ENEL525 Project: Lesson 1 Exercise for Submission
November 17, 2023

In Lesson 1, your task is to submit only Exercise #3, which can be found in the second part of the lesson (slide 46). To help you complete this submission, instructions and additional notes are provided below. Additionally, for your convenience, the exercise is also outlined on the second page of this document.

Assignment: Lesson 1, Exercise #3
Submission Due Date: November 24, by midnight (5% of your final project grade)

Please submit a single PDF file to the D2L Dropbox containing screenshots of the following:

1. Your complete code executed without any errors (warnings may be present as long as the code functions correctly). Include the "Problems" tab in PyCharm or the executed cells in Jupyter to show that the code is running correctly.
2. A scatter plot of the dataset with points color-coded according to their label, including labeled x and y axes.
3. An image showing the two CSV files "dataset.csv" and "targets.csv" in your directory, confirming their creation.
4. A screenshot of the partial output from using the print command to read the "dataset.csv" file (full output is not required, just a snippet).

Additional Information:

- You can use Jupyter or Google Colab for coding in Python online, which connects you to a remote server. Keep in mind that if you're inactive for extended periods, the server might time out, resulting in an error.
- If this happens, you can: Save your work in a text file or a .py file. Start a new instance at Jupyter's website. Paste your code into a new python3.ipynb file or run the .py file from the terminal to continue where you left off.
- If you prefer, you may also work offline on your own computer using Python version 3.7.0 or later and an IDE like PyCharm 2021. Detailed instructions for setting up your environment can be found in the Lesson 1 videos.

Useful links:

- Pycharm 2021 (Lesson 1 videos explain how to do so in details).
- Pycharm: https://www.jetbrains.com/pycharm/
- Python: https://www.python.org/downloads/
- Jupyter lab to execute python in browser: https://jupyter.org/try
- Google Colab to execute python in browser: https://colab.research.google.com/?utm_source=scs-index

If you have any feedback to improve future lessons or questions regarding the content/exercises, please email: aliadib.arnab@ucalgary.ca

Exercise 3: Dataset Simulation for Two Distinct Classes

Building on Exercise 2, your task now includes some adjustments to create a dataset representing two different classes:

1. Construct an array named "target_1" filled with zeros for the first 1000 data points (these represent the adult class in terms of height and weight).
2. Generate another 1000 data points with a mean height of 50 and mean weight of 25, each group having standard deviations of 10 and 5, respectively. Assign these the value of 1 in a "target_2" array to denote the child class.
3. Merge both sets of data and targets to form a dataset array with a shape of [2000,2] and a target array of [2000,1].
    - Visualize the data in a scatter plot, using different colors for each target label.
    - Remember to label the axes appropriately as height (in kg) for the x-axis and weight (in cm) for the y-axis.
4. Convert both arrays into a Pandas DataFrame and output them as two separate CSV files named "dataset.csv" and "targets.csv".
5. Finally, use the appropriate command to read back the "dataset.csv" file to verify it can be processed correctly. You only need to display a portion of the output using the print command.