HPE DSI 311 – Introduction to Machine Learning – Spring 2023
Homework Assignment #3
Due Monday (February 27), 11:59 pm (Central)

Your assignment is to create a Jupyter notebook that demonstrates how to do the following (use methods discussed in the class materials shared so far):

Load the dataset in the file named winequality_white.csv and set up a classification problem: predicting the quality value (y variable with seven classes labeled 3, 4, 5, …, 9) based on the values of **all** the other eleven variables (acidity, alcohol, pH, etc.). The goal of this assignment is to see how an MLPClassifier performs compared to a more classical ML model.

1. Train and tune the MLPClassifier (via cross-validation) using at least three different combinations of architecture choices (e.g., number of layers, # of neurons per layer, activation function). (6 points)
2. Study and describe the performance impact of varying at least three different combinations of optimizer parameter values (e.g., solver, epoch, learning rate) for one of the architectures in Step 1. (6 points)
3. Test the performance of the best MLPClassifier from Steps 1 and 2, using two scoring methods of your choice. Discuss in detail your results. (6 points)
4. Train and tune a different classifier that is not a neural network; compare the MLPClassifier test results from Step 3 to that classifier. Discuss in detail your results. (4 points)

What to submit: Please name your h/w submission as follows:
311_lastName_firstName_assignmentNumber.ipynb

How to submit: Please submit homework in Moodle.