Nama : Dina Etikasari

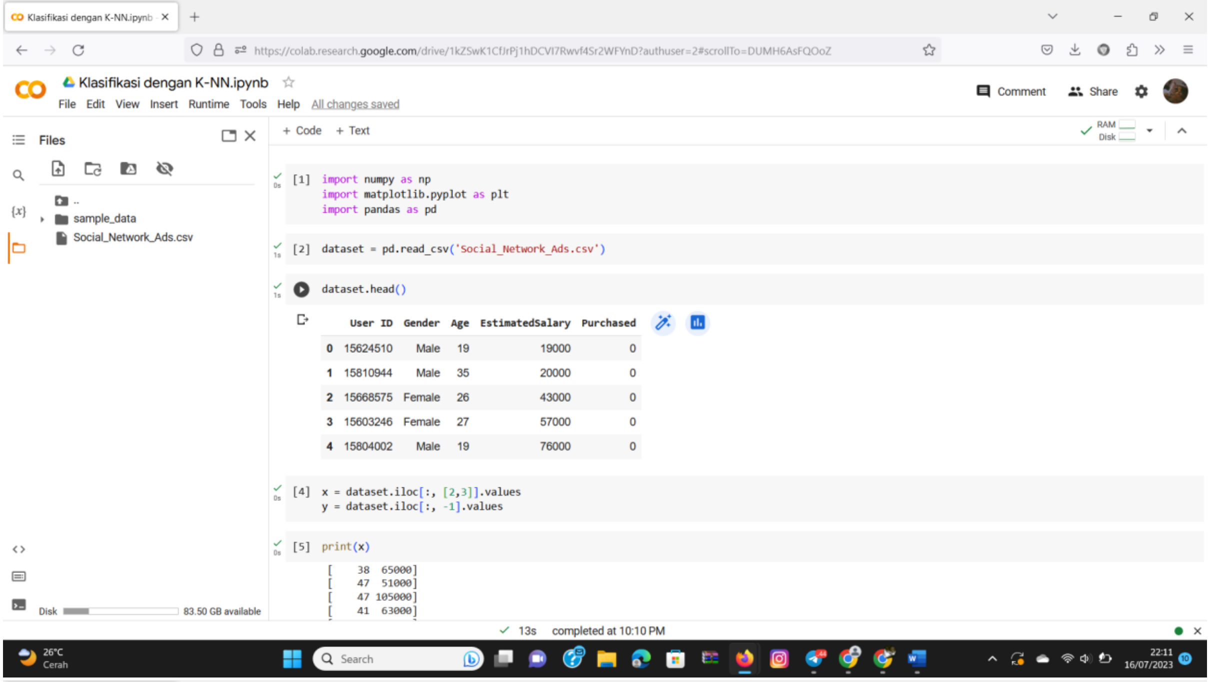
NIM : A11.2020.13191

Kelas : A11.4614

Link Colab :

https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ

Link Github : https://github.com/Dinaeks19/DATA_MINING/tree/main/LATIHAN%206


Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Files
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available
13s completed at 10:10 PM
26°C
Cerah
Search
22:11
16/07/2023

```
[1] import numpy as np
    import matplotlib.pyplot as plt
    import pandas as pd
```

```
[2] dataset = pd.read_csv('Social_Network_Ads.csv')
```

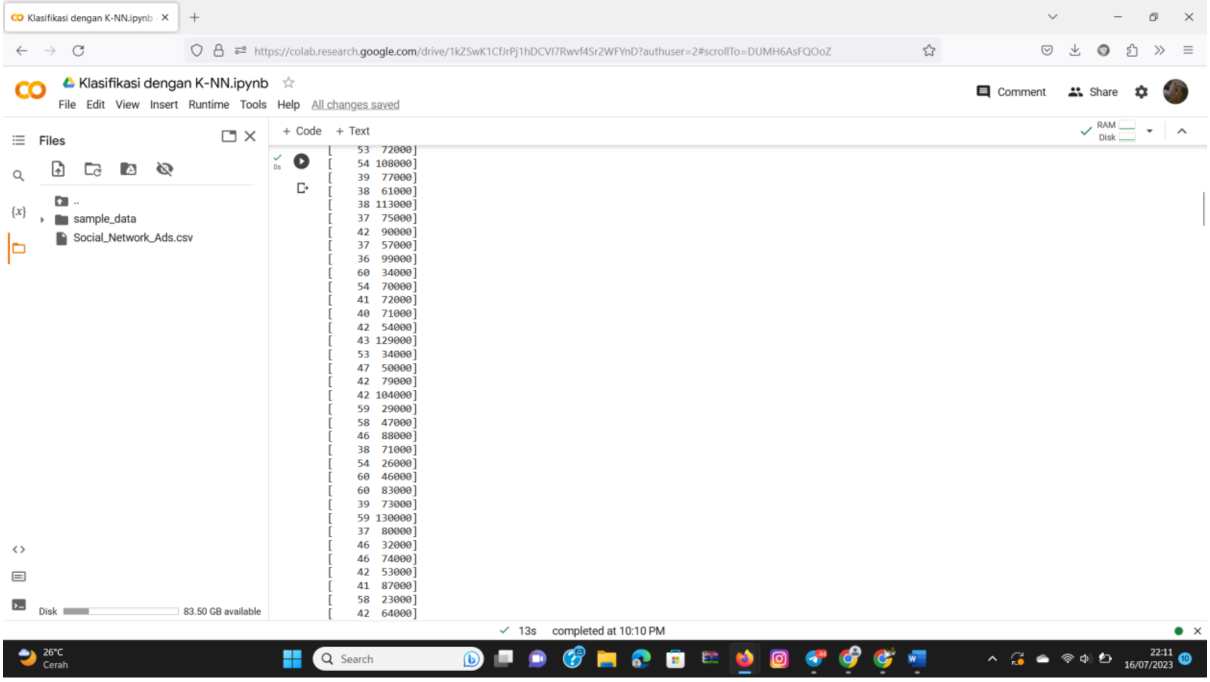
```
dataset.head()
```

	User ID	Gender	Age	EstimatedSalary	Purchased
0	15624510	Male	19	19000	0
1	15810944	Male	35	20000	0
2	15668575	Female	26	43000	0
3	15603246	Female	27	57000	0
4	15804002	Male	19	76000	0

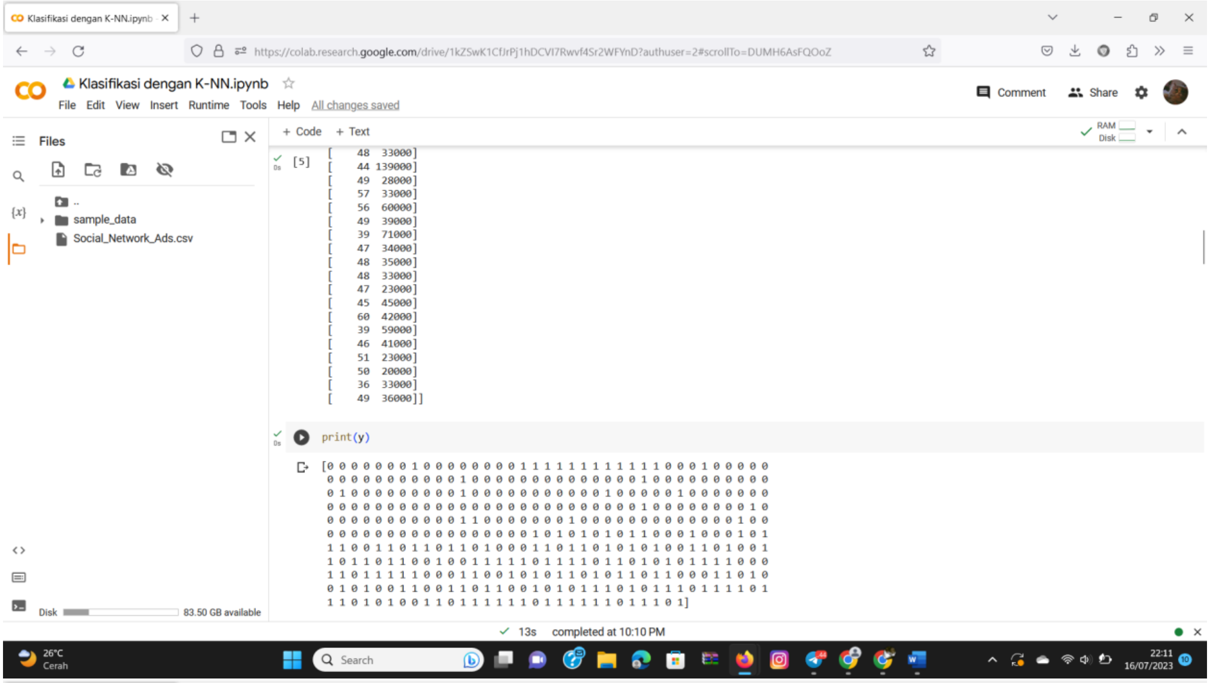
```
[4] x = dataset.iloc[:, [2,3]].values
    y = dataset.iloc[:, -1].values
```

```
[5] print(x)
```

```
[     38  65000]
[     47  51000]
[     47 105000]
[     41  63000]
```


Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Files
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available
13s completed at 10:10 PM
26°C
Cerah
Search
22:11
16/07/2023

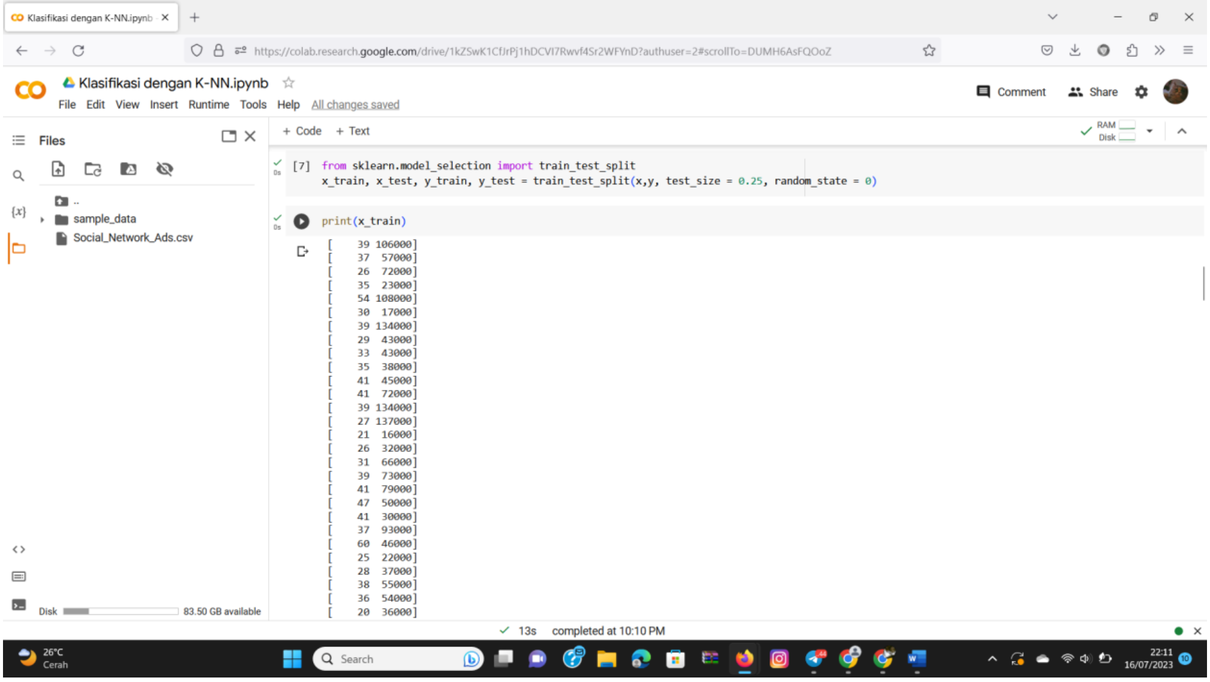
```
[     53  72000]
[     54 108000]
[     39  77000]
[     38  61000]
[     38 113000]
[     37  75000]
[     42  90000]
[     37  57000]
[     36  99000]
[     60  34000]
[     54  70000]
[     41  72000]
[     40  71000]
[     42  54000]
[     43 129000]
[     53  34000]
[     47  50000]
[     42  79000]
[     42 104000]
[     59  29000]
[     58  47000]
[     46  88000]
[     38  71000]
[     54  26000]
[     60  46000]
[     60  83000]
[     39  73000]
[     59 130000]
[     37  80000]
[     46  32000]
[     46  74000]
[     42  53000]
[     41  87000]
[     58  23000]
[     42  64000]
```


Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
Klasifikasi dengan K-NN.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
+ Code + Text
RAM Disk
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available
13s completed at 10:10 PM
26°C Cerah
Search
22:11 16/07/2023

```
[5]
 [    48  33000]
 [    44 139000]
 [    49  28000]
 [    57  33000]
 [    56  60000]
 [    49  39000]
 [    39  71000]
 [    47  34000]
 [    48  35000]
 [    48  33000]
 [    47  23000]
 [    45  45000]
 [    60  42000]
 [    39  59000]
 [    46  41000]
 [    51  23000]
 [    50  20000]
 [    36  33000]
 [    49  36000]]
```

```
print(y)
```

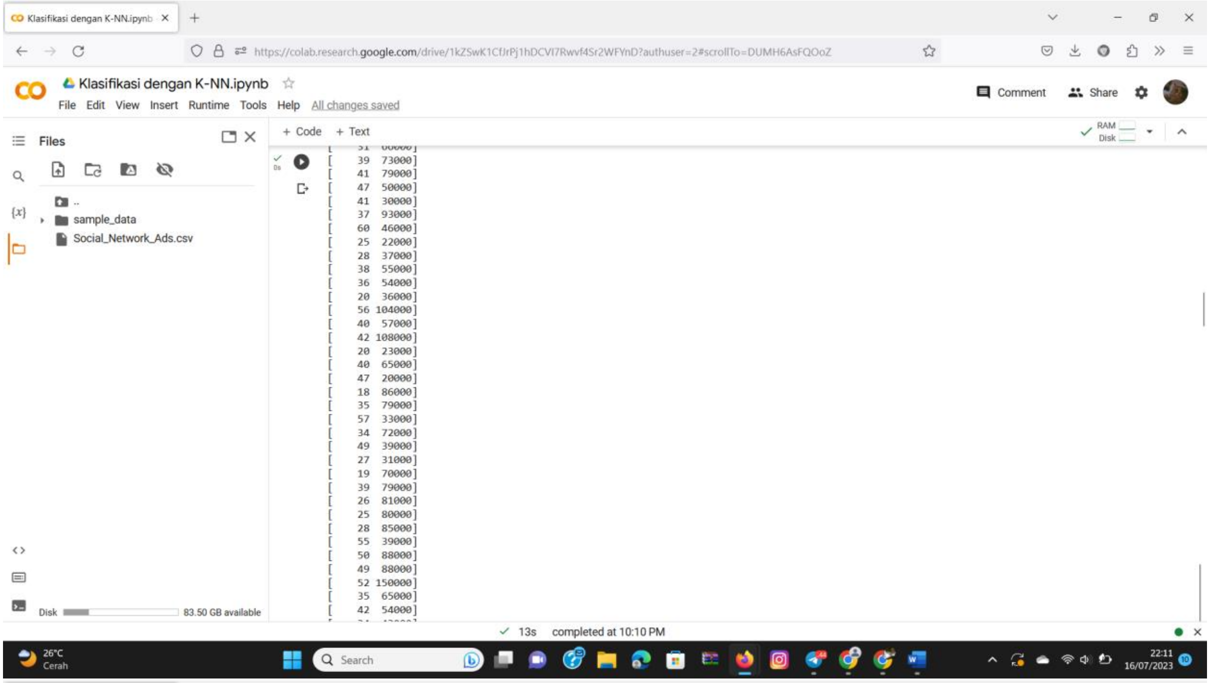
```
[0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 1 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0
 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0
 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 1 0 1 1 0 0 0 1 0 0 0 1 0 1
 1 1 0 0 1 1 0 1 1 0 1 1 0 1 0 0 0 1 1 0 1 1 0 1 0 1 0 1 0 0 1 1 0 1 0 0 1
 1 0 1 1 0 1 1 0 0 1 0 0 1 1 1 1 1 0 1 1 1 1 0 1 1 0 1 0 1 0 1 1 1 1 0 0 0
 1 1 0 1 1 1 1 1 0 0 0 1 1 0 0 1 0 1 0 1 1 0 1 0 1 1 0 1 1 0 0 0 1 1 0 1 0
 0 1 0 1 0 0 1 1 0 0 1 1 0 1 1 0 0 1 0 1 0 1 1 1 0 1 0 1 1 1 0 1 1 1 1 0 1
 1 1 0 1 0 1 0 0 1 1 0 1 1 1 1 1 1 0 1 1 1 1 1 1 0 1 1 1 0 1]
```


Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
Klasifikasi dengan K-NN.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
+ Code + Text
RAM Disk
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available
13s completed at 10:10 PM
26°C Cerah
Search
22:11 16/07/2023

```
[7] from sklearn.model_selection import train_test_split
    x_train, x_test, y_train, y_test = train_test_split(x,y, test_size = 0.25, random_state = 0)
```

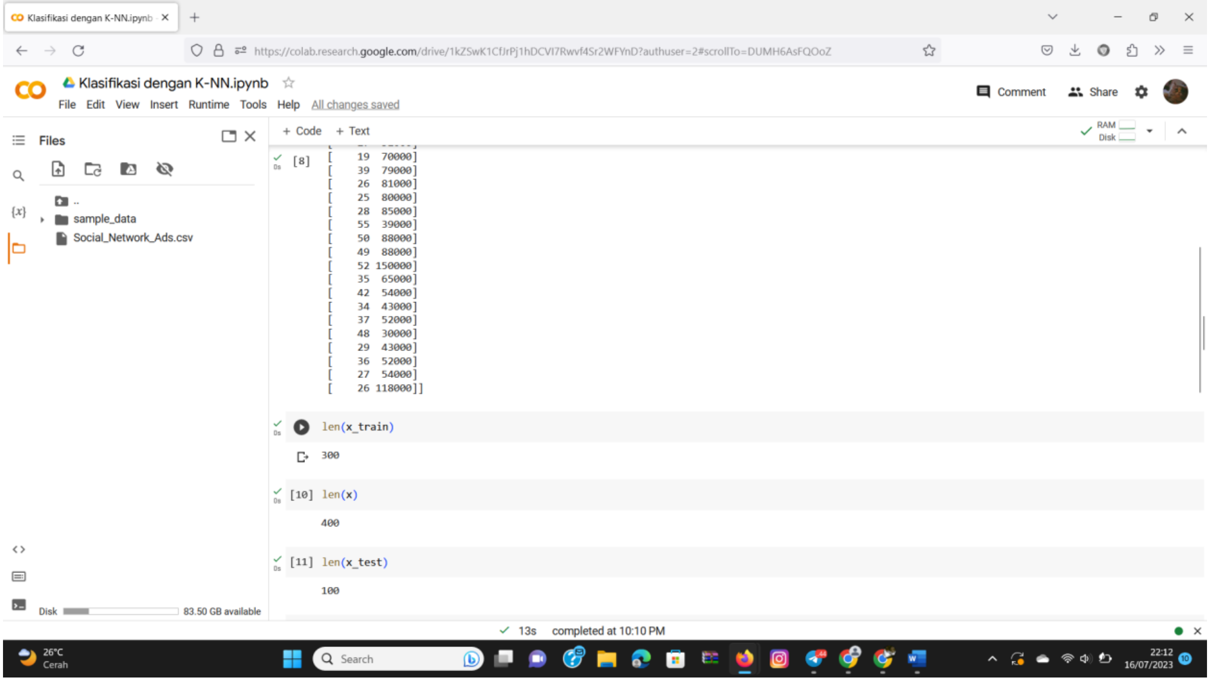
```
print(x_train)
```

```
[    39 106000]
[    37  57000]
[    26  72000]
[    35  23000]
[    54 108000]
[    30  17000]
[    39 134000]
[    29  43000]
[    33  43000]
[    35  38000]
[    41  45000]
[    41  72000]
[    39 134000]
[    27 137000]
[    21  16000]
[    26  32000]
[    31  66000]
[    39  73000]
[    41  79000]
[    47  50000]
[    41  30000]
[    37  93000]
[    60  46000]
[    25  22000]
[    28  37000]
[    38  55000]
[    36  54000]
[    20  36000]
```



Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
Klasifikasi dengan K-NN.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available
13s completed at 10:10 PM
26°C
Cerah
Search
22:11
16/07/2023

```
 [    39  73000]
 [    41  79000]
 [    47  50000]
 [    41  30000]
 [    37  93000]
 [    60  46000]
 [    25  22000]
 [    28  37000]
 [    38  55000]
 [    36  54000]
 [    20  36000]
 [    56 104000]
 [    40  57000]
 [    42 108000]
 [    20  23000]
 [    40  65000]
 [    47  20000]
 [    18  86000]
 [    35  79000]
 [    57  33000]
 [    34  72000]
 [    49  39000]
 [    27  31000]
 [    19  70000]
 [    39  79000]
 [    26  81000]
 [    25  80000]
 [    28  85000]
 [    55  39000]
 [    50  88000]
 [    49  88000]
 [    52 150000]
 [    35  65000]
 [    42  54000]
```



Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
Klasifikasi dengan K-NN.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available
13s completed at 10:10 PM
26°C
Cerah
Search
22:12
16/07/2023

```
[8]
 [    19  70000]
 [    39  79000]
 [    26  81000]
 [    25  80000]
 [    28  85000]
 [    55  39000]
 [    50  88000]
 [    49  88000]
 [    52 150000]
 [    35  65000]
 [    42  54000]
 [    34  43000]
 [    37  52000]
 [    48  30000]
 [    29  43000]
 [    36  52000]
 [    27  54000]
 [    26 118000]]
```

```
len(x_train)
```

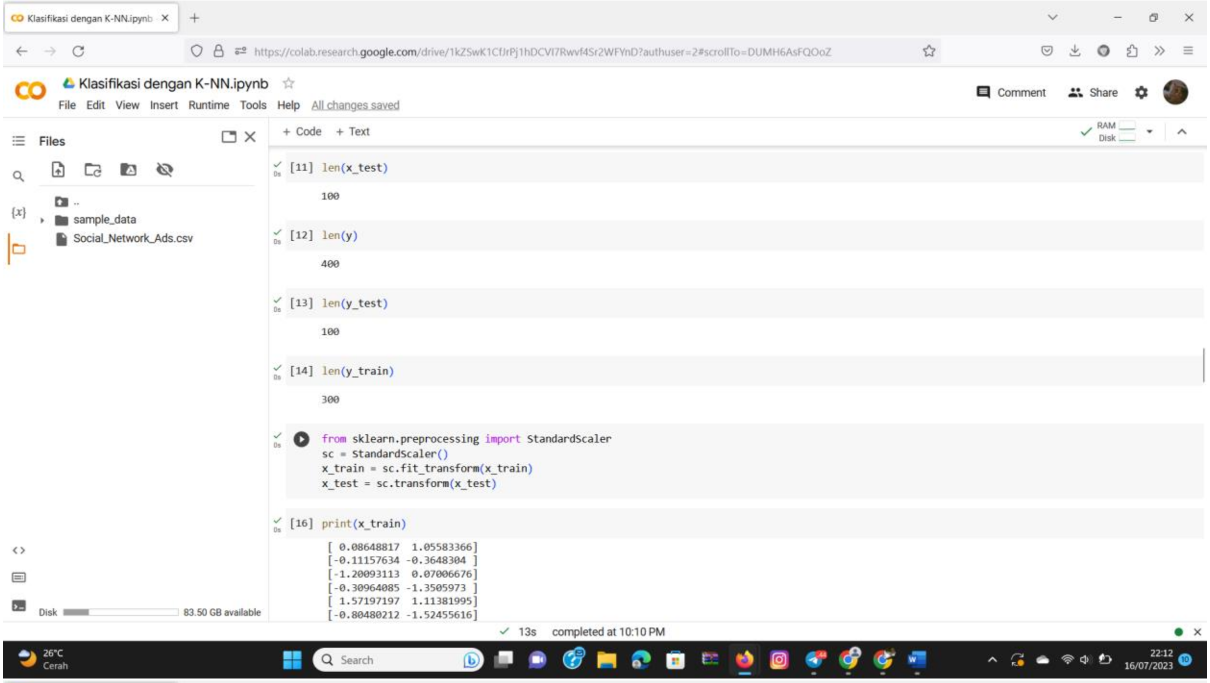
```
300
```

```
[10] len(x)
```

```
400
```

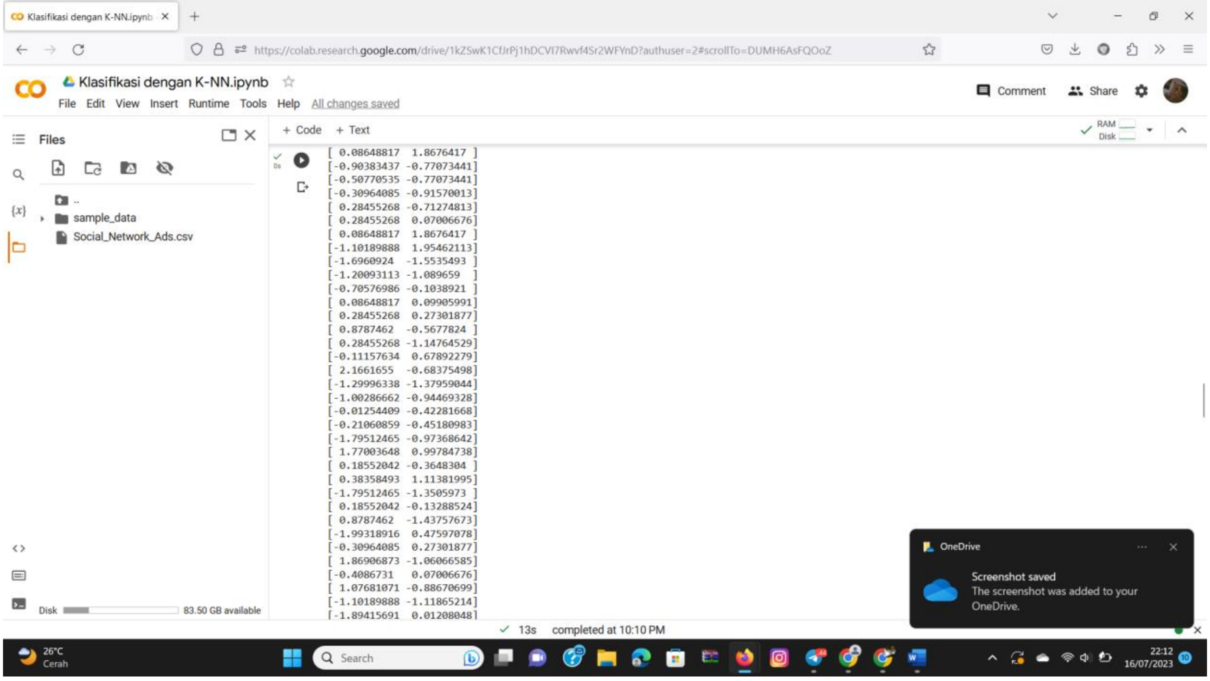
```
[11] len(x_test)
```

```
100
```


Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available
+ Code + Text
RAM
Disk
26°C
Cerah
Search
22:12
16/07/2023

```
[11] len(x_test)

     100

[12] len(y)

     400

[13] len(y_test)

     100

[14] len(y_train)

     300

     from sklearn.preprocessing import StandardScaler
     sc = StandardScaler()
     x_train = sc.fit_transform(x_train)
     x_test = sc.transform(x_test)

[16] print(x_train)

      [ 0.08648817  1.05583366]
      [-0.11157634 -0.3648304 ]
      [-1.20093113  0.07006676]
      [-0.30964085 -1.3505973 ]
      [ 1.57197197  1.11381995]
      [-0.80480212 -1.52455616]
```

✓ 13s completed at 10:10 PM

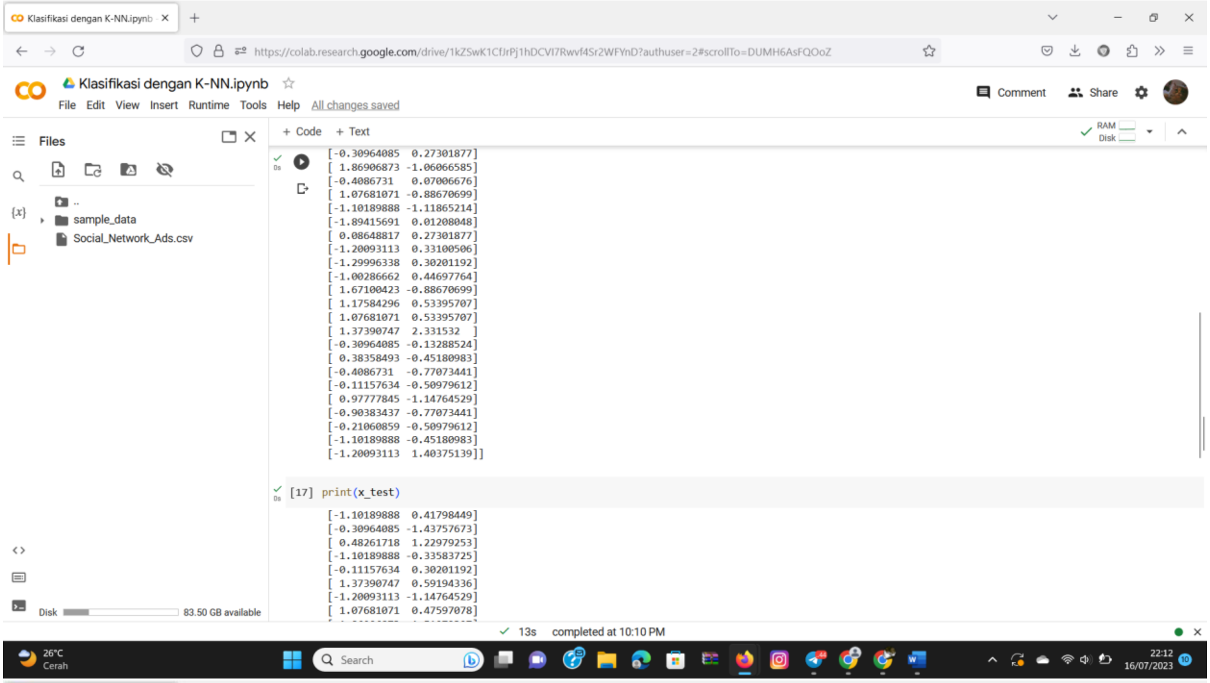

Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available
+ Code + Text
RAM
Disk
26°C
Cerah
Search
22:12
16/07/2023

```
 [ 0.08648817  1.8676417 ]
 [-0.90383437 -0.77073441]
 [-0.50770535 -0.77073441]
 [-0.30964085 -0.91570013]
 [ 0.28455268 -0.71274813]
 [ 0.28455268  0.07006676]
 [ 0.08648817  1.8676417 ]
 [-1.10189888  1.95462113]
 [-1.6960924  -1.5535493 ]
 [-1.20093113 -1.089659  ]
 [-0.70576986 -0.1038921 ]
 [ 0.08648817  0.09905991]
 [ 0.28455268  0.27301877]
 [ 0.8787462  -0.5677824 ]
 [ 0.28455268 -1.14764529]
 [-0.11157634  0.67892279]
 [ 2.1661655  -0.68375498]
 [-1.29996338 -1.37959044]
 [-1.00286662 -0.94469328]
 [-0.01254409 -0.42281668]
 [-0.21060859 -0.45180983]
 [-1.79512465 -0.97368642]
 [ 1.77003648  0.99784738]
 [ 0.18552042 -0.3648304 ]
 [ 0.38358493  1.11381995]
 [-1.79512465 -1.3505973 ]
 [ 0.18552042 -0.13288524]
 [ 0.8787462  -1.43757673]
 [-1.99318916  0.47597078]
 [-0.30964085  0.27301877]
 [ 1.86906873 -1.06066585]
 [-0.4086731   0.07006676]
 [ 1.07681071 -0.88670699]
 [-1.10189888 -1.11865214]
 [-1.89415691  0.01208048]
```

OneDrive

Screenshot saved
The screenshot was added to your OneDrive.

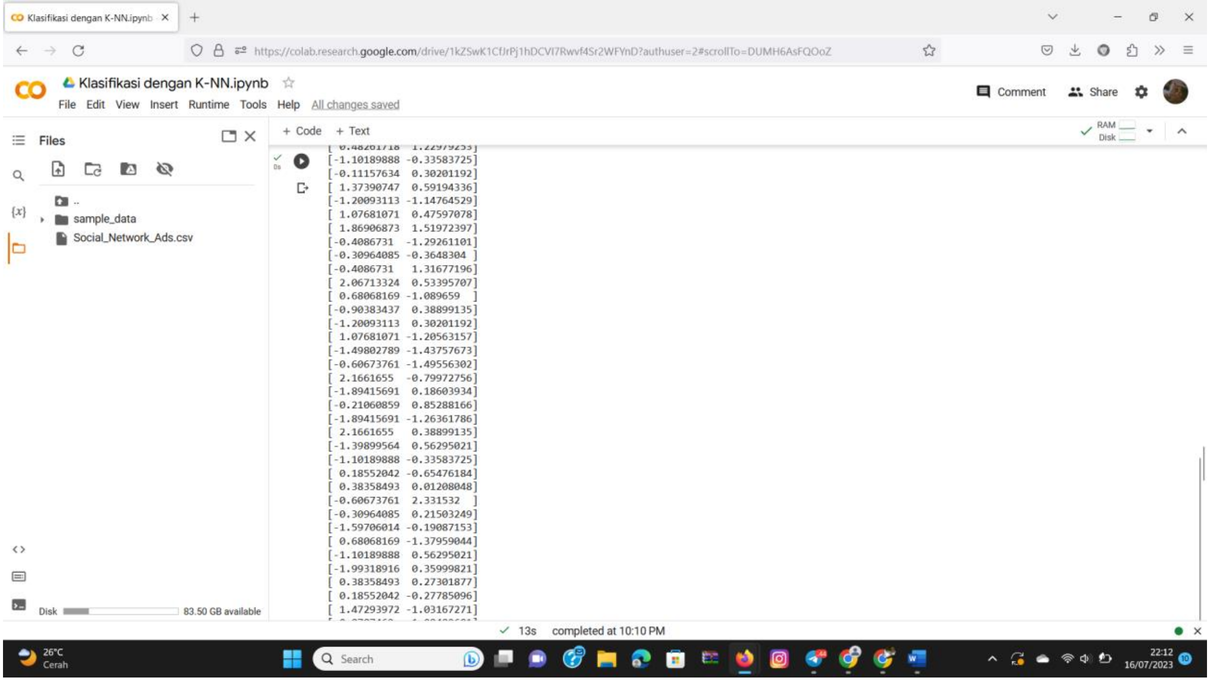
✓ 13s completed at 10:10 PM


Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
+ Code + Text
RAM Disk
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available
13s completed at 10:10 PM
26°C Cerah
Search
22:12 16/07/2023

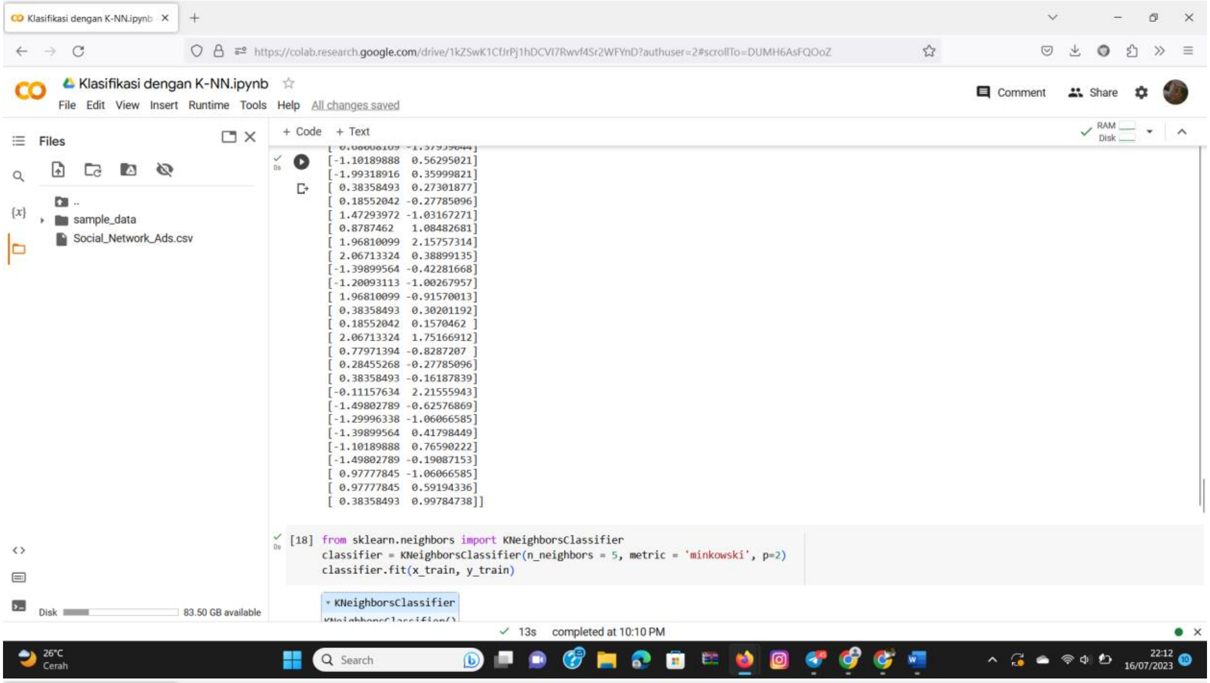
```
[-0.30964085  0.27301877]
[ 1.86906873 -1.06066585]
[-0.4086731   0.07006676]
[ 1.07681071 -0.88670699]
[-1.10189888 -1.11865214]
[-1.89415691  0.01208048]
[ 0.08648817  0.27301877]
[-1.20093113  0.33100506]
[-1.29996338  0.30201192]
[-1.00286662  0.44697764]
[ 1.67100423 -0.88670699]
[ 1.17584296  0.53395707]
[ 1.07681071  0.53395707]
[ 1.37390747  2.331532  ]
[-0.30964085 -0.13288524]
[ 0.38358493 -0.45180983]
[-0.4086731  -0.77073441]
[-0.11157634 -0.50979612]
[ 0.97777845 -1.14764529]
[-0.90383437 -0.77073441]
[-0.21060859 -0.50979612]
[-1.10189888 -0.45180983]
[-1.20093113  1.40375139]]
```

```
[17] print(x_test)
```

```
[-1.10189888  0.41798449]
[-0.30964085 -1.43757673]
[ 0.48261718  1.22979253]
[-1.10189888 -0.33583725]
[-0.11157634  0.30201192]
[ 1.37390747  0.59194336]
[-1.20093113 -1.14764529]
[ 1.07681071  0.47597078]
```


Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
+ Code + Text
RAM Disk
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available
13s completed at 10:10 PM
26°C Cerah
Search
22:12 16/07/2023

```
[ 0.48261718  1.22979253]
[-1.10189888 -0.33583725]
[-0.11157634  0.30201192]
[ 1.37390747  0.59194336]
[-1.20093113 -1.14764529]
[ 1.07681071  0.47597078]
[ 1.86906873  1.51972397]
[-0.4086731  -1.29261101]
[-0.30964085 -0.3648304 ]
[-0.4086731   1.31677196]
[ 2.06713324  0.53395707]
[ 0.68068169 -1.089659  ]
[-0.90383437  0.38899135]
[-1.20093113  0.30201192]
[ 1.07681071 -1.20563157]
[-1.49802789 -1.43757673]
[-0.60673761 -1.49556302]
[ 2.1661655  -0.79972756]
[-1.89415691  0.18603934]
[-0.21060859  0.85288166]
[-1.89415691 -1.26361786]
[ 2.1661655   0.38899135]
[-1.39899564  0.56295021]
[-1.10189888 -0.33583725]
[ 0.18552042 -0.65476184]
[ 0.38358493  0.01208048]
[-0.60673761  2.331532  ]
[-0.30964085  0.21503249]
[-1.59706014 -0.19087153]
[ 0.68068169 -1.37959044]
[-1.10189888  0.56295021]
[-1.99318916  0.35999821]
[ 0.38358493  0.27301877]
[ 0.18552042 -0.27785096]
[ 1.47293972 -1.03167271]
```

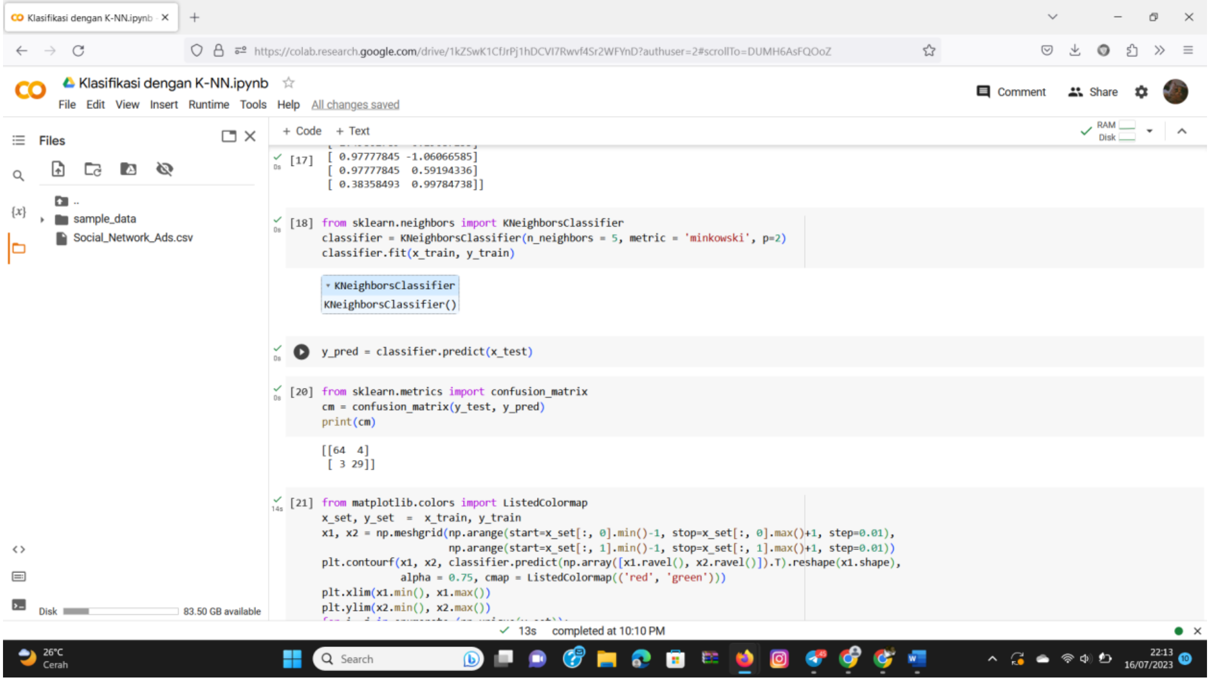

Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
Klasifikasi dengan K-NN.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
+ Code + Text
RAM Disk
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available

```
 [-1.10189888  0.56295021]
 [-1.99318916  0.35999821]
 [ 0.38358493  0.27301877]
 [ 0.18552042 -0.27785096]
 [ 1.47293972 -1.03167271]
 [ 0.8787462   1.08482681]
 [ 1.96810099  2.15757314]
 [ 2.06713324  0.38899135]
 [-1.39899564 -0.42281668]
 [-1.20093113 -1.00267957]
 [ 1.96810099 -0.91570013]
 [ 0.38358493  0.30201192]
 [ 0.18552042  0.1570462 ]
 [ 2.06713324  1.75166912]
 [ 0.77971394 -0.8287207 ]
 [ 0.28455268 -0.27785096]
 [ 0.38358493 -0.16187839]
 [-0.11157634  2.21555943]
 [-1.49802789 -0.62576869]
 [-1.29996338 -1.06066585]
 [-1.39899564  0.41798449]
 [-1.10189888  0.76590222]
 [-1.49802789 -0.19087153]
 [ 0.97777845 -1.06066585]
 [ 0.97777845  0.59194336]
 [ 0.38358493  0.99784738]]
```

```
[18] from sklearn.neighbors import KNeighborsClassifier
     classifier = KNeighborsClassifier(n_neighbors = 5, metric = 'minkowski', p=2)
     classifier.fit(x_train, y_train)
```

```
▾ KNeighborsClassifier
```

✓ 13s completed at 10:10 PM
26°C
Cerah
Search
22:12
16/07/2023


Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
Klasifikasi dengan K-NN.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
+ Code + Text
RAM Disk
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available

```
[17] [ 0.97777845 -1.06066585]
     [ 0.97777845  0.59194336]
     [ 0.38358493  0.99784738]]
```

```
[18] from sklearn.neighbors import KNeighborsClassifier
     classifier = KNeighborsClassifier(n_neighbors = 5, metric = 'minkowski', p=2)
     classifier.fit(x_train, y_train)
```

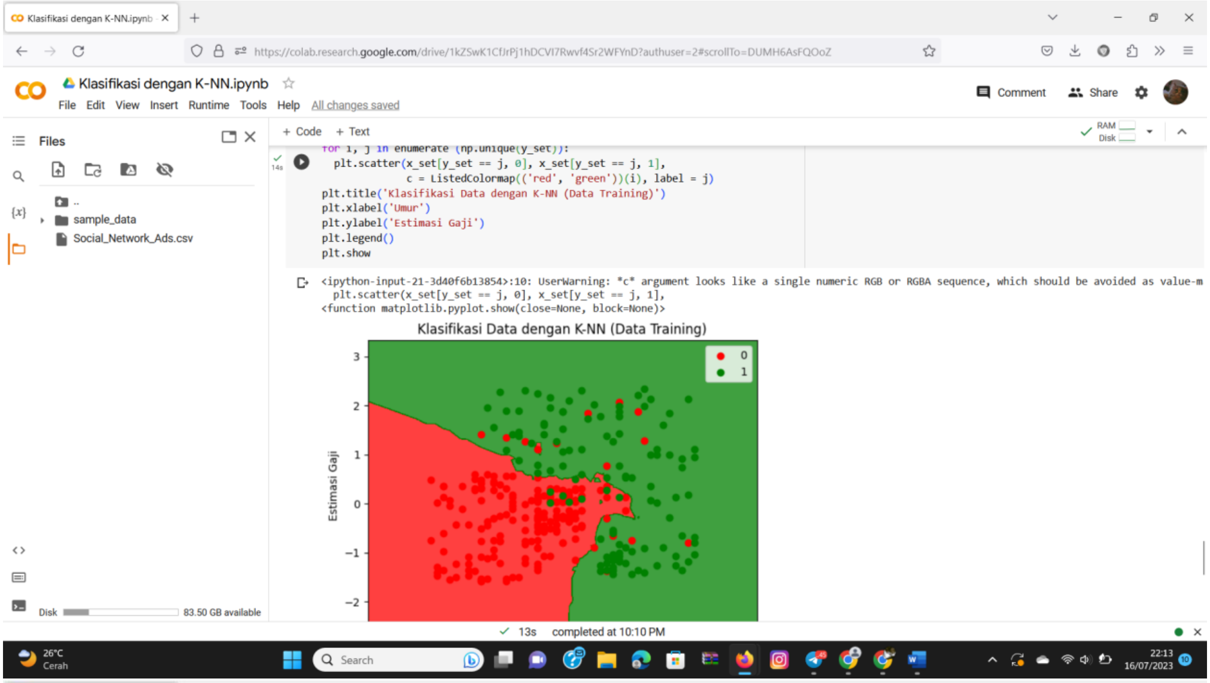
```
▾ KNeighborsClassifier
KNeighborsClassifier()
```

```
y_pred = classifier.predict(x_test)
```

```
[20] from sklearn.metrics import confusion_matrix
     cm = confusion_matrix(y_test, y_pred)
     print(cm)
```

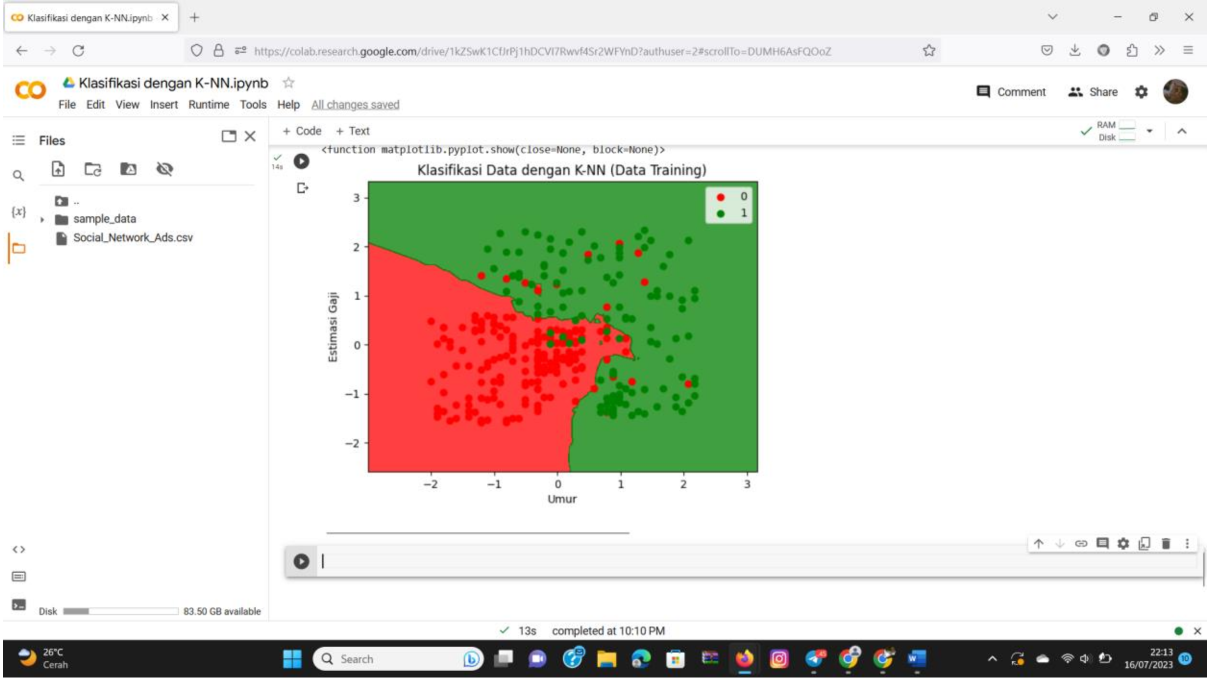
```
[[64  4]
 [ 3 29]]
```

```
[21] from matplotlib.colors import ListedColormap
     x_set, y_set  =  x_train, y_train
     x1, x2 = np.meshgrid(np.arange(start=x_set[:, 0].min()-1, stop=x_set[:, 0].max()+1, step=0.01),
                          np.arange(start=x_set[:, 1].min()-1, stop=x_set[:, 1].max()+1, step=0.01))
     plt.contourf(x1, x2, classifier.predict(np.array([x1.ravel(), x2.ravel()]).T).reshape(x1.shape),
                  alpha = 0.75, cmap = ListedColormap(('red', 'green')))
     plt.xlim(x1.min(), x1.max())
     plt.ylim(x2.min(), x2.max())
```


Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
Klasifikasi dengan K-NN.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available
+ Code + Text
RAM Disk
Klasifikasi Data dengan K-NN (Data Training)
Estimasi Gaji
3
2
1
0
−1
−2
0
1
✓ 13s completed at 10:10 PM
26°C Cerah
Search
22:13 16/07/2023

```
for i, j in enumerate (np.unique(y_set)):
  plt.scatter(x_set[y_set == j, 0], x_set[y_set == j, 1],
              c = ListedColormap(('red', 'green'))(i), label = j)
plt.title('Klasifikasi Data dengan K-NN (Data Training)')
plt.xlabel('Umur')
plt.ylabel('Estimasi Gaji')
plt.legend()
plt.show
```

```
<ipython-input-21-3d40f6b13854>:10: UserWarning: *c* argument looks like a single numeric RGB or RGBA sequence, which should be avoided as value-m
  plt.scatter(x_set[y_set == j, 0], x_set[y_set == j, 1],
<function matplotlib.pyplot.show(close=None, block=None)>
```


Klasifikasi dengan K-NN.ipynb
https://colab.research.google.com/drive/1kZSwK1CfJrPj1hDCVI7Rwvf4Sr2WFYnD?authuser=2#scrollTo=DUMH6AsFQOoZ
Klasifikasi dengan K-NN.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
sample_data
Social_Network_Ads.csv
Disk 83.50 GB available
+ Code + Text
RAM Disk
Klasifikasi Data dengan K-NN (Data Training)
Estimasi Gaji
3
2
1
0
−1
−2
0
1
−2
−1
0
1
2
3
Umur
✓ 13s completed at 10:10 PM
26°C Cerah
Search
22:13 16/07/2023

```
<function matplotlib.pyplot.show(close=None, block=None)>
```