EXPERIMENT 5

Feature Scaling in Given Dataset

Aim:

To understand the importance of feature scaling by performing a experiment.

Algorithm:

1. Import the required libraries — numpy, pandas, and preprocessing tools from sklearn.

2. Load the dataset using pd.read\_csv() and fill missing values in the Country column with its mode.

3. Separate the dataset into features (X) and labels (Y).

4. Handle missing numerical data in Age and Salary columns using the mean imputation strategy.

5. Apply One-Hot Encoding to the Country column to convert categorical values into numeric form.

6. Concatenate the encoded features with numerical columns to form the final feature set.

7. Perform Standardization using StandardScaler to scale features based on mean and standard deviation.

8. Perform Normalization using MinMaxScaler to scale all feature values between 0 and 1.

9. Display the transformed datasets after both scaling methods.

Program:

import numpy as np

import pandas as pd

df=pd.read\_csv("C:\pre\_process\_data.csv")

df.Country.fillna(df.Country.mode()[0],inplace=True)

features=df.iloc[:,:-1].values

label=df.iloc[:,-1].values

from sklearn.impute import SimpleImputer

age=SimpleImputer(strategy="mean",missing\_values=np.nan)

Salary=SimpleImputer(strategy="mean",missing\_values=np.nan)

age.fit(features[:,[1]])

Salary.fit(features[:,[2]])

SimpleImputer()

features[:,[1]]=age.transform(features[:,[1]])

features[:,[2]]=Salary.transform(features[:,[2]])

print(features)

from sklearn.preprocessing import OneHotEncoder

oh = OneHotEncoder(sparse\_output=False)

Country=oh.fit\_transform(features[:,[0]])

print("\n")

print(Country)

final\_set=np.concatenate((Country,features[:,[1,2]]),axis=1)

print("\n")

print(final\_set)

from sklearn.preprocessing import StandardScaler

sc=StandardScaler()

sc.fit(final\_set)

feat\_standard\_scaler=sc.transform(final\_set)

print("\n")

print(feat\_standard\_scaler)

from sklearn.preprocessing import MinMaxScaler

mms=MinMaxScaler(feature\_range=(0,1))

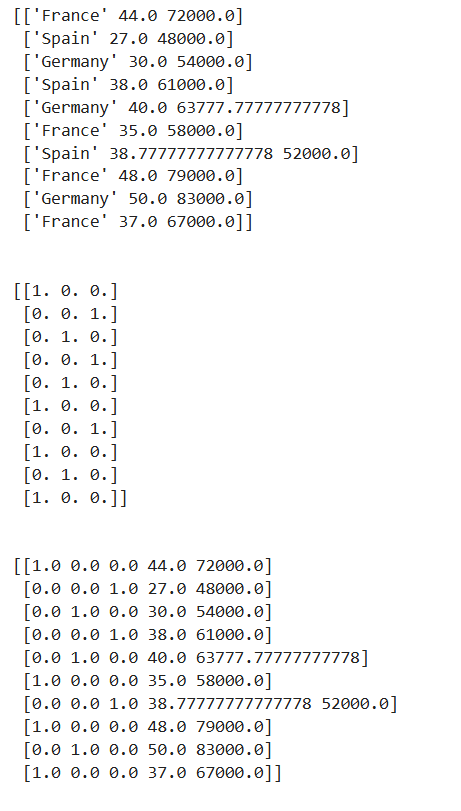
mms.fit(final\_set)

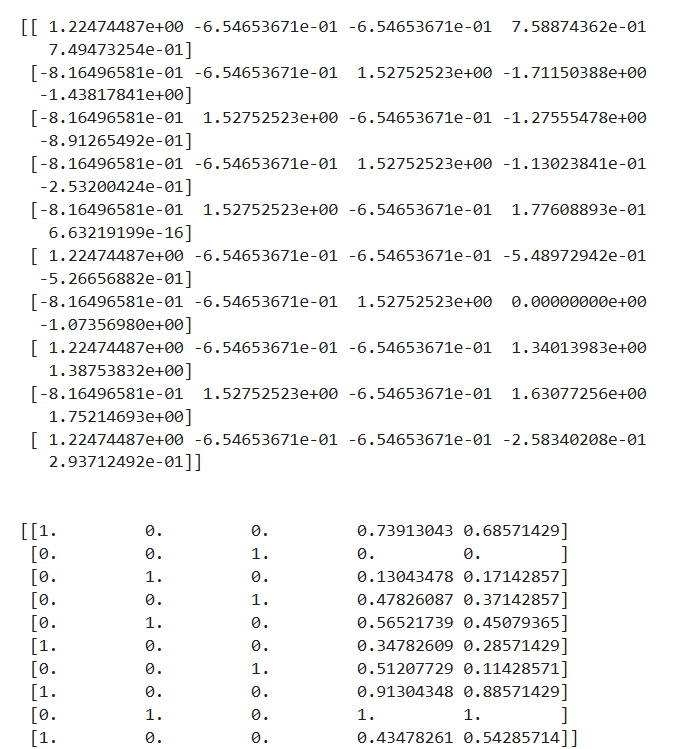
feat\_minmax\_scaler=mms.transform(final\_set)

print("\n")

print(feat\_minmax\_scaler)

Output:





Result:

Hence a python program to standardize the values and perform feature scaling is written and executed successfully.