Project Deliverable 3

Due Date: April 30, 2023

This assignment consists of two parts and for each of them there is a separate submission on Canvas.

Part I of the assignment is a peer review. Each student must **submit an individual review** of her/his peers by listing names of each team member and evaluating their contribution with a maximum of 5 points.

For example:

Students	Points
Team Member 1	4
Team Member 2	5
Team Member 3	3.5

Part I of the assignment will be graded based on the average review points received for each student.

Part II

The purpose of this assignment is to check the robustness of your model. **You will be provided with a new test data set. Please note you need to train the models with the previously given training data set.** For **Part II** of this assignment each team needs to provide 3 deliverables: a written report, a csv file with predictions for the data samples in this new highly imbalanced testing data set and a working code with the details specified below.

1. The report is going to be a continuation of the report you created for project deliverable 2. This report must include the contents from project deliverables 1 and 2 (6-page report) and then add the following sections (no more than additional 2 pages, i.e., a total of no more than 8 pages):
    a. Description of the new predictive models (implement at least 1 new model that you did not use in the previous deliverables and we did not cover in the classes) and provide justification for the following:
        i. model selection
        ii. resampling method(s) utilized during training
        iii. performance metric used to gauge the prediction accuracy of the model
    b. Representation of results: confusion matrix, accuracy, sensitivity, specificity, ROC curve and any other relevant plots
    c. Conclusion: describe why the selected model produces such results
2. Obtain the best model on the training data set and then apply that model on the test data set
    a. Report the predicted (y_hat) values for the samples in the **new test data set** in a csv format
3. A working code should be in python notebook format and satisfy the following criteria:
    a. The code should run without any modification from the TA/Instructor
    b. The code should include comments describing each of the steps