

Teams and Channels | General | Machine Learning - ColabML practice - Colab

colab.research.google.com/drive/1D1Dhy_mW7Ifh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_data

HR_comma_sep.csv

heart_disease.csv

Disk75.09 GB available

+ Code + Text

[2] import pandas as pd
a=pd.read_csv('HR_comma_sep.csv')
a

	satisfaction_level	last_evaluation	number_project	average_monthly_hours	time_spent_company	Work_accident	left	promoti
0	0.38	0.53	2	157	3	0	1	
1	0.80	0.86	5	262	6	0	1	
2	0.11	0.88	7	272	4	0	1	
3	0.72	0.87	5	223	5	0	1	
4	0.37	0.52	2	159	3	0	1	
...
14994	0.40	0.57	2	151	3	0	1	
14995	0.37	0.48	2	160	3	0	1	
14996	0.37	0.53	2	143	3	0	1	
14997	0.11	0.96	6	280	4	0	1	
14998	0.37	0.52	2	158	3	0	1	

14999 rows x 10 columns

0s

completed at 8:04 PM

20:06

28-11-2024

Teams and Channels | General | Machine Learning - ColabML practice - Colab

colab.research.google.com/drive/1DlDhy_mW7Ifh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_data

HR_comma_sep.csv

heart_disease.csv

Disk75.09 GB available

+ Code + Text

[6] b['income']=b['salary'].apply(lambda x:1 if x=='low' else 0)

b

	satisfaction_level	last_evaluation	number_project	average_monthly_hours	time_spend_company	Work_accident	left	promoti
0	0.38	0.53	2	157	3	0	1	
1	0.80	0.86	5	262	6	0	1	
2	0.11	0.88	7	272	4	0	1	
3	0.72	0.87	5	223	5	0	1	
4	0.37	0.52	2	159	3	0	1	
...
14994	0.40	0.57	2	151	3	0	1	
14995	0.37	0.48	2	160	3	0	1	
14996	0.37	0.53	2	143	3	0	1	
14997	0.11	0.96	6	280	4	0	1	
14998	0.37	0.52	2	158	3	0	1	

0s completed at 8:04 PM

20:0628-11-2024

Teams and Channels | General | Machine Learning - ColabML practice - Colab

colab.research.google.com/drive/1DlDhy_mW7lfh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_dataHR_comma_sep.csvheart_disease.csv

Disk75.09 GB available

+ Code + Text

149970.110.965280401
149980.370.522158301
14999 rows × 10 columns

[7] c=b.drop(['salary','income'],axis=1)
c

	satisfaction_level	last_evaluation	number_project	average_monthly_hours	time_spent_company	Work_accident	left	promotion
0	0.38	0.53	2	157	3	0	1	
1	0.80	0.86	5	262	6	0	1	
2	0.11	0.88	7	272	4	0	1	
3	0.72	0.87	5	223	5	0	1	
4	0.37	0.52	2	159	3	0	1	
...
14994	0.40	0.57	2	151	3	0	1	
14995	0.37	0.48	2	160	3	0	1	
14996	0.37	0.56	2	160	3	0	1	

0s completed at 8:04 PM

20:0628-11-2024

Teams and Channels | General | Machine Learning - ColabML practice - Colab

colab.research.google.com/drive/1DlDhy_mW7lfh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_dataHR_comma_sep.csvheart_disease.csv

Disk75.09 GB available

+ Code + Text

14999 rows x 8 columns

```
d=b['income']
d
```

income	
0	1
1	0
2	0
3	1
4	1
...	...
14994	1
14995	1
14996	1
14997	1

0s completed at 8:04 PM

20:0628-11-2024

Teams and Channels | General | Machine Learning - Colab | ML practice - Colab

colab.research.google.com/drive/1DlDhy_mW7lfh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

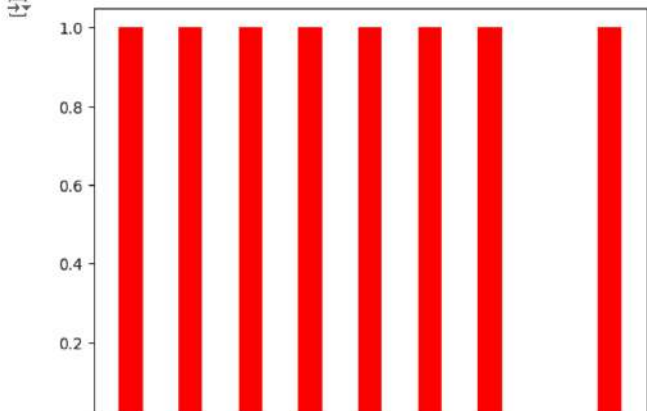
Files

sample_data
HR_comma_sep.csv
heart_disease.csv

75.09 GB available

+ Code + Text

```
[17] import matplotlib.pyplot as plt
plt.bar(c['time_spend_company'],d,width=0.4,color='red')
plt.show()
```



Category	time_spend_company
1	1.0
2	1.0
3	1.0
4	1.0
5	1.0
6	1.0
7	1.0
8	1.0

0s completed at 8:04 PM

Search

20:06 28-11-2024

Teams and Channels | General | Machine Learning - ColabML practice - Colab

colab.research.google.com/drive/1D1Dhy_mW7Ifh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

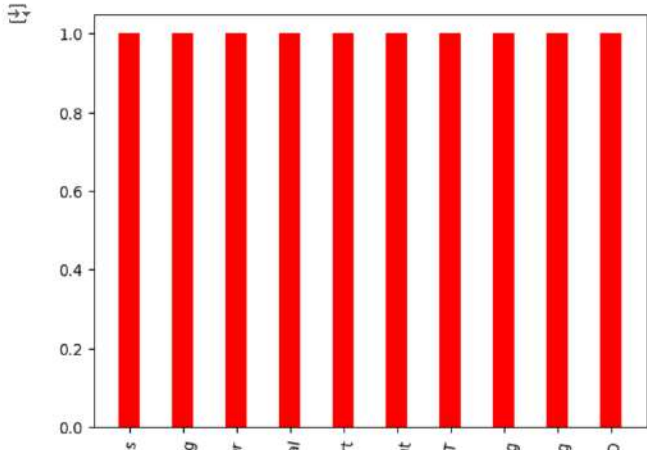
Files

sample_dataHR_comma_sep.csvheart_disease.csv

Disk75.09 GB available

+ Code + Text

[20] plt.bar(e,d,width=0.4,color='red')
plt.xticks(rotation=80)
plt.show()



Category	Value
es	1.0
ng	1.0
hr	1.0
al	1.0
xt	1.0
nt	1.0
fr	1.0
rg	1.0
rg	1.0
ld	1.0

0s completed at 8:04 PM

Search

20:0628-11-2024

Teams and Channels | General | Machine Learning - Colab | ML practice - Colab

colab.research.google.com/drive/1D1Dhy_mW7lfh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_data

HR_comma_sep.csv

heart_disease.csv

+ Code + Text

```
[10] from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(c,d,test_size=0.3)

from sklearn.linear_model import LogisticRegression
z=LogisticRegression()
y=z.fit(x_train,y_train)
y

/usr/local/lib/python3.10/dist-packages/sklearn/linear_model/_logistic.py:469: ConvergenceWarning: lbfgs failed to converge (status
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
n_iter_i = _check_optimize_result(
LogisticRegression
LogisticRegression()
```

[23] y.score(x_test,y_test)

0s completed at 8:04 PM

75.09 GB available

20:06 28-11-2024

Teams and Channels | General | Machine Learning - Colab | ML practice - Colab

colab.research.google.com/drive/1D1Dhy_mW7Ifh5IQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_data

HR_comma_sep.csv

heart_disease.csv

+ Code + Text

[23] y.score(x_test,y_test)

0.546

[24] y.predict(x_test)

array([1, 0, 0, ..., 0, 0, 0])

import pandas as pd

a=pd.read_csv('heart_disease.csv')

a

	Gender	age	education	currentSmoker	cigsPerDay	BPMeds	prevalentStroke	prevalentHyp	diabetes	totChol	sysBP	diaBP
0	Male	39	postgraduate	0	0.0	0.0	no	0	0	195.0	106.0	70.0
1	Female	46	primaryschool	0	0.0	0.0	no	0	0	250.0	121.0	81.0
2	Male	48	uneducated	1	20.0	0.0	no	0	0	245.0	127.5	80.0
3	Female	61	graduate	1	30.0	0.0	no	1	0	225.0	150.0	95.0
4	Female	46	graduate	1	23.0	0.0	no	0	0	285.0	130.0	84.0

0s completed at 8:04 PM

75.09 GB available

20:07 28-11-2024

Teams and Channels | General | Machine Learning - ColabML practice - Colab

colab.research.google.com/drive/1D1Dhy_mW7Ifh5IQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_dataHR_comma_sep.csvheart_disease.csv

Disk75.09 GB available

+ Code + Text

[24] y.predict(x_test)

array([1, 0, 0, ..., 0, 0, 0])

import pandas as pd

a=pd.read_csv('heart_disease.csv')

a

	Gender	age	education	currentSmoker	cigsPerDay	BPMeds	prevalentStroke	prevalentHyp	diabetes	totChol	sysBP	diaBP
0	Male	39	postgraduate	0	0.0	0.0	no	0	0	195.0	106.0	70.0
1	Female	46	primaryschool	0	0.0	0.0	no	0	0	250.0	121.0	81.0
2	Male	48	uneducated	1	20.0	0.0	no	0	0	245.0	127.5	80.0
3	Female	61	graduate	1	30.0	0.0	no	1	0	225.0	150.0	95.0
4	Female	46	graduate	1	23.0	0.0	no	0	0	285.0	130.0	84.0
...
4233	Male	50	uneducated	1	1.0	0.0	no	1	0	313.0	179.0	92.0
4234	Male	51	graduate	1	43.0	0.0	no	0	0	207.0	126.5	80.0
4235	Female	48	primaryschool	1	20.0	NaN	no	0	0	248.0	131.0	72.0

0s completed at 8:04 PM

20:0728-11-2024

Teams and Channels | General | Machine Learning - ColabML practice - Colab

colab.research.google.com/drive/1D1Dhy_mW7Ifh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_dataHR_comma_sep.csvheart_disease.csv

4238 rows x 16 columns

```
from sklearn.preprocessing import LabelEncoder
z=LabelEncoder()
z
```

LabelEncoder

```
[32] y=a
      y['Heart_stroke']=z.fit_transform(y['Heart_ stroke'])
      y
```

perDay	BPMeds	prevalentStroke	prevalentHyp	diabetes	totChol	sysBP	diaBP	BMI	heartRate	glucose	Heart_stroke	Heart_stroke
0.0	0.0	no	0	0	195.0	106.0	70.0	26.97	80.0	77.0	No	0
0.0	0.0	no	0	0	250.0	121.0	81.0	28.73	95.0	76.0	No	0
20.0	0.0	no	0	0	245.0	127.5	80.0	25.34	75.0	70.0	No	0

0s completed at 8:04 PM

75.09 GB available

20:07 28-11-2024

Teams and Channels | General | Machine Learning - ColabML practice - Colab

colab.research.google.com/drive/1DlDhy_mW7Ifh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_dataHR_comma_sep.csvheart_disease.csv

Disk75.09 GB available

+ Code + Text

```
[32] y=a
      y['Heart_stroke']=z.fit_transform(y['Heart_stroke'])
      y
```

perDay	BPMeds	prevalentStroke	prevalentHyp	diabetes	totChol	sysBP	diaBP	BMI	heartRate	glucose	Heart_stroke	Heart_stroke
0.0	0.0	no	0	0	195.0	106.0	70.0	26.97	80.0	77.0	No	0
0.0	0.0	no	0	0	250.0	121.0	81.0	28.73	95.0	76.0	No	0
20.0	0.0	no	0	0	245.0	127.5	80.0	25.34	75.0	70.0	No	0
30.0	0.0	no	1	0	225.0	150.0	95.0	28.58	65.0	103.0	yes	1
23.0	0.0	no	0	0	285.0	130.0	84.0	23.10	85.0	85.0	No	0
...
1.0	0.0	no	1	0	313.0	179.0	92.0	25.97	66.0	86.0	yes	1
43.0	0.0	no	0	0	207.0	126.5	80.0	19.71	65.0	68.0	No	0
20.0	NaN	no	0	0	248.0	131.0	72.0	22.00	84.0	86.0	No	0
15.0	0.0	no	0	0	210.0	126.5	87.0	19.16	86.0	NaN	No	0
0.0	0.0	no	0	0	260.0	132.5	82.0	21.47	80.0	107.0	No	0

0s completed at 8:04 PM

Search

ENG IN

20:0728-11-2024

Teams and Channels | General | xMachine Learning - Colab xML practice - Colab x

colab.research.google.com/drive/1DlDhy_mW7lfh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_dataHR_comma_sep.csvheart_disease.csv

75.09 GB available

Next steps: View recommended plotsNew interactive sheet

[34] b=y.drop(['Heart_stroke','Heart_stroke','Gender','education','prevalentStroke'],axis=1)

	age	currentSmoker	cigsPerDay	BPMeds	prevalentHyp	diabetes	totChol	sysBP	diaBP	BMI	heartRate	glucose
0	39	0	0.0	0.0	0	0	195.0	106.0	70.0	26.97	80.0	77.0
1	46	0	0.0	0.0	0	0	250.0	121.0	81.0	28.73	95.0	76.0
2	48	1	20.0	0.0	0	0	245.0	127.5	80.0	25.34	75.0	70.0
3	61	1	30.0	0.0	1	0	225.0	150.0	95.0	28.58	65.0	103.0
4	46	1	23.0	0.0	0	0	285.0	130.0	84.0	23.10	85.0	85.0
...
4233	50	1	1.0	0.0	1	0	313.0	179.0	92.0	25.97	66.0	86.0
4234	51	1	43.0	0.0	0	0	207.0	126.5	80.0	19.71	65.0	68.0
4235	48	1	20.0	NaN	0	0	248.0	131.0	72.0	22.00	84.0	86.0
4236	44	1	15.0	0.0	0	0	210.0	126.5	87.0	19.16	86.0	NaN

0s completed at 8:04 PM

Search

ENG IN20:0728-11-2024

Teams and Channels | General | x Machine Learning - Colab x ML practice - Colab x

colab.research.google.com/drive/1DlDhy_mW7lfh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
- HR_comma_sep.csv
- heart_disease.csv

Next steps: View recommended plots New interactive sheet

[36] b.isnull().sum()

	0
age	0
currentSmoker	0
cigsPerDay	29
BPMeds	53
prevalentHyp	0
diabetes	0
totChol	50
sysBP	0
diaBP	0
BMI	19
heartRate	1

0s completed at 8:04 PM

20:07 28-11-2024

Teams and Channels | General | Machine Learning - Colab | ML practice - Colab

colab.research.google.com/drive/1D1Dhy_mW7Ifh5IQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_data

HR_comma_sep.csv

heart_disease.csv

+ Code + Text

heartRate 1

glucose 388

dtype: int64

b.fillna(b.mean(),inplace=True)

[38] b.isnull().sum()

0

age 0

currentSmoker 0

cigsPerDay 0

BPMeds 0

prevalentHyp 0

diabetes 0

totChol 0

sysBP 0

0s completed at 8:04 PM

Disk 75.09 GB available

20:07 28-11-2024

Teams and Channels | General | Machine Learning - ColabML practice - Colab

colab.research.google.com/drive/1DlDhy_mW7Ifh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_data

HR_comma_sep.csv

heart_disease.csv

+ Code + Text

[35] c=y['Heart_stroke']

c

Heart_stroke

0	0
1	0
2	0
3	1
4	0
...	...
4233	1
4234	0
4235	0
4236	0
4237	0

4238 rows × 1 columns

0s completed at 8:04 PM

1

Search

ENG IN

20:07 28-11-2024

Teams and Channels | General | x Machine Learning - Colab x ML practice - Colab x

colab.research.google.com/drive/1D1Dhy_mW7Ifh5IQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

- sample_data
- HR_comma_sep.csv
- heart_disease.csv

4238 rows x 1 columns

dtype: int64

```
[40] from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(b,c,test_size=0.2)

[45] from sklearn.preprocessing import StandardScaler
g=StandardScaler()
h=g.fit_transform(x_train.values)
h
```

```
array([[ -0.18075604, -0.97898189, -0.75125662, ..., -1.15010198,
        -0.49119145,  0.11707434],
       [ 1.69027593,  1.02146935,  1.76198796, ...,  0.13821623,
        -0.49119145, -0.22231004],
       [ 1.80721542, -0.97898189, -0.75125662, ...,  1.49973434,
        -0.07705769,  1.0503814 ],
       ...,
       [ 1.45639693, -0.97898189, -0.75125662, ..., -0.10578343,
        -0.90532522, -0.0950409 ],
       [ 0.28700195, -0.97898189, -0.75125662, ..., -0.25706322,
        -0.15988444, -0.22231004],
       [-0.99933252, -0.97898189, -0.75125662, ..., -1.73082118,
```

0s completed at 8:04 PM

75.09 GB available

20:07 28-11-2024

Teams and Channels | General | xMachine Learning - Colab xML practice - Colab x

colab.research.google.com/drive/1DdHy_mW7Ifh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

GmailYouTube | meaning in the Ca...WhatsApp (1) Teams and Chan...Python Basic.ipynb...TimeSeries - ColabMachine Learning - ...

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_dataHR_comma_sep.csvheart_disease.csv

+ Code + Text

```
-0.15988444, -0.22231004],
[-0.99933252, -0.97898189, -0.75125662, ..., -1.73082118,
-0.90532522, -0.94350186]])

[47] from sklearn.linear_model import LogisticRegression
u=LogisticRegression()
v=u.fit(x_train,y_train)
v

/usr/local/lib/python3.10/dist-packages/sklearn/linear_model/_logistic.py:469: ConvergenceWarning: lbfgs failed to converge (status
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
n_iter_1 = _check_optimize_result(
LogisticRegression
LogisticRegression()
```

[48] v.score(x_test,y_test)

0s completed at 8:04 PM

Disk75.09 GB available

0s completed at 8:04 PM

Search

ENG IN20:0728-11-2024

[illegible]

The screenshot displays a Google Colab notebook interface. The top navigation bar shows the URL `colab.research.google.com/drive/1DlDhy_mW7Ifh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn`. The notebook title is "ML practice". The left sidebar shows a file explorer with a folder named "sample_data" containing two files: "HR_comma_sep.csv" and "heart_disease.csv". The main code area contains two code cells. The first cell, executed successfully (indicated by a green checkmark), defines a confusion matrix `i` for a binary classification problem. The matrix is a 2x2 array: $\begin{bmatrix} 713 & 3 \\ 127 & 5 \end{bmatrix}$. The second cell, also executed successfully, imports `seaborn` and generates a heatmap of the confusion matrix `i` with `annot=True`. The resulting heatmap shows two classes on the x and y axes. The diagonal elements are highlighted in light orange, representing correct classifications: 713 for the top-left cell and 5 for the bottom-right cell. The off-diagonal elements are dark blue, representing misclassifications: 3 for the top-right cell and 127 for the bottom-left cell. A color bar on the right indicates the count of instances, ranging from 400 to 700. The bottom status bar shows the notebook is "completed at 8:04 PM" with 0 seconds of runtime.

Teams and Channels | General | Machine Learning - ColabML practice - Colab

colab.research.google.com/drive/1D1Dhy_mW7Ifh5lQdniTEAqy7GJc1dCp9#scrollTo=I9buA7AYMRqn

ML practice

File Edit View Insert Runtime Tools Help All changes saved

Files

sample_data

HR_comma_sep.csv

heart_disease.csv

sns.heatmap(i,annot=True)

<Axes: >

01

7.1e+023

1.3e+025

01

0s completed at 8:04 PM

20:07 28-11-2024

