

MACHINE LEARNING

ASSIGNMENT-1

NAME:- Ahmed Eltayeb

Uni.Rollno: 1917753

CLASS: Btech (CSE 3)

Submitted to: Mrs. Shewta ma'am

► Dataset: iris(download .csv file from Kaggle)

This case study will use data science and machine learning to classify Iris flowers into 3 species:

- Iris-setosa
- Iris-versicolor
- Iris-virginica

Features:

- petal length
- Petal width
- Sepal length
- Sepal width

Please perform the following steps to complete this case study:

1. Create a new empty File in spyder

2. Import all the modules required for:

- numpy
- pandas
- matplotlib
- sklearn/scikit

3. Read the Iris.csvfileintoaPandasDataSetcalled:iris

- Use the pandas read_csv method.

Note:(**Make sure you only have one index column**)

4. Use the describe method to display some stats about the data.

5. Prepare your X and y, using appropriate variable names:

- X: Drop the Species column.
- y: Specify the Species column.

6. Split the data into training and testing data.

- Use sklearn train_test_split to split the data.

7. Create the model and fit to the training data.

- Use the fit method to fit it to the training data.

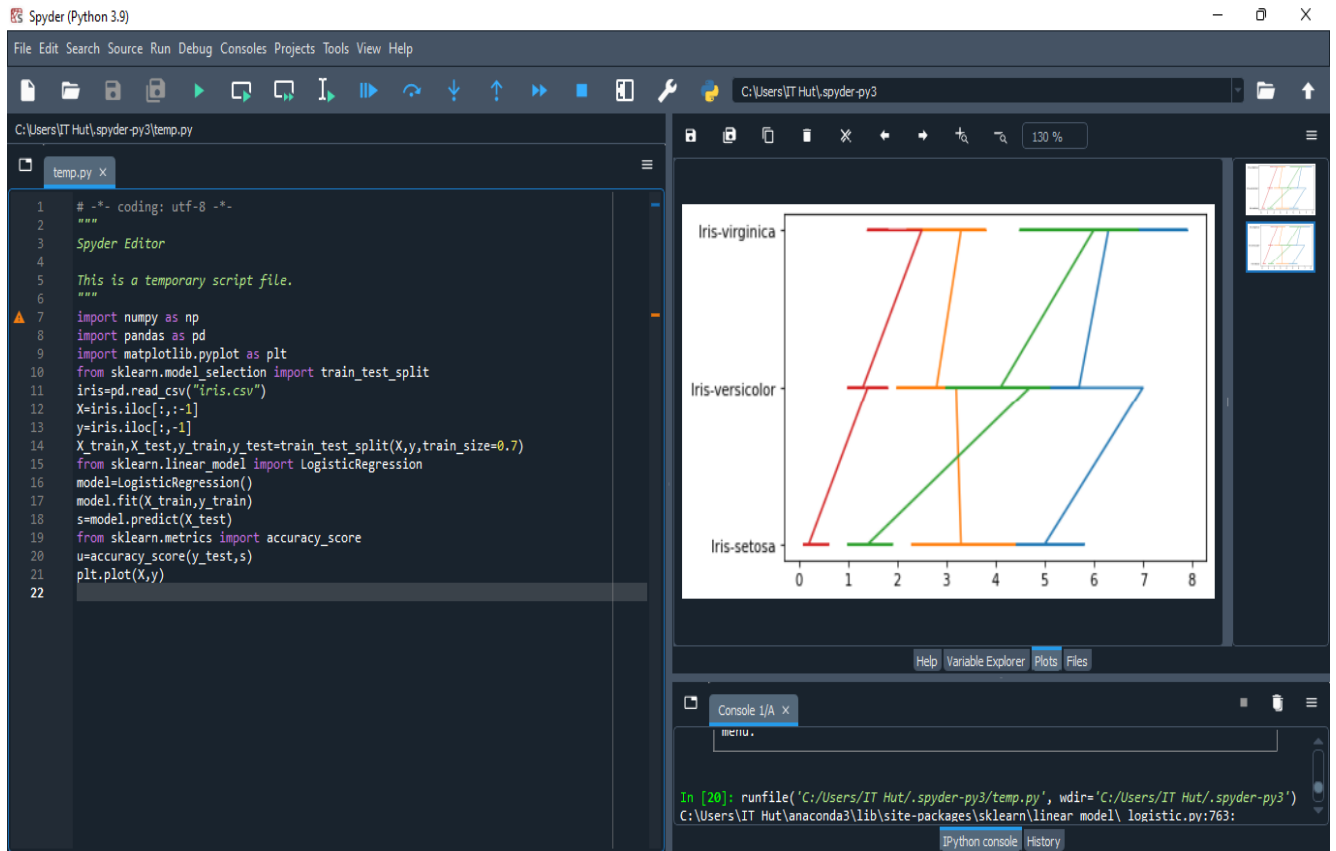
8. Predict values based on testing data.

- Use the predict method to predict values with the x testing data and store them in a variable.

9. Print out the classification report for the y test data and the predictions.

10. Support your Case studies with plotting graphs by using matplotlib

Ans:



Spyder (Python 3.9)

File Edit Search Source Run Debug Consoles Projects Tools View Help

C:\Users\IT Hut\.spyder-py3

temp.py x

```
1 # -*- coding: utf-8 -*-
2 """
3 Spyder Editor
4 This is a temporary script file.
5 """
6
7 import numpy as np
8 import pandas as pd
9 import matplotlib.pyplot as plt
10 from sklearn.model_selection import train_test_split
11 iris=pd.read_csv("iris.csv")
12 X=iris.iloc[:, :-1]
13 y=iris.iloc[:, -1]
14 X_train,X_test,y_train,y_test=train_test_split(X,y,train_size=0.7)
15 from sklearn.linear_model import LogisticRegression
16 model=LogisticRegression()
17 model.fit(X_train,y_train)
18 s=model.predict(X_test)
19 from sklearn.metrics import accuracy_score
20 u=accuracy_score(y_test,s)
21 plt.plot(X,y)
22
```

Name	Type	Size	Value
iris	DataFrame	(150, 5)	Column names: sepal_length, sep...
model	linear_model._logistic.LogisticRegression	1	LogisticRegression object of sk...
s	Array of object	(45,)	ndarray object of numpy module
u	float64	1	0.9333333333333333
X	DataFrame	(150, 4)	Column names: sepal_length, sep...
X_test	DataFrame	(45, 4)	Column names: sepal_length, sep...
X_train	DataFrame	(105, 4)	Column names: sepal_length, sep...
y	Series	(150,)	Series object of pandas.core.series module
y_test	Series	(45,)	Series object of pandas.core.series module
y_train	Series	(105,)	Series object of pandas.core.series module

Help Variable Explorer Plots Files

Console 1/A x

menu.

In [20]: runfile('C:/Users/IT Hut/.spyder-py3/temp.py', wdir='C:/Users/IT Hut/.spyder-py3')

C:\Users\IT Hut\anaconda3\lib\site-packages\sklearn\linear_model\logistic.py:763:

Python console History

So, Accuracy of model is 92% .