

matplotlib - Colaboratorytumert\_dataset - Colaboratorytumert\_data\_by\_svm - Col

colab.research.google.com/drive/1t0uxrhnhOoIVEKAQUWrgFudXahKQDIdu#scrollTo=S8PwwKGF2qOl

GmailYouTubeMapsNewsTranslateWhatsAppDownloadsJCO/OR User...earPages - NCS[Se...ZType - Typin...ln(1) Feed | Link...GitHubKaggle: Your...

tumert\_dataset

FileEditViewInsertRuntimeToolsHelpAll changes saved

Table of contents

Section

{x}

+ Code+ Text

✓

import numpy as npimport pandas as pddata=pd.read\_csv("/content/KNNAlgorithmDataset.csv")data

✓

id diagnosis radius\_mean texture\_mean perimeter\_mean area\_mean smoothness\_mean compactness\_mean concavity\_mean concave points\_mean ... texture\_worst perimeter\_worst area\_worst smoothness\_worst

0 842302 M 17.99 10.38 122.80 1001.0 0.11840 0.27760 0.30010 0.14710 ... 17.33 184.60 2019.0 0.1622

1 842517 M 20.57 17.77 132.90 1326.0 0.08474 0.07864 0.08690 0.07017 ... 23.41 158.80 1956.0 0.1238

2 84300903 M 19.69 21.25 130.00 1203.0 0.10960 0.15990 0.19740 0.12790 ... 25.53 152.50 1709.0 0.1444

3 84348301 M 11.42 20.38 77.58 386.1 0.14250 0.28390 0.24140 0.10520 ... 26.50 98.87 567.7 0.2098

4 84358402 M 20.29 14.34 135.10 1297.0 0.10030 0.13280 0.19800 0.10430 ... 16.67 152.20 1575.0 0.1374

...

564 926424 M 21.56 22.39 142.00 1479.0 0.11100 0.11590 0.24390 0.13890 ... 26.40 166.10 2027.0 0.1410

565 926682 M 20.13 28.25 131.20 1261.0 0.09780 0.10340 0.14400 0.09791 ... 38.25 155.00 1731.0 0.1166

566 926954 M 16.60 28.08 108.30 858.1 0.08455 0.10230 0.09251 0.05302 ... 34.12 126.70 1124.0 0.1139

567 927241 M 20.60 29.33 140.10 1265.0 0.11780 0.27700 0.35140 0.15200 ... 39.42 184.60 1821.0 0.1650

568 92751 B 7.76 24.54 47.92 181.0 0.05263 0.04362 0.00000 0.00000 ... 30.37 59.16 268.6 0.0899

569 rows x 33 columns

✓

data.isna().sum()

✓

data\_upt=pd.get\_dummies(data,{'diagnosis'})data\_upt

[12]

cnt=data.groupby('diagnosis')['diagnosis'].count()cnt

[23]

del data\_upt['id']

[24]

del data\_upt['diagnosis\_M']

0s completed at 8:43 AM

matplotlib - Colaboratorytumert\_dataset - Colaboratorytumert\_data\_by\_svm - Col

colab.research.google.com/drive/1t0uxrhnhOoiVEKAQUWrgFudXahKQDIdu#scrollTo=BQ-ERSGw3Gd

GmailYouTubeMapsNewsTranslateWhatsAppDownloadsJCO / OR User...earPages - NCS[Se...ZType - Typin... (1) Feed | Link...GitHubKaggle: Your...

tumert\_dataset

FileEditViewInsertRuntimeToolsHelp

Files

{x}

sample\_data

KNNAgorithmDataset.csv

569

927241

M

20.60

29.33

140.10

1265.0

0.11780

0.27700

0.35140

0.15200

...

39.42

184.00

1821.0

0.1050

568 92751 B 7.76 24.54 47.92 181.0 0.05263 0.04362 0.00000 0.00000 ... 30.37 59.16 268.6 0.0899

569 rows x 33 columns

data.isna().sum()

id

0

diagnosis

0

radius\_mean

0

texture\_mean

0

perimeter\_mean

0

area\_mean

0

smoothness\_mean

0

compactness\_mean

0

concavity\_mean

0

concave points\_mean

0

symmetry\_mean

0

fractal\_dimension\_mean

0

radius\_se

0

texture\_se

0

perimeter\_se

0

area\_se

0

smoothness\_se

0

compactness\_se

0

concavity\_se

0

concave points\_se

0

symmetry\_se

0

fractal\_dimension\_se

0

radius\_worst

0

texture\_worst

0

perimeter\_worst

0

area\_worst

0

smoothness\_worst

0

compactness\_worst

0

concavity\_worst

0

concave points\_worst

0

symmetry\_worst

0

fractal\_dimension\_worst

0

Unnamed: 32

569

dtype: int64

[13] data\_upd=pd.get\_dummies(data,{'diagnosis'})

data\_upd

[12] cnt=data.groupby('diagnosis') ['diagnosis'].count()

0s completed at 8:57 AM

matplotlib - Colaboratorytumert\_dataset - Colaboratorytumert\_data\_by\_svm - Colab

colab.research.google.com/drive/1t0uxrhhOoIVEKAQUWrgFudXahKQDIdu#scrollTo=K\_UknW6zwzf

GmailYouTubeMapsNewsTranslateWhatsAppDownloadsJCO/OR User...Pages - NCS[Se...ZType - Typin...ln(1) Feed | Link...GitHubKaggle: Your...

tumert\_dataset

FileEditViewInsertRuntimeToolsHelpSaving...

Files

-.configsample\_dataKNNAAlgorithmDataset.csv

+ Code+ Text

[38]

concaity\_worst0  
concaie points\_worst0  
symetry\_worst0  
fractal\_dimension\_worst0  
Unnamed: 32569  
dtype: int64

data\_upt=pd.get\_dummies(data,{'diagnosis'}) #used to encoding  
data\_upt

idradius\_mean texture\_mean perimeter\_mean area\_mean smoothness\_mean compactness\_mean concaity\_mean concave points\_mean symmetry\_mean ... area\_worst smoothness\_worst compactness\_worst concav

084230217.9910.38122.801001.00.118400.277600.300100.147100.2419...2019.00.162200.66560

184251720.5717.77132.901326.00.084740.078640.086900.070170.1812...1956.00.123800.18660

28430090319.6921.25130.001203.00.109600.159900.197400.127900.2069...1709.00.144400.42450

38434830111.4220.3877.58386.10.142500.283900.241400.105200.2597...567.70.209800.86630

48435840220.2914.34135.101297.00.100300.132800.198000.104300.1809...1575.00.137400.20500

... ..

56492642421.5622.39142.001479.00.111000.115900.243900.138900.1726...2027.00.141000.21130

56592668220.1328.25131.201261.00.097800.103400.144000.097910.1752...1731.00.116600.19220

56692695416.6028.08108.30858.10.084550.102300.092510.053020.1590...1124.00.113900.30940

56792724120.6029.33140.101265.00.117800.277000.351400.152000.2397...1821.00.165000.86810

568927517.7624.5447.92181.00.052630.043620.000000.000000.1587...268.60.089960.06444

569 rows x 34 columns

[12] cnt=data.groupby('diagnosis')['diagnosis'].count()  
cnt

[23] del data\_upt['id']

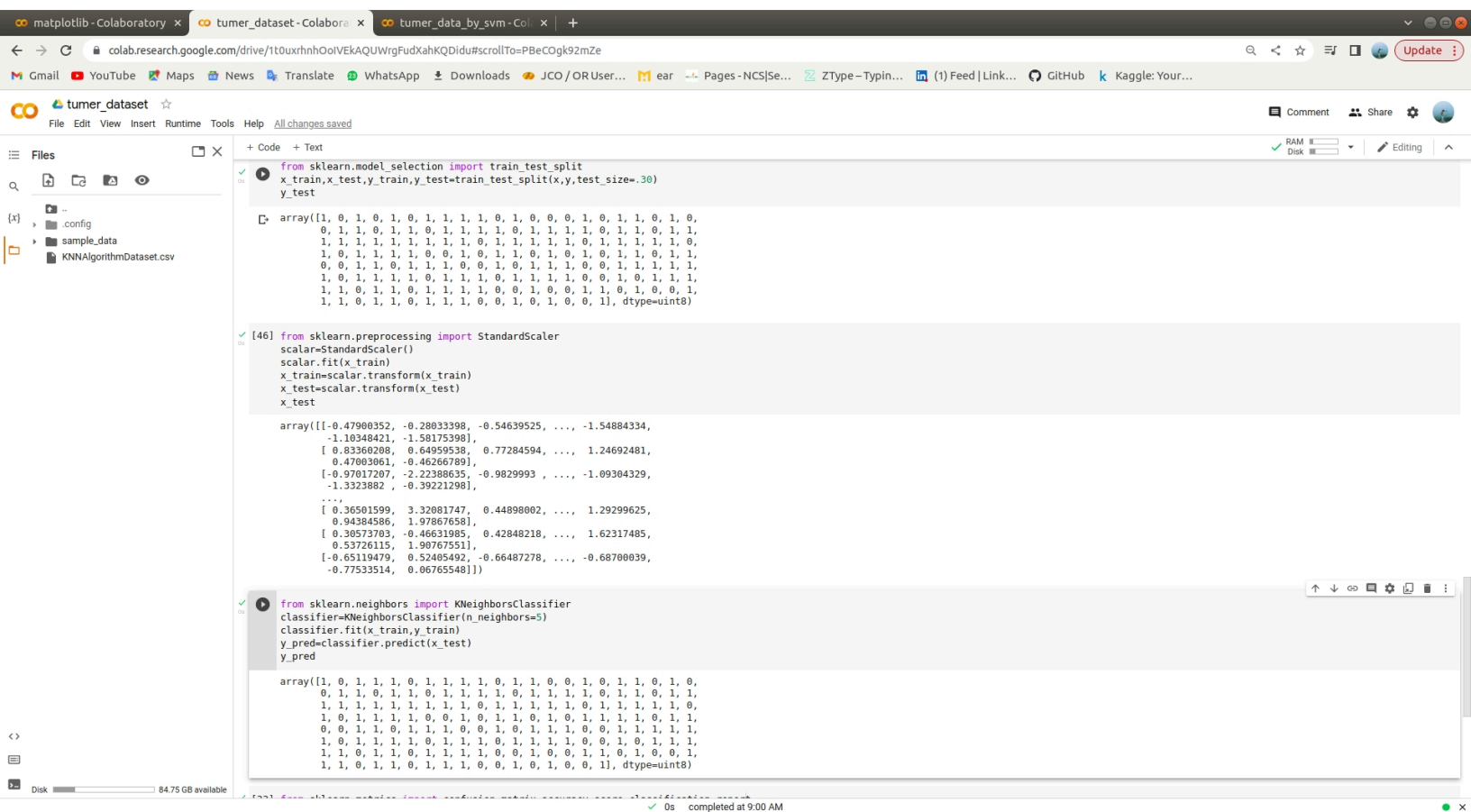
[24] del data\_upt['diagnosis\_M']

[25] del data\_upt['Unnamed: 32']

Disk84.76 GB available

0s completed at 8:58 AM





```
1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1,  
1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 1,  
1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, dtype=uint8)
```

```
from sklearn.metrics import confusion_matrix, accuracy_score, classification_report  
a=confusion_matrix(y_test, y_pred)  
score=accuracy_score(y_test, y_pred)  
score
```

0.9824561403508771