Graduate Certificate in Artificial Intelligence with Machine Learning
AIGC 5002 - Machine Learning and Deep Learning
Fall 2023

Lab 1:
September 13, 2023
• Linear Regression and WEKA

Submission guidelines:

- For this lab, you will need to submit 1 PDF file by the end of lab time.
- After you complete all the exercises, convert your Jupyter Notebook (.ipynb) to PDF. Name the PDF as follows: firstname_lastname_LAB1.pdf
- Go to the course Blackboard → Labs folder → Lab Exercise 1 and submit the pdf.

Part 1: Linear Regression in Python

1. Search the web for a dataset with one dependent and one independent variables and with a linear nature in a field of your interest (Finance, gaming, healthcare, IoT, robotics, retail, etc..). (You can search Kaggle.com using phrases like "Linear regression")
2. Download the dataset to your PC.
3. In your notebook, create a markdown cell "Importing the dataset."
4. Create a code cell that will import the dataset into a Pandas DataFrame, as it was shown in the class demo.
5. Visualize the dataset and confirm that there is a linear relationship between the variables.
6. Split the data set to training and testing data (70:30 ratio).
7. Visualize the points in both sets.
8. Fit a linear regression model and evaluate it on the testing dataset.

Part 2: WEKA Installation

Installation

1. Download Weka:
    a. Visit the official Weka download page: Weka Download
    b. Depending on your operating system, choose the right version. For most users, the primary versions will be:
    c. Weka for Windows
    d. Weka for Mac OS X
    e. Weka for Linux
2. Installation (For Windows):
    a. Once downloaded, locate the .exe file (it should be in your Downloads folder or the directory you saved it to).

b. Double-click the .exe file to start the installation.
c. Follow the on-screen instructions, accept the terms, and finish the installation.

3. Running Weka:
   a. Go to the Start menu or Applications folder, and you should find Weka. Click on it to run.

Part 3: Linear Regression and Weka

1. Fit a linear regression model to the same dataset, and compare the results between Python and WEKA.

Enjoy!
