BLM4800 – Introduction to Data Mining

Project Report Instructions

- **Report Instructions:** This project can be completed as a group or individually. More work is expected from group projects. Groups can consist of a maximum of two or three students.
  - After selecting the dataset, you should explore and analyze the data.
  - You will need to develop data mining models that will fit the dataset in your analysis.
  - You should create visualizations and exploratory graphs to help express the findings of your analysis.
  - You can use the following pages as an example of sample scripts to provide ideas for data analysis and visualization.
    - https://github.com/wehrley/wehrley.github.io/blob/master/SOUPTONUTS.md
    - https://statsguys.wordpress.com/2014/01/03/first-post/

  - **In your own words**, write a report in research report format describing the data set you selected and the models you used to analyze the data. Plagiarism will not be tolerated and plagiarism will be severely punished.
  - Consider good reading flow in the report and pay attention to visualizations to increase readability.
  - Submit your report only through the online.yildiz.edu.tr system.
  - A report of 5 to 12 pages is expected.

  Report Format Instructions:
  - Your completed report should be designed in IEEE conference paper format with 2 columns. It should look similar to this file:
  - https://www.ieee.org/conferences/publishing/templates.html
  - For an example of the format for the completed report, you can refer to the article "Distdgl: distributed graph neural network training for billion-scale graphs"

- **Deadline for submission of the report: 28 May 2023**
- **Reports submitted after the deadline will be penalized.**