import pandas as pd

import matplotlib.pyplot as plt

from sklearn.cluster import KMeans

from sklearn.preprocessing import StandardScaler

data = pd.read\_csv("iris.csv")

X = data.drop(columns=["Id", "Species"], errors="ignore")

scaler = StandardScaler()

X\_scaled = scaler.fit\_transform(X)

inertia = []

K\_range = range(1, 11)

for k in K\_range:

kmeans = KMeans(n\_clusters=k, random\_state=42, n\_init=10)

kmeans.fit(X\_scaled)

inertia.append(kmeans.inertia\_)

plt.figure(figsize=(8, 5))

plt.plot(K\_range, inertia, 'bo-')

plt.xlabel("Number of Clusters (k)")

plt.ylabel("Inertia (Sum of Squared Distances)")

plt.title("Elbow Method for Optimal k (Iris Dataset)")

plt.show()

k\_opt = 3

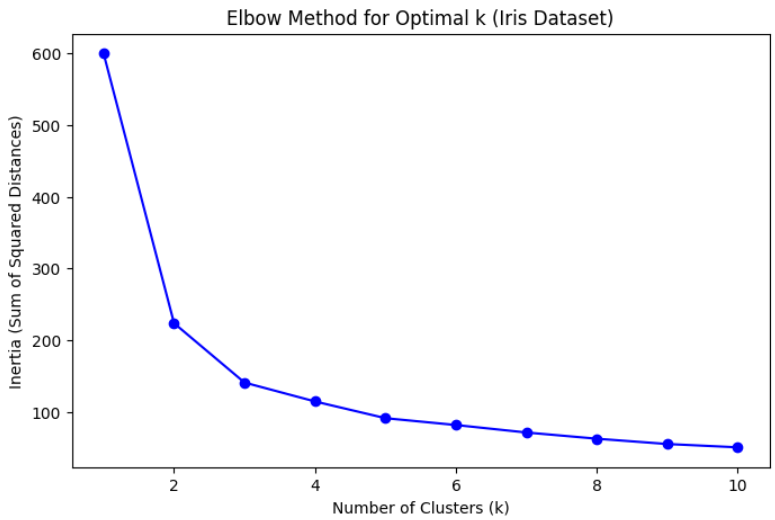
kmeans = KMeans(n\_clusters=k\_opt, random\_state=42, n\_init=10)

clusters = kmeans.fit\_predict(X\_scaled)

data["Cluster"] = clusters

print(data.head())

Output:-



Id SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm Species \

0 1 5.1 3.5 1.4 0.2 Iris-setosa

1 2 4.9 3.0 1.4 0.2 Iris-setosa

2 3 4.7 3.2 1.3 0.2 Iris-setosa

3 4 4.6 3.1 1.5 0.2 Iris-setosa

4 5 5.0 3.6 1.4 0.2 Iris-setosa

Cluster

0 1

1 1

2 1

3 1

4 1