Problem 1:

Problem | Result

Read the sklearn dataset

Send Feedback

scikit-learn comes with a few small standard datasets that do not require downloading any file from some external website.

These datasets are useful to quickly illustrate the behaviour of the various algorithms implemented in scikit-learn. They are however often too small to be representative of real-world machine learning tasks.

One such dataset is the "Iris dataset".

Your work is to load this dataset from the sklearn library and print the first five rows using the head method of the pandas' library.

Output

Print the first 5 rows of iris dataset(after loading from sklearn toy dataset).

```
## Open and read data file as specified in the question
## Print the required output in given format

from sklearn import datasets

iris = datasets.load_iris()

x = iris.data
y = iris.target

import pandas as pd
df = pd.DataFrame(x)
df.columns = iris.feature_names
print(df.head())
```

Problem 2:

Problem | Result

Split the dataset

Send Feedback

In the previous problem we loaded the iris dataset. Now use this loaded dataset to split the training and testing dataset.

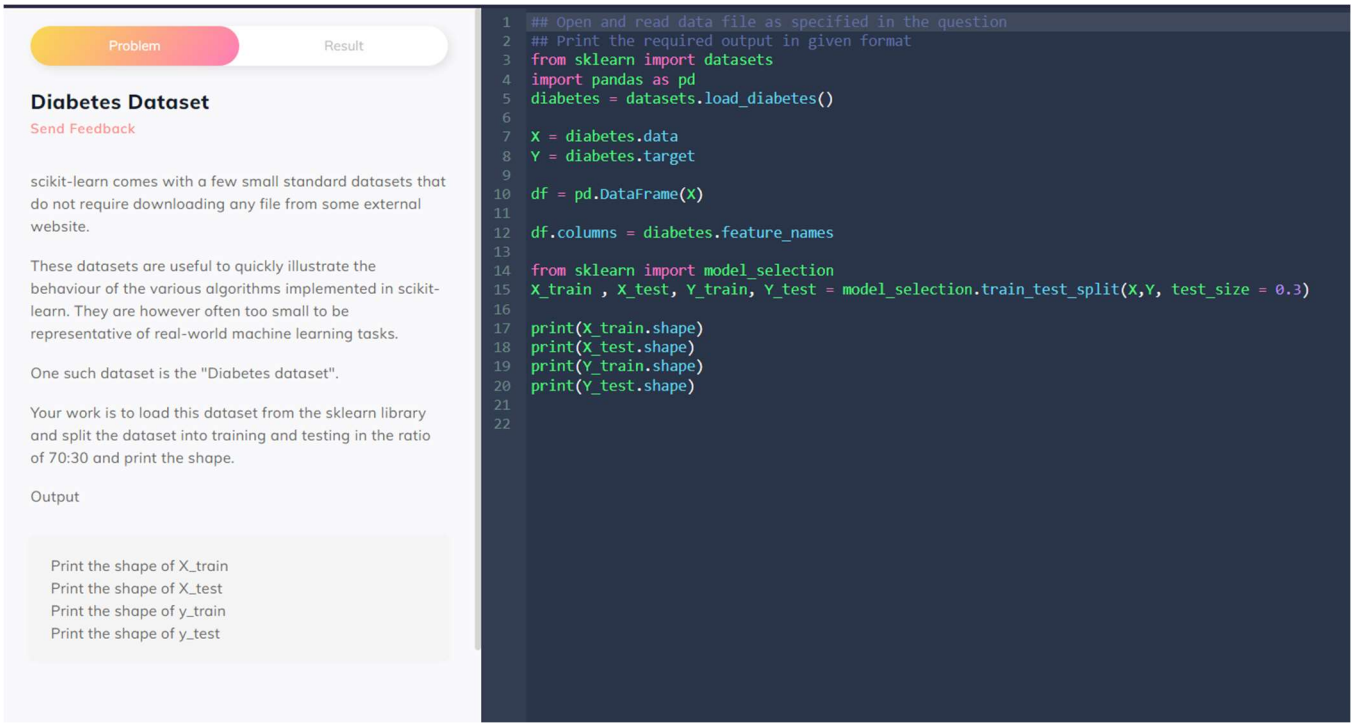
Your work is to again load this dataset from the sklearn library and split the dataset into training and testing in the ratio of 70:30 and print the shape.

Output

Print the shape of X_train
Print the shape of X_test
Print the shape of y_train
Print the shape of y_test

```
## Open and read data file as specified in the question
## Print the required output in given format

from sklearn import datasets
import pandas as pd

iris = datasets.load_iris()
X = iris.data
Y = iris.target

df = pd.DataFrame(X)


df.columns = iris.feature_names

from sklearn import model_selection

X_train, X_test, Y_train, Y_test = model_selection.train_test_split(X,Y, test_size = 0.3)

print(X_train.shape)
print(X_test.shape)
print(Y_train.shape)
print(Y_test.shape)

```

Problem 3:


Problem
Result
Diabetes Dataset
Send Feedback
scikit-learn comes with a few small standard datasets that do not require downloading any file from some external website.
These datasets are useful to quickly illustrate the behaviour of the various algorithms implemented in scikit-learn. They are however often too small to be representative of real-world machine learning tasks.
One such dataset is the "Diabetes dataset".
Your work is to load this dataset from the sklearn library and split the dataset into training and testing in the ratio of 70:30 and print the shape.
Output
Print the shape of X_train
Print the shape of X_test
Print the shape of y_train
Print the shape of y_test
Open and read data file as specified in the question
Print the required output in given format
from sklearn import datasets
import pandas as pd
diabetes = datasets.load_diabetes()
X = diabetes.data
Y = diabetes.target
df = pd.DataFrame(X)
df.columns = diabetes.feature_names
from sklearn import model_selection
X_train , X_test, Y_train, Y_test = model_selection.train_test_split(X,Y, test_size = 0.3)
print(X_train.shape)
print(X_test.shape)
print(Y_train.shape)
print(Y_test.shape)