Date: 17-Mar-2023

Q. Implement the **K-Means Clustering technique from scratch (without using any built-in kmeans() function from the Python libraries) using the Iris dataset**.

Dataset Description: We have shared the *'IRIS.csv'* file which is to be used for implementing this exercise. Refer to *iris.NAMES* file for thorough feature descriptions.

Feature details in brief:

The dataset contains a set of 150 records with 5 features - Petal Length, Petal Width, Sepal Length, Sepal width, and Class(Species):

1. sepal length in cm

2. sepal width in cm

3. petal length in cm

4. petal width in cm

5. class:

       -- Iris Setosa

       -- Iris Versicolour

       -- Iris Virginica

Clustering Details: You should use only the first four features (Petal Length, Petal Width, Sepal Length & Sepal width) to build the clusters using K-means. The number of clusters, **K**, must be found using the "Elbow method". You can use the Sum of Squared Error as the objective function for measuring the "accuracy" of the K-Means Clustering algorithm. Then, using this accuracy measure compare your K-Means model against the model defined using the *KMeans()* function in the module *sklearn.cluster* of the *sklearn* library in Python.