

```
from sklearn.preprocessing import LabelEncoder
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
from pandas.plotting import scatter_matrix
import pandas as pd
import math as m
import random

def datacleaner(data):
    return data.astype(np.int64)#converting float values to integer

def data_cutter(data):
    #cutting the data into 70-30 portions and also cleaning it incase it doesn't contain only integer part
    total_size=len(data)
    train_size=m.floor(0.7*total_size)
    test_size=total_size-train_size
    train_data=data.head(train_size)
    test_data=data.tail(total_size-train_size)
    train_data=datacleaner(train_data)#cleaning process
    test_data=datacleaner(test_data)
    return train_data,test_data

def LinearRegression_predictor(train_x,train_y,test_x,test_y):
    obj=LinearRegression()
    obj.fit(train_x,train_y)
    Yp=obj.predict(test_x)
```

```

print("R square value is:",obj.score(train_x,train_y))

def main():
    path="WinePredictor.csv"
    attributes=['Alcohol','Malic acid','Ash','Alcalinity of ash',
'Magnesium','Total phenols','Flavanoids','Nonflavanoid phenols','
Proanthocyanins','Color intensity','Hue','OD280/OD315 of diluted
wines','Proline']

    #read features of the data and cut them into train set and te
st set
    data_label=pd.read_csv(path,usecols=attributes)
    print(data_label.info)#print the data's characteristics
    train_label,test_label=data_cutter(data_label)
    print(train_label.info())#this is after the data is cleaned a
nd cut
    #into test and trainset,we are printing taining set's info he
re

    #read target of the data and cut them into train set and test
set
    data_target=pd.read_csv(path,usecols=['Class'])
    train_target,test_target=data_cutter(data_target)
    Linear_Regression_predictor(train_label,train_target,test_lab
el,test_target)
    #passing it to linear regression function for fit and accurac
y calculation

if __name__=="__main__":

```

```
main()
```