Final Project

Objectives

- Apply what you have learned to analyze real dataset
- Extend your learning
- Interpret the results of data analysis
- Present your work

Application Datasets

- **UC Irvine Machine Learning Repository**
  https://archive.ics.uci.edu/ml/datasets.php

1. Steel Plates Faults Dataset (27 attributes, 1941 instances)
https://archive.ics.uci.edu/ml/datasets/steel+plates+faults

Steel plates faults are classified into 7 types.

2. Concrete Compressive Strength Dataset (9 attributes, 1030 instances)
https://archive.ics.uci.edu/ml/datasets/Concrete+Compressive+Strength

Concrete compressive strength and related variables.

Tasks

- Choose one from the given two data sets
- Formulate an appropriate problem
- Propose a strategy/procedure to solve the problem
- Perform the analysis (training & test)
- Write a report

Report

- 1. Power Point slides (pdf file)
  - Title page: Course name, project title, your names, student ID, date
  - Background, problem formulation, strategy/methods, justification, data analysis, results, interpretation, discussion
  - Page limit: 12 slides (including title page); extra slides will be ignored.
  - DON'T include codes in the slides.
- 2. Appendix (pdf file): codes used in your data analysis
- 3. Presentation (video file)
  - Present your slides as if you are in front of the teacher and classmates, and record your presentation.
  - All members should participate in the presentation.
  - Start with a summary of team member contributions
  - Time limit: 10 minutes

Submission Deadline

- Canvas--→Assignments
  - “Project_slides”: submit ppt slides
  - “Project_presentation”: submit video file
  - “Project_appendix”: submit appendix file

Submission deadline: November 27, Sunday @10:00 PM

Requirements

- Reading references related with the chosen dataset is allowed.
- Your report must contain something that is not covered in lectures.
- This is a group assignment. Each group only needs to do one submission.
- Each group is required to do independent data analysis and report writing.
- You can ask TA for help ONLY if you have difficulty in downloading a dataset and importing the data to software.
- Grading on project will be based on the quality of data analysis (60%) and quality of presentation (40%).