Deadline for Project Submission is Dec 09, 2023 5:00 pm Pacific Time

Please submit –

1. A Short write– up (less than 2 pages, 12 font 1.5 spacing). It should contain the following sections (a) Definition of the Problem, (b) Rational of target variable selection, (c) A short note on which kind of machine learning is suitable for this problem, (d) Conclusion with a at least one visual (e) Scope for future work (f) Did you go beyond the expectation and deserve the extra 5 points ? If so, why and what extra work did you do ?
2. Jupyter Notebook with code that goes all the way from reading in the data to final output.
3. A random set of up to 10 students will be asked to remotely code review by the TA's

1. Problem Understanding and Definition (5 points)
- Clear definition of the problem statement and its significance.
- Identification of the target variable and relevant features.
- Proper framing of the problem as a supervised, unsupervised, or reinforcement learning task.

2. Data Collection and Preprocessing (5 points)
- Explanation of data sources and data collection methods.
- Data preprocessing steps such as handling missing values, data normalization, and feature engineering.
- Justification for data preprocessing choices.

3. Model Selection and Implementation (5 points)
- Appropriate selection of machine learning models for the given task.
- Implementation of the selected models using a suitable programming language or framework.
- Proper parameter tuning to optimize model performance.

4. Model Evaluation and Interpretation (5 points)
- Rigorous evaluation of model performance using appropriate metrics (e.g., accuracy, precision, recall, F1-score, etc.).
- Interpretation of model results and insights gained from the analysis.

5. Presentation and Communication (5 points)
- Clarity and organization of the project report or presentation.

- Ability to effectively communicate technical concepts to a non-technical audience.
- Use of visual aids and examples to support explanations.

6. Creativity and Innovation (5 points)
- Introduction of original ideas or novel approaches to solving the problem.
- Not merely copy pasting, massaging an existing notebook like from kaggle.

7. Code Quality and Documentation (5 points)
- Readability and organization of the code.
- Sufficient comments and documentation to explain the code logic.

8. Bonus Points (up to 5 points)
- Extra points for exceptional performance, going beyond the basic requirements, or addressing additional challenges.

Maximum marks is 40. Late submissions accrue 10 marks penalty and please note no exceptions will be made. Please allow a buffer time to allow for any last minute emergencies.