores are obtained in exams and assignments.
- Programming assignment is due at 11:59pm on the specified date. **Late submission of programming assignments will receive the penalty of -20% per day** including Saturday and Sunday.
- Discussions on programming assignments are allowed, but all programming assignments must be independent work. Any forms of cheating on the programming assignments, project and exams will cause a penalty of getting an F as the final grade to the university regulation guidelines.

Topics & Tentative Schedule

Week	Topic	Reading	Assignment Due
1	Introduction	Chap. 1	
2	Frequent Pattern Mining	Chap. 6	
3		Chap. 7	Assignment 1
4	Clustering	Chap. 10	
5		Chap. 10	Assignment 2
6	Graph Data Mining	Chap. 11	Assignment 3
7		Chap. 11	
8	Sequence Data Mining	Chap. 13	Assignment 4
9		Chap. 13	Assignment 5
10	Classification	Chap. 8	
11		Chap. 9	Assignment 6
12	Data Preprocessing	Chap. 2	Assignment 7
13		Chap. 3	
14	Project Presentation		Project
15	Study Days		