import pandas as pd

col\_list=["sno","name","rollno","marks","gender"]

df=pd.read\_csv("info.csv",usecols=col\_list)

print(df)

print("-------------------------------------------------------------")

from sklearn.preprocessing import LabelEncoder

gencode=LabelEncoder()

df.gender=gencode.fit\_transform(df.gender)

print("Encoding Complete")

print(df)

print("-------------------------------------------------------------")

from sklearn import preprocessing

df.marks=preprocessing.scale(df.marks)

Marksscale=preprocessing.scale(df.marks)

print("Scaling of marks is completed")

print(df)

print("-------------------------------------------------------------")

Marksshape=Marksscale.reshape(-1,1)

Marksbin=preprocessing.Binarizer(threshold=0.5).transform(Marksshape)

df['marks']=Marksbin

print("Transformation of marks is completed")

print(df)

print("-------------------------------------------------------------")

duplicate=pd.concat([df]\*2,ignore\_index=True)

print("Duplication of records")

print(duplicate)

print("-------------------------------------------------------------")

duprem=pd.DataFrame.drop\_duplicates(duplicate)

print("After duplication removal")

print(duprem)

df['cgpa']=df['cgpa'].fillna(0)

print(df)

df1=df.copy()

df1['cgpa'].fillna(df1['cgpa'].mean(),inplace=True)

printf(df1)