PROGRAMMING ASSIGNMENT 1

Designed by: Srikar Kodavati

Instructor: Dr. Wei Ding

Setup:

- Open Programming_Assignment_1.ipynb in colab by file/ open notebook/ upload, choose the file file from downloads.
- Upload winequality-red.csv to Google drive, make sure both Google colab and Google drive are logged in with same account.
- Import libraries and mount Google drive using the code given.

Description:

The dataset contains physiochemical properties which determine the wine quality. In this assignment we are classifying wine based on these parameters.

Programming:

1. Load the dataset from Google drive and select the given

columns in the function get_col().

- Read the data into a data frame using read_csv().
- Create the data frame with three columns 'density', 'pH', 'alcohol'

2. Normalize the three columns using min-max normalization technique in the function norm_data() .

3. Use K Means algorithm to classify the wine data into 3 classes and plot the data in the function classify() .

- Apply K means algorithm on the normalized data use number of classes as 3. Tip: use KMeans() API.
- Save the predicted class in the data frame.
- Plot the three classes in the same graph, and the graph should be a 3D plot. Tip: use fig.add_subplot() and scatter() functions

Grading criteria:

Part 1: Processing data frame (0 to 25 pts).
- Reading data (12.5 pts).
- Selecting columns (12.5 pts).

Part 2: Normalizing data (0-25 pts).

Part 3: Predicting and plotting data (0-50 pts).
- Implementing KMeans model (20 pts).
- Plotting the output data (30 pts).

Submission Deadline: 4:00 PM September 21st , 2023.
Late submission are not accepted and will be awarded zero points.

Submission:

Download the notebook as .ipynb file (File/ download/ download .ipynb). Upload this file on blackboard in the programming assignment 1 section.