

Regresi Logistik		
Praktikan	Aslab	
Nama: xxxx	Annur Hangga Prihadi	065001800028
Nim: xxxx	Faiz Kumara	065001800003

PRAKTIKUM 7

DATA SAINS DAN ANALITIK

Topik pertemuan praktikum ke-tujuh adalah mengolah data penilaian suatu item dari kafe X menggunakan Regresi Logistik untuk menentukan apakah item tersebut perlu di-bundle atau tidak?

Source Code:

https://github.com/hanggaa/PrakDSDA/blob/main/Prak_7_Regresi_Logistik.ipynb

Latihan 1

1. Memasang library yang dibutuhkan

```
In [1]: import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn import metrics
import seaborn as sn
import matplotlib.pyplot as plt
```

2. Membaca file

```
In [2]: data = pd.read_csv("C:/Users/hangg/Downloads/Random Aslab/DSDA/Material/Order_Bundle_WP.csv", sep=";", #Sesuai
```

3. Menampilkan n data

```
In [3]: data.head(4)
```

```
Out[3]:
```

	Id	Penilaian_dari_pelanggan	Penilaian_dari_karyawan	Penilaian_dari_petinggi	Perlu_dipaket?
0	1	880	3	3	1
1	2	750	1	3	1
2	3	700	1	1	1
3	4	751	1	2	1

4. Menampilkan tipe data kolom

```
In [4]: data.dtypes

Out[4]: Id                int64
        Penilaian_dari_pelanggan  int64
        Penilaian_dari_karyawan   int64
        Penilaian_dari_petinggi   int64
        Perlu_dipaket?          int64
        dtype: object
```

5. Menampilkan deskripsi kolom

```
In [5]: data.describe()

Out[5]:
```

	Id	Penilaian_dari_pelanggan	Penilaian_dari_karyawan	Penilaian_dari_petinggi	Perlu_dipaket?
count	30.000000	30.000000	30.000000	30.000000	30.000000
mean	15.500000	671.400000	2.733333	2.166667	0.700000
std	8.803408	99.421291	1.229896	0.833908	0.466092
min	1.000000	504.000000	1.000000	1.000000	0.000000
25%	8.250000	595.000000	1.250000	1.250000	0.000000
50%	15.500000	691.000000	3.000000	2.000000	1.000000
75%	22.750000	745.500000	4.000000	3.000000	1.000000
max	30.000000	880.000000	4.000000	3.000000	1.000000

6. Menentukan variabel independen dan dependen

```
In [6]: X = data[['Penilaian_dari_pelanggan', 'Penilaian_dari_karyawan', 'Penilaian_dari_petinggi']]
        y = data['Perlu_dipaket?']
```

7. Mengatur data training dan testing

```
In [12]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.3, random_state = 0)
```

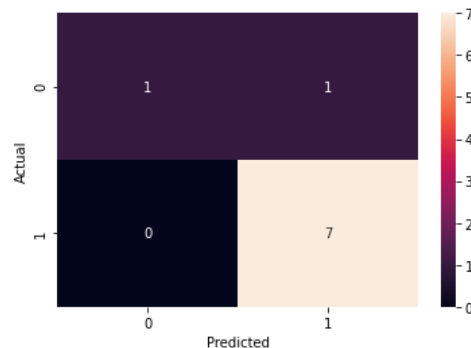
8. Menerapkan regresi logistik

```
In [13]: logistic_regression = LogisticRegression()
        logistic_regression.fit(X_train, y_train)
        y_pred = logistic_regression.predict(X_test)
```

9. Menampilkan confusion matrix

```
In [16]: confusion_matrix = pd.crosstab(y_test, y_pred, rownames=['Actual'], colnames=['Predicted'])  
sn.heatmap(confusion_matrix, annot=True)
```

```
Out[16]: <AxesSubplot:xlabel='Predicted', ylabel='Actual'>
```



10. Menampilkan akurasi dari prediksi

```
In [17]: print('Accuracy: ', metrics.accuracy_score(y_test, y_pred))  
plt.show()
```

```
Accuracy: 0.8888888888888888
```

11. Melihat data yang diuji

```
In [18]: print(X_test)
```

	Penilaian_dari_pelanggan	Penilaian_dari_karyawan	Penilaian_dari_petinggi
2	700	1	1
28	796	4	2
13	649	4	3
10	632	3	1
26	627	4	1
24	695	3	3
27	770	4	2
11	800	2	1
17	536	3	3

12. Melihat hasil prediksi

```
In [19]: print(y_pred)  
[1 1 1 1 1 1 1 1 1 1]
```

13. Menguji data yang baru

```
In [21]: new_predict = {'Penilaian_dari_pelanggan': [720, 600],  
                        'Penilaian_dari_karyawan': [2, 4],  
                        'Penilaian_dari_petinggi': [1, 3]}  
df2 = pd.DataFrame(new_predict, columns= ['Penilaian_dari_pelanggan', 'Penilaian_dari_karyawan', 'Penilaian_dari_petinggi'])
```

14. Menerapkan regresi logistik pada data yang baru

```
In [22]: y_pred=logistic_regression.predict(df2)
```

15. Menampilkan data baru yang akan diuji

```
In [23]: print (df2)
```

	Penilaian_dari_pelanggan	Penilaian_dari_karyawan	Penilaian_dari_petinggi
0	720	2	1
1	600	4	3

16. Menampilkan hasil data prediksi data baru yang baru saja diuji

```
In [24]: print (y_pred)
```

```
[1 1]
```

Latihan 2

1. Cari hasil prediksi uji data baru dengan ketentuan berikut

Penilaian_dari_pelanggan	Penilaian_dari_karyawan	Penilaian_dari_petinggi
560	2	1
660	2	2
700	4	2
600	1	1

Clue:

Ubah bagian yang ditandai dengan data di atas

```
new_predict = {'Penilaian_dari_pelanggan': [720, 600],  
               'Penilaian_dari_karyawan': [2, 4],  
               'Penilaian_dari_petinggi': [1, 3]  
              }
```

Lampiran Screenshot hasil

[Input screenshot disini](#)

Bagaimana interpretasi hasil prediksi di atas?

[Ketik interpretasi disini](#)