08/08/2024, 19:37 Classification - Colab

import pandas as pd
from sklearn import datasets
X,Y=datasets.load_iris(return_X_y=True,as_frame=True)

X.describe()

→		sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
	count	150.000000	150.000000	150.000000	150.000000
	mean	5.843333	3.057333	3.758000	1.199333
	std	0.828066	0.435866	1.765298	0.762238
	min	4.300000	2.000000	1.000000	0.100000
	25%	5.100000	2.800000	1.600000	0.300000
	50%	5.800000	3.000000	4.350000	1.300000
	75%	6.400000	3.300000	5.100000	1.800000
	max	7.900000	4.400000	6.900000	2.500000

X.isna().sum()

sepal length (cm) 0
sepal width (cm) 0
petal length (cm) 0
petal width (cm) 0
dtype: int64

from sklearn.linear_model import LogisticRegression
model=LogisticRegression(tol=0.1,solver='lbfgs')

Y

```
0 0

1 0

2 0

3 0

4 0

...

145 2

146 2

147 2

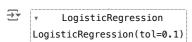
148 2

149 2

Name: target, Length: 150, dtype: int64
```

from sklearn.model_selection import train_test_split
X_train,X_test,Y_train,Y_test=train_test_split(X,Y,random_state=96,test_size=0.3)

model.fit(X_train,Y_train)



model.score(X_train,Y_train)

→ 0.9714285714285714

model.score(X_test,Y_test)

Y_predict=model.predict(X_test)

from sklearn.metrics import r2_score

r2_score(Y_test,Y_predict)

→ 0.964171974522293