import numpy as np

import pandas as pd

""" 3. Load the dataset as dataframe using pandas """

df = pd.read\_csv('/content/sample\_data/glass source classification dataset.csv')

df

""" 4. Handle missing values if needed """

df.isnull().sum()

from sklearn.impute import SimpleImputer

impute = SimpleImputer(missing\_values=np.nan, strategy='mean')

impute.fit(df[['Ca']])

df['Ca'] = impute.transform(df[['Ca']])

df.isnull().sum()

""" 5.Encode Categorical Features """

from sklearn.preprocessing import LabelEncoder

enc = LabelEncoder()

df['Ba'] = enc.fit\_transform(df['Ba'])

df['Fe'] = enc.fit\_transform(df['Fe'])

df['Type'] = enc.fit\_transform(df['Type'])

df

""" 6. Scale all the values between 0-1 with proper scaling technique """

from sklearn.preprocessing import MinMaxScaler

scaler= MinMaxScaler()

scaler.fit(df)

MinMaxScaler()

df\_scaled = scaler.transform(df)

print("per-feature minimum after scaling:\n {}".format(df\_scaled.min(axis=0)))

print("per-feature maximum after scaling:\n {}".format(df\_scaled.max(axis=0)))

""" 7. Split the dataset into features and labels."""

features = df[["RI","Na","Mg","Al","Si","K","Ca","Ba","Fe"]]

features.head()

lebels = df[["Type"]]

lebels.head()