CME 4403 Term Project –
First Submission: May 4, 2021, 23:59
Final Submission: June 15, 2021, 23:59

Please combine source codes and report in one .docx file (e.g., YourName.docx) then submit it to the Classroom.

- The project will be coded in **Python**.
- You need to implement **two machine-learning models** taught in the course to perform a classification task (for two or more target classes) on a real data set.
- You should use a large data set selected from any domain. Possible web sites for real data sets:
  http://homepages.inf.ed.ac.uk/rbf/IAPR/researchers/MLPAGES/mldat.htm
  https://archive.ics.uci.edu/ml/datasets.html
  https://www.kaggle.com/datasets
- You should analyze a unique data set, i.e., not used by another one. Enter your dataset link into the google sheet. **Do not use banned datasets** given in pdf file.
- An image registration / classification problem is also acceptable.
- Apply a cross-validation scheme while training and evaluating the model performance.
- Change model parameters to observe if there is a performance improvement on the classification task.
- Report performance of the model by using different evaluation metrics e.g., accuracy, precision, recall, etc.
- **First submission:** Write down a project progress report and Prepare a short presentation (3-4 minutes) which should describe the data set, plots, pre-processing steps
- **Final submission:** Write down a final project report, which should describe the data set, pre-processing steps, machine-learning model, parameter tuning, and performance results with plots. Prepare a presentation (6-8 minutes) to explain

your data set, data cleaning stages, machine learning model, parameter tuning, results etc.

- All students should attend the project presentations. Otherwise, their projects will not be graded.