Assignment 10

You are provided with a dataset from USA Forensic Science Service which has description of 6 types of glass; defined in terms of their oxide content (i.e. Na, Fe, K, etc). Your task is to use K-Nearest Neighbor (KNN) classifier to classify the glasses.

The original dataset is available at

(https://archive.ics.uci.edu/ml/datasets/glass+identification). For detailed description on the attributes of the dataset,

please refer to the original link of the dataset in the UCI ML repository.

But the shared drive folder have the dataset for your convenience

perform exploratory data analysis on the dataset using Python Pandas, including dropping irrelevant fields for predicted values, and standardization of each attribute.

Following data cleaning, two Scikit-Learn KNN models should be created for two different distance metrics: Square Euclidean and Manhattan distance. The performance of the two models using different distance metrics should be compared in terms of accuracy to the test data and Scikit-Learn Classification Report.

Assignment 11 DataPreprocessing,
Assignment 12 Logistic Regression,
Assignment 13 K-NN,
Assignment 14 SVM,
Assignment 15 Decision Trees,
Assignment 16 Random Forest

Refer the same folder for the .ipynb files and datasets