DATA602 - Principle of Data Science (Section PCS2) Assignment 4 Instructions

Due Date: October 29, 2024, by 11:59 pm

Fall 2024 – Dr. Fardina Alam

This assignment consists of the following 3 Parts.

1. **Part01 - Google Colab/ Jupyter Notebook ML Classification Coding Assignment (55 pts)**
2. **Part 02 - Google Colab/ Jupyter Notebook Advanced ML Classification Coding Assignment (40 pts)**
   a. The assignment, dataset, and its instructions are linked here.
   b. Make a copy of these notebook files and complete your assignment.

Submission Instructions: You need to submit both **notebook (.ipynb)** files to Gradescope. Rename the file as follows:
'Homework4_Part01_Classification_DATA602_UID_LastName.ipynb' and
'Homework4_Part02_Classification_DATA602_UID_LastName.ipynb'.

Make sure all outputs are shown. **NO CREDIT** will be given to **notebooks** that do not show code output.
DO NOT modify/remove ANY parts of the questions—answer when indicated.

---

3. Paper Writing Assignment (Points 15): Read the Chapter "Choosing the Right Classification Model" and address the following questions:

- **Summarize the Key Points:** What are the main objectives of the chapter?

- **Challenges in Model Selection:** Why is selecting the right classification model often difficult? Discuss the factors that contribute to this complexity.

- **Write a summary** of the strengths and weaknesses of all the ML models mentioned in the chapter.

Instructions: For each question, respond in no more than one-two paragraphs. Ensure your answers are both concise and clear.

Submission Instructions: Submit a **PDF** file to Gradescope. Using **Overleaf** is preferred. Rename the file as follows: 'Homework04_PaperWriting03_DATA602_UID_LastName.pdf".

*****Note:** Submit the two files individually on Gradescope and not as a zip

REMEMBER: The use of AI tools in generating responses will result in a score of zero and will be followed by an academic violation review.