Data Mining
Fall 2024
Term Project

Goal

Find a dataset (or multiple datasets) of your interest, set the goals for data mining, then conduct a full data mining process to mine the selected dataset by using Data Mining techniques. The project should be implemented under the following **requirements**:

1. You may use any languages like **Python, Java, C++** as well as **open-source libraries/tools like scikit-learn, Weka, etc**.
2. You must use <u>at least</u> two of the following three types of data mining techniques in the project: **I) Link Analysis** (e.g., Associations Mining, Sequential Pattern Mining, Time Series Mining); **II) Classification**; **III) Clustering**.

Dataset

You may use any interesting dataset that is **publicized**. The followings are some good reference sources to start with (not limited to the following):

- Kaggle dataset (https://www.kaggle.com/datasets)
- Kdnuggets (https://www.kdnuggets.com)

Requirements

1. **Team Size:**
   - Default 4 people in each team (may vary dynamically as needed).
2. **Important Dates:**

   <u>Registration</u>
   - Team registration: **2024/09/16(Mon.) 10:00 a.m. ~ 2024/09/20(Fri.) 11:59 a.m.**
   - Topic & dataset registration: **2024/09/23(Mon.) 10:00 a.m. ~ 2024/09/27(Fri.) 11:59 p.m.**

   <u>Proposal Presentation</u> (on-site, play the video)
   - Proposal slides & video upload: before **2024/10/13 (Sun.) 11:59 a.m.**
   - Proposal presentation: **2024/10/16(Wed.) & 2024/10/23(Wed.)**
   - Proposal revision submission: before **2024/10/30(Wed.) 11:59 p.m.**

   <u>Final presentation</u> (on-site, play the video) & Report
   - Final presentation slides & video upload: before **2024/12/01(Sun.) 11:59 a.m.**
   - Final presentation: **2024/12/04(Wed.) & 2023/12/11(Wed.)**
   - Final report submission: before **2024/12/20(Fri.) 11:59 p.m.**

*More details and instructions will be announced on E3.