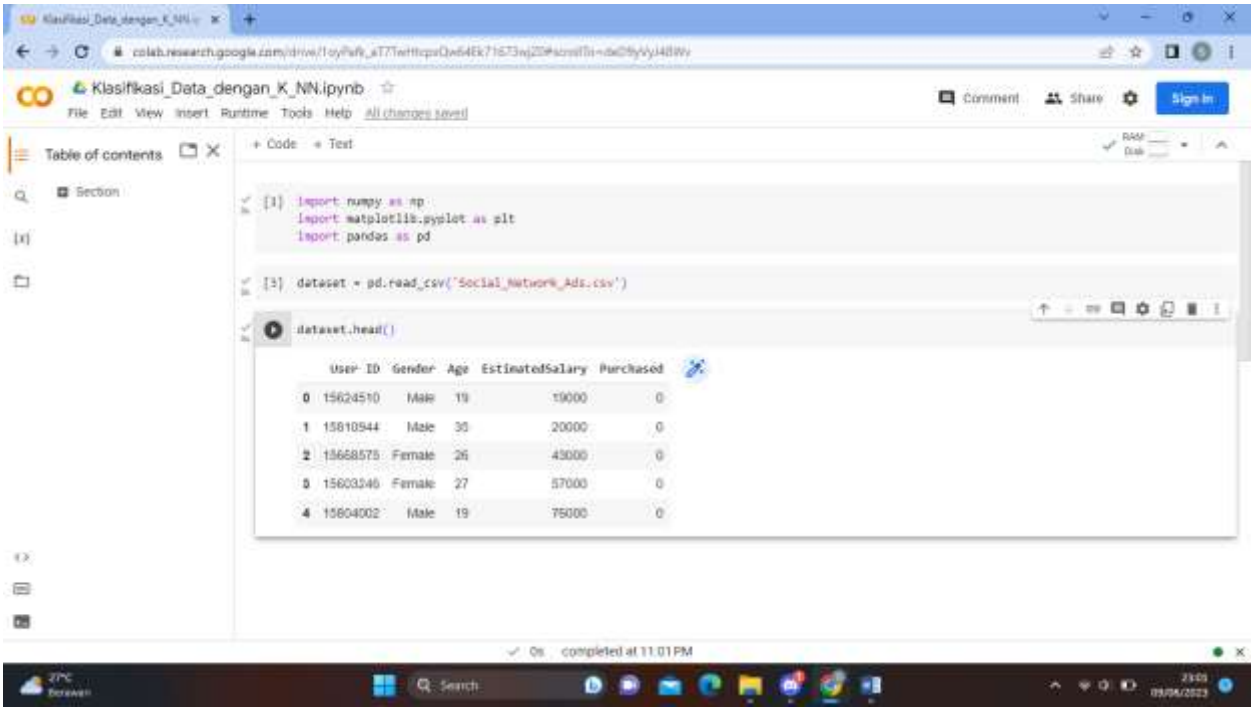
Nama : Shine Vardina Victoria Sudrajat

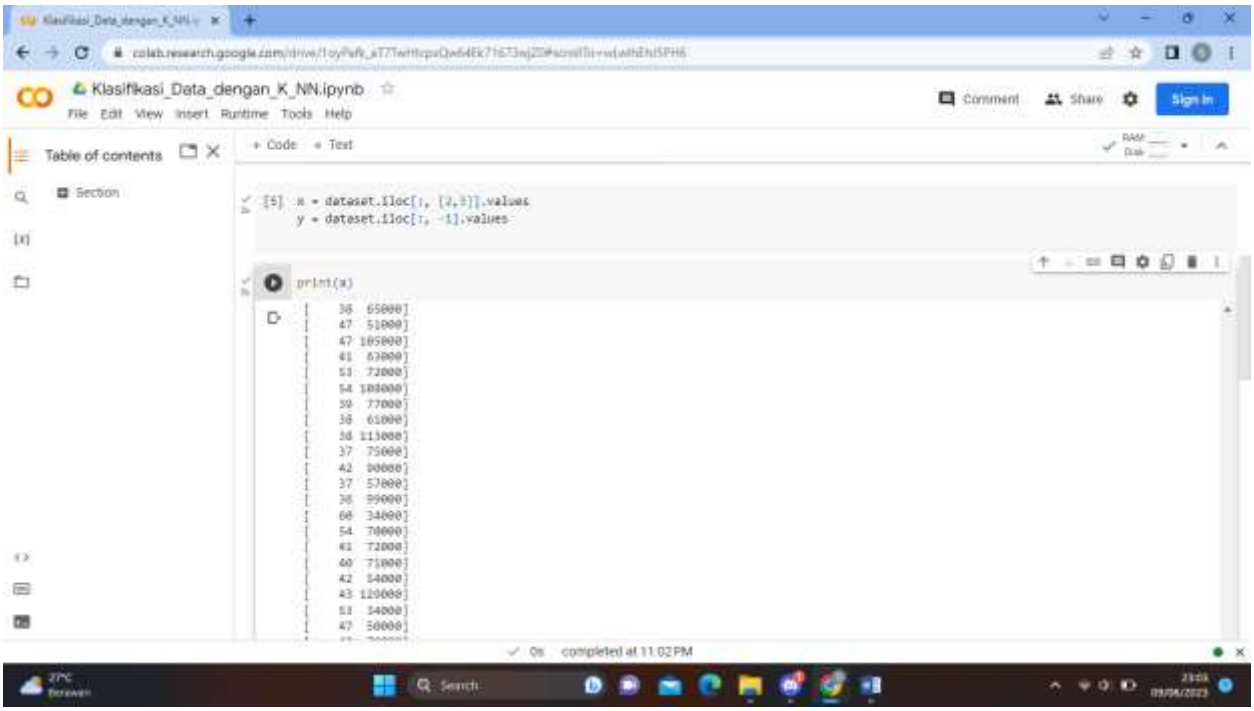
NIM : A11.2020.13189

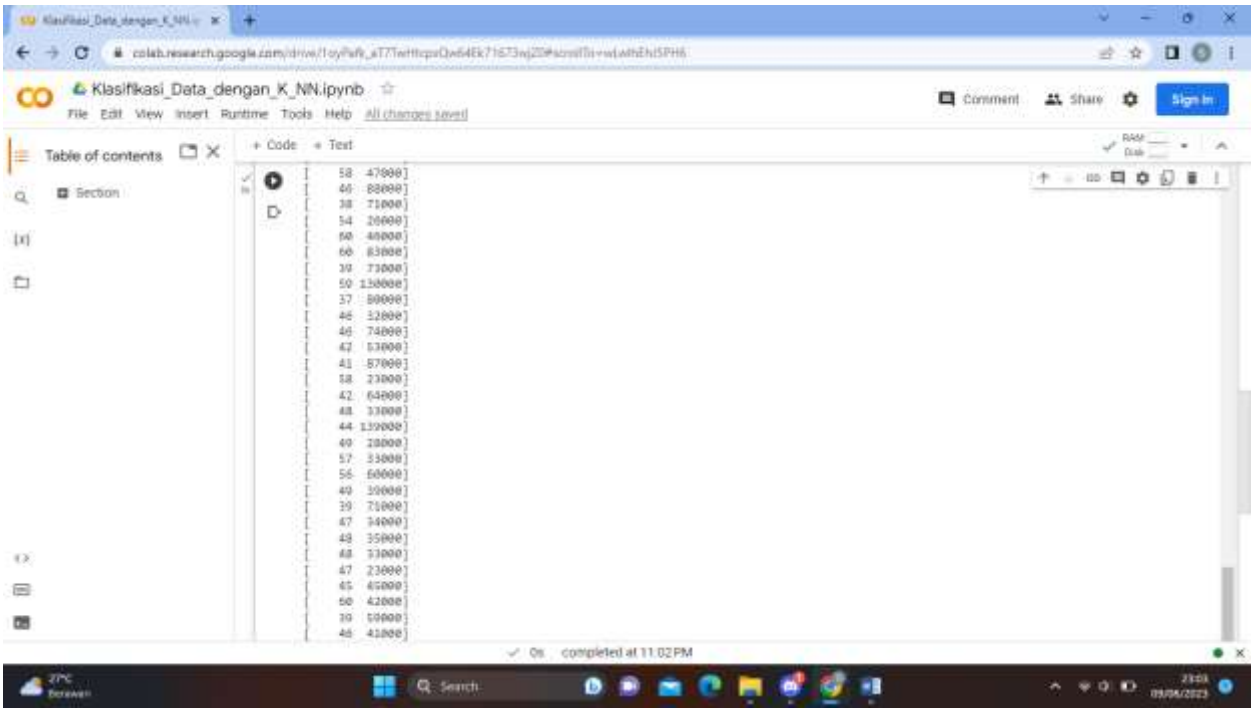
Kelas : A11.4614

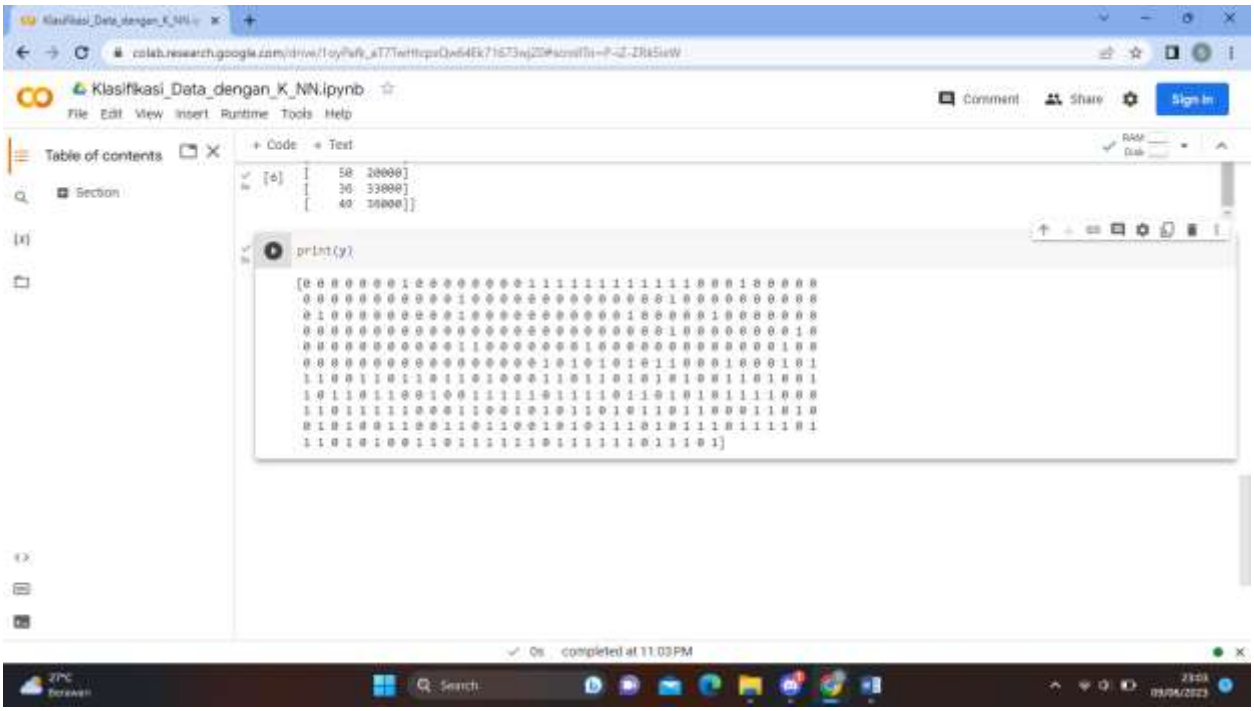
Link Colab : https://colab.research.google.com/drive/1oyPafk_aT7TwHtcpxQw64Ek71673wjZ0?usp=sharing

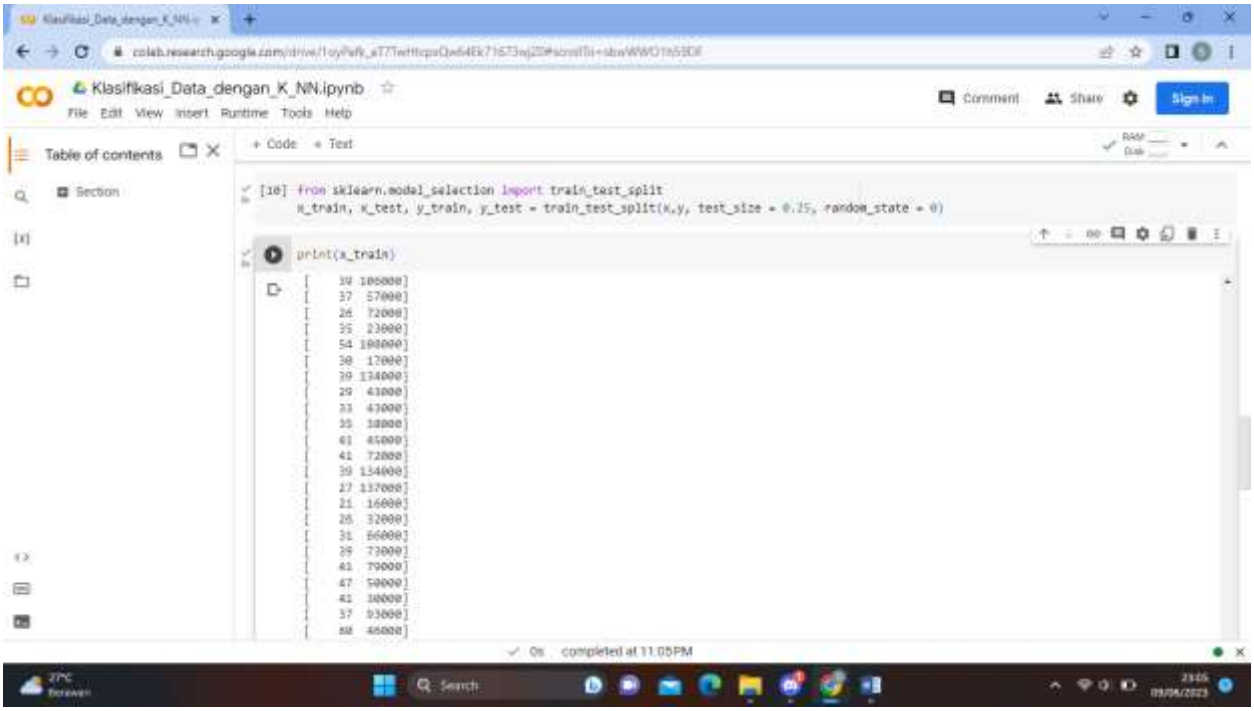
Link Github : https://github.com/Shinevv27/DataMining/tree/main/Latihan6

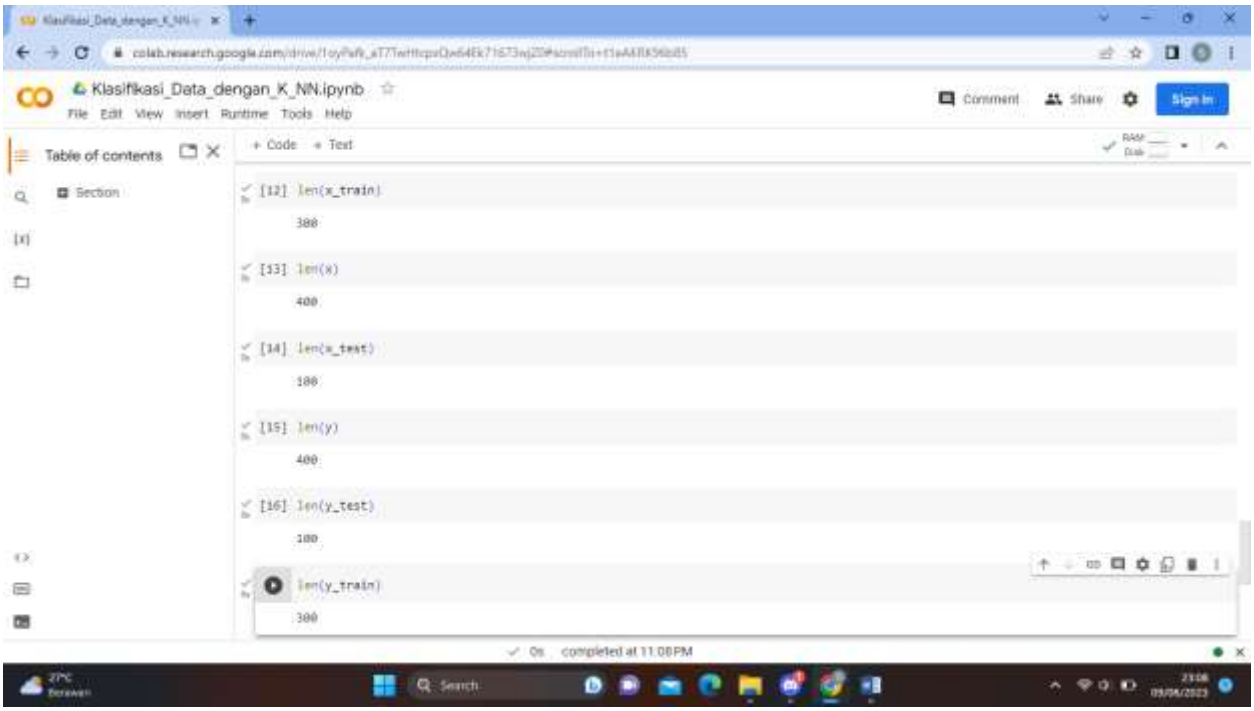


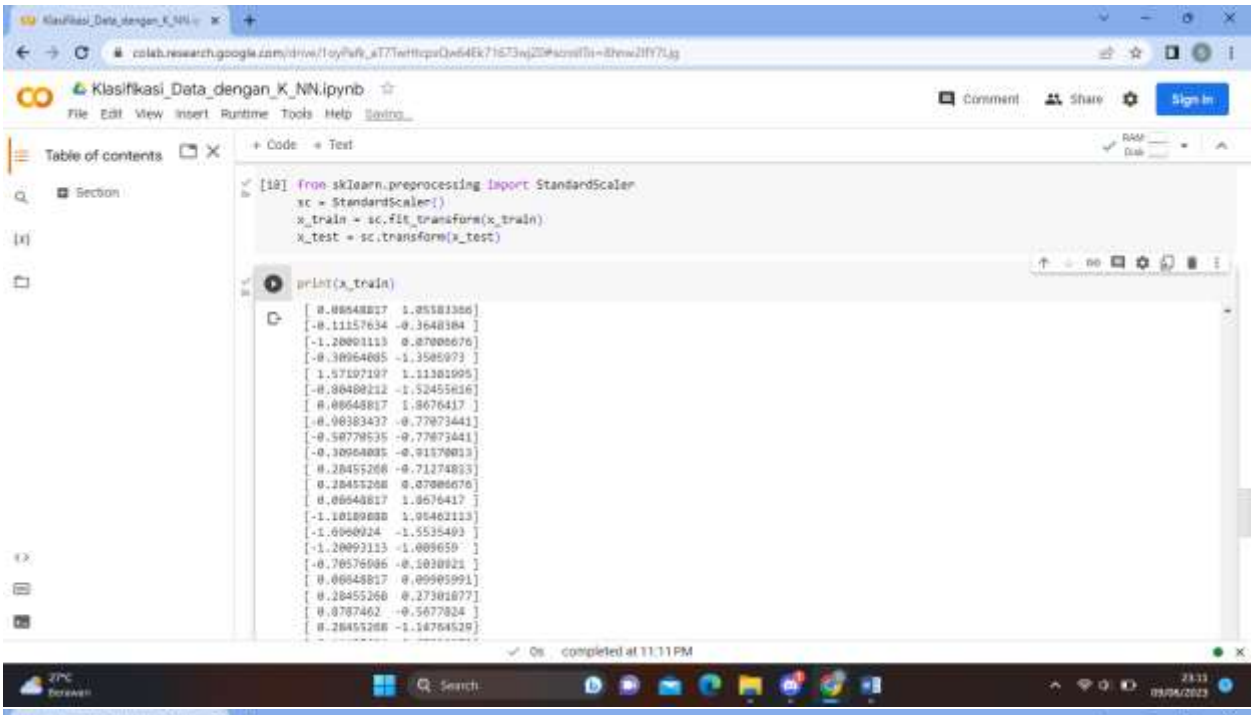

Klasifikasi_Data_dengan_K_NN.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
Sign in
Table of contents
Section
+ Code
+ Text
[5] x = dataset.iloc[:, [2,3]].values
y = dataset.iloc[:, -1].values
print(x)
[38 65000]
[47 51000]
[47 105000]
[41 63000]
[53 72000]
[54 108000]
[39 77000]
[38 61000]
[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]
0s completed at 11:02 PM
27°C
Search

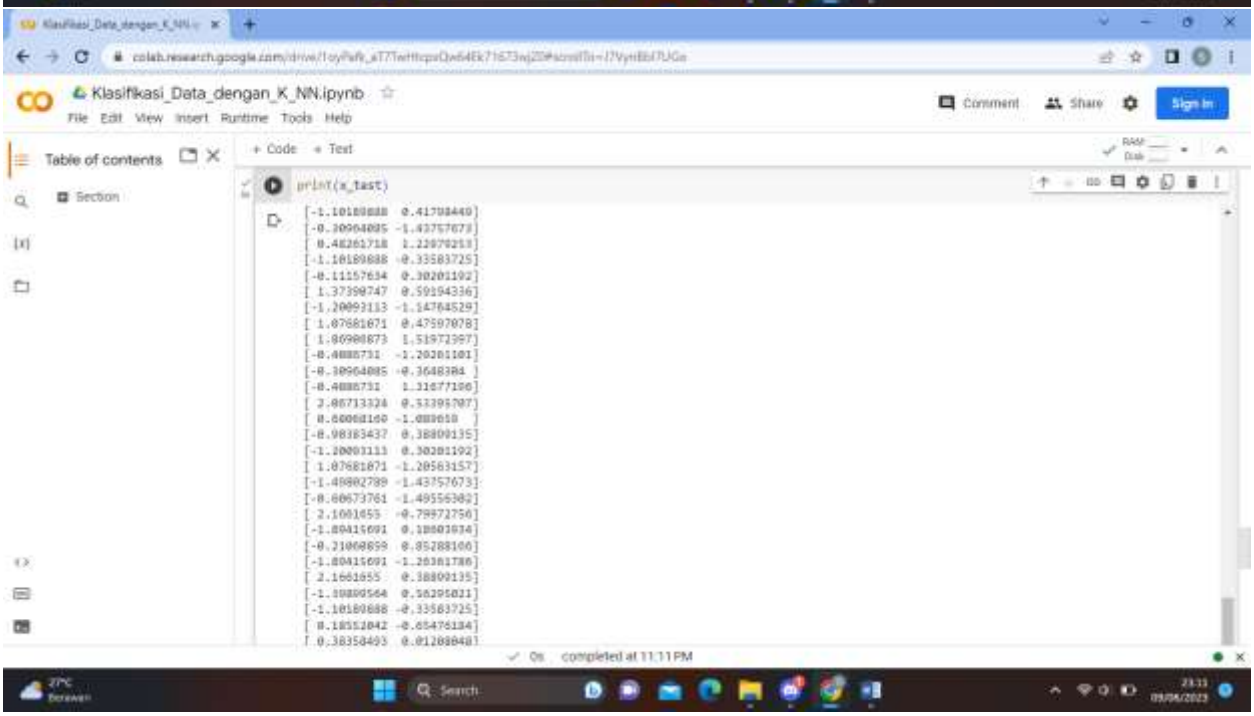

Klasifikasi_Data_dengan_K_NN.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Sign in
Table of contents
Section
+ Code
+ Text
[58 47000]
[46 88000]
[38 71000]
[54 26000]
[60 46000]
[60 83000]
[39 73000]
[59 130000]
[37 80000]
[46 32000]
[46 74000]
[42 53000]
[41 87000]
[58 23000]
[42 64000]
[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]
0s completed at 11:02 PM
27°C
Search

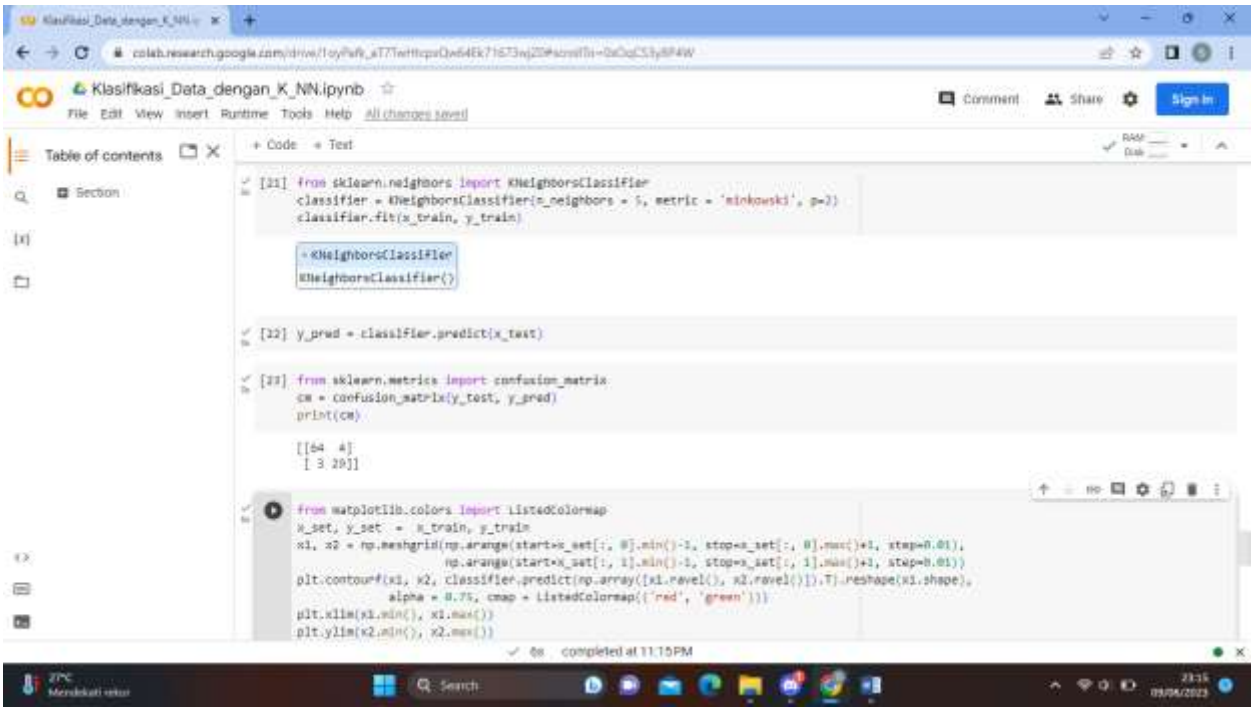

Klasifikasi_Data_dengan_K_NN.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
Sign in
Table of contents
Section
+ Code
+ Text
[6] [58 20000]
[36 33000]
[49 36000]]
print(y)
[0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 0 0 0 1 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 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 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 0 1 0 0 0 0 0 0 0 0 1 0
0 0 0 0 0 0 0 0 0 0 0 0 1 1 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 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 1 0 0 0
1 1 0 1 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 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 1 0 1 1 1 0 1]
completed at 11:03 PM
Search


Klasifikasi_Data_dengan_K_NN.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
Sign in
Table of contents
Section
+ Code
+ Text
[10] 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]
[61 65000]
[41 72000]
[39 134000]
[27 137000]
[21 16000]
[26 32000]
[31 66000]
[39 73000]
[41 79000]
[47 50000]
[41 30000]
[37 93000]
[60 46000]
completed at 11:05 PM
Search


Klasifikasi_Data_dengan_K_NN.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
Sign in
Table of contents
Section
+ Code + Text
[12] len(x_train)
300
[13] len(x)
400
[14] len(x_test)
100
[15] len(y)
400
[16] len(y_test)
100
len(y_train)
300
0s completed at 11:08PM
Search


Klasifikasi_Data_dengan_K_NN.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
Sign in
Table of contents
Section
+ Code + Text
[10] from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
x_train = sc.fit_transform(x_train)
x_test = sc.transform(x_test)
print(x_train)
[0.08648817 1.05583366]
[-0.11157634 -0.3648304]
[-1.20093113 0.07008676]
[-0.30964085 -1.3505973]
[1.57197197 1.11301995]
[-0.80480212 -1.52455616]
[0.08648817 1.0676417]
[-0.98383437 -0.77073441]
[-0.50778535 -0.77073441]
[-0.30964085 -0.91570013]
[0.28455268 -0.71174033]
[0.28455268 0.07008676]
[0.08648817 1.0676417]
[-1.10189888 1.95462113]
[-1.6960924 -1.5535493]
[-1.20093113 -1.009659]
[-0.70576986 -0.1030821]
[0.08648817 0.09905991]
[0.28455268 0.27301877]
[0.0787462 -0.5677824]
[0.28455268 -1.14764529]
0s completed at 11:11 PM
27°C
Search
23:11
09/06/2023


Klasifikasi_Data_dengan_K_NN.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
Sign in
Table of contents
Section
+ Code + Text
print(x_test)
[-1.10189888 0.41703449]
[-0.30964085 -1.43757673]
[0.48261718 1.22070213]
[-1.10189888 -0.33583725]
[-0.11157634 0.30201192]
[1.37390747 0.59194336]
[-1.20093113 -1.14764529]
[1.07681071 0.47597078]
[1.86908873 1.51972397]
[-0.4886731 -1.29281181]
[-0.30964085 -0.3648304]
[-0.4886731 1.31677196]
[2.06713324 0.53395707]
[0.68068169 -1.089659]
[-0.98383437 0.38809135]
[-1.20093113 0.30201192]
[1.07681071 -1.20563157]
[-1.49802789 -1.43757673]
[-0.60673761 -1.49556302]
[2.1661655 -0.79972756]
[-1.89411691 0.18681034]
[-0.21068659 0.85288166]
[-1.89411691 -1.26361786]
[2.1661655 0.38809135]
[-1.39890564 0.56295021]
[-1.10189888 -0.33583725]
[0.18552042 -0.65476184]
[0.38358493 0.01208048]
0s completed at 11:11 PM
27°C
Search
23:11
09/06/2023


Klasifikasi_Data_dengan_K_NN.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Sign in
Table of contents
Section
+ Code + Text
[21] from sklearn.neighbors import KNeighborsClassifier
classifier = KNeighborsClassifier(n_neighbors = 5, metric = 'minkowski', p=2)
classifier.fit(x_train, y_train)
KNeighborsClassifier
KNeighborsClassifier()
[22] y_pred = classifier.predict(x_test)
[23] from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test, y_pred)
print(cm)
[[64 4]
[3 29]]
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())
0s completed at 11:15 PM
Search


Klasifikasi_Data_dengan_K_NN.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Sign in
Table of contents
Section
+ Code + Text
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())
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-24-3d40f6b13854>:10: UserWarning: *c* argument looks like a single numeric RGB or RGBA sequence, which should be avoide
plt.scatter(x_set[y_set == j, 0], x_set[y_set == j, 1],
<function matplotlib.pyplot.show(close=None, block=None)>
Klasifikasi Data dengan K-NN (Data Training)
0
1
3
2
completed at 11:15 PM
Search


Klasifikasi_Data_dengan_K_NN.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Sign in
Table of contents
Section
+ Code + Text
<ipython-input-24-3d40f6b13854>:10: UserWarning: *c* argument looks like a single numeric RGB or RGBA sequence, which should be avoide
plt.scatter(x_set[y_set == j, 0], x_set[y_set == j, 1],
<function matplotlib.pyplot.show(close=None, block=None)>
Klasifikasi Data dengan K-NN (Data Training)
0
1
Estimasi Gaji
Umur
3
2
1
0
−1
−2
−2
−1
0
1
2
3
completed at 11:15 PM
Search