CARNEGIE MELLON UNIVERSITY
DATA, INFERENCE & APPLIED MACHINE LEARNING (COURSE 18-875)
KAGGLE COMPETITION

INSTRUCTIONS

- Submissions should be made via canvas.
- **Single** Python/MATLAB code file(.ipynb or .m) **[Do not Submit checkpoints for .ipynb]**. In addition, each line of code should be documented by text. This demonstrates that the code is unique and owned by the student
- Competition report(.pdf) with full evidence that student participated in a competition and demonstrate a full understanding of each step in the process including textual descriptions of each result (statistics, table, graph etc) represents and insights that can be gained + a screenshot of your ranking on kaggle leaderboard
- Indicate the libraries you have used in your code at the beginning of the report (After the title page)
- Data files (as given)

Submission process:

1. Put source code **file and data files** in a single folder
2. Name of the folder should be the same as your andrew ID
3. **Zip this folder and attach the zipped file on assignment submission page (CANVAS)**
4. After attaching zipped file, click on "Add Another File" from assignment submission page and **attach your report**
5. Submit your assignment

N.B. This process will allow us to compile your reports in **Turnitin** to check for plagiarism.

Specific reasons for a submission being classified as incomplete include:

- Failure to correctly name your folder with your Andrew ID, report, and code file with andrewID_DIAML_Kaggle. For example, mcsharry_DIAML_Kaggle.
- A missing report describing the steps, results, and insights
- A missing dataset required for running the code
- A missing code file such as .ipynb or .m file
- An error in the file path needed to run the code

The submission deadline is **on Friday 9th, December 2022 16:59 Eastern Time (ET) / Friday 9th, December 2022 23:59 Rwandan Time (CAT).**

Kaggle Competition

You are encouraged to enter the Kaggle challenge referencing the Titanic data set. At the end of this course, extra-credit will be given to students based on their final score on the challenge, coinciding with the deadline for this final assignment. Go to this link https://www.kaggle.com/c/titanic-gettingStarted and follow the instructions to register and enter the challenge. After this assignment, you should compare all models and decide which is most appropriate for classifying survival on the Titanic. Please submit a short report explaining the steps that you followed and include your score from Kaggle in the report. Provide evidence that you have achieved this score by including a zipped folder which contains your code and data files just like in assignments as well as a screenshot of the score.