GRADED ASSIGNMENT

Course	:	JOINT RP-NVIDIA AI INNOVATION & TRAINING PROGRAMME (INTAKE 03)
Module Name	:	Machine Learning Fundamentals
Due Date/Time	:	24 June 2022, 23:59 hours

INSTRUCTIONS TO STUDENT:

1. Develop a machine learning application for **classification** (developed using the Python programming language) with the following functions:
   - Utilises any of the following datasets:
     - Student dataset (https://archive.ics.uci.edu/ml/datasets/Student+Performance)
     - University dataset (https://archive.ics.uci.edu/ml/datasets/University)
     - Flags dataset (https://archive.ics.uci.edu/ml/datasets/Flags)
     - Any other suitable dataset with real-life data (i.e., not randomly generated) with a minimum of 100 data instances and using not less than three features.
   - Reads the dataset from a file into an appropriate internal representation (e.g., Pandas Dataframe).
   - Pre-processes the data if required.
   - Splits the dataset into training and testing subsets.
   - Creates an instance of an appropriate Machine Learning model.
   - Trains the model.
   - Evaluates the trained model and presents the results of the model performance in an appropriate format (with the aid of tables, charts, statistics, metrics, etc.).

2. From the dataset you have selected, you should determine your own objective for your Machine Learning application for **classification**, deciding which attribute(s) should be the target, and which are the suitable features.

3. Your submission should include a one-page write-up with the following information:
    a. A summary on the objectives, features of your application (one paragraph).
    b. The dataset used, the source, the targets and the features.
    c. A summary of the results obtained (one-paragraph).
    d. Clear instructions on how to use your application (if appropriate, not required for Jupyter Notebook submissions).

4. Grading Criteria:
    a. Application (**80%**)
        i. Completeness – relevant components of ML application, as described in the bullet points in (1) above, are present (60%).
        ii. Relevant explanation of code in the form of comments (10%).
        iii. Program executes smoothly with minimal error (10%).

    b. Write-up (**20%**)
        i. Application description, see 3a – 3b above (10%).
        ii. Summary of results and instructions, see 3c – 3d above (10%).
        iii. Marks are given for completeness, clarity and quality.

    c. Bonus (worth 5%)
        i. Extra marks will be given for helpful and informative output or explanation.
        ii. Your bonus points can be used for “top-up” of graded assignment score if it is less than 80%.

5. Submit softcopy of your work via email to koay_seng_tian@rp.edu.sg, and jimmy_goh@rp.edu.sg in zip format which should include – uploads to Microsoft Teams.
    a. your Python code
    b. your Dataset used, or URLs to your Dataset
    c. your write-up

6. Name your notebook file/zip file clearly, using:

    AIITP-03-MachineLearningGA_Your Name

7. Deadline for submission is **24 June 2022, 23:59 hours**. Please start your assignment early. You may submit your assignment before the deadline.

8. Please note that late submission may be penalised according to the lateness of the submission.

Please note that plagiarism will result in significant grade deduction for the submitted work.

~~~ END OF PAPER ~~~