

UAS KONSEP APLIKASI DATA MAINING
REGRESI DENGAN MENGGUNAKAN ALGORITMA KNN
(SOAL NOMOR 4A)



Aminurachma Aisyah Nilatika
17.52.001

KEMENTRIAN RISET DAN TEKNOLOGI PENDIDIKAN TINGGI
SEKOLAH TINGGI MANAJEMEN INFORMATIKA KOMPUTER
PRADNYA PARAMITA
MALANG

2020

1. Yang pertama adalah melakukan import library, yang pertama adalah pandas nanti di sebut sebagai pd dan numpy sebagai np, seperti syntax berikut ini :

```
In [1]: import pandas as pd  
import numpy as np  
%matplotlib inline
```

2. Lalu sebelum kita melanjutkan nya, mari kita mengecek versi pandas yang kita gunakan dengan syntax berikut ini :

```
In [2]: pd.__version__
```

Dengan output :

```
Out[2]: '1.0.1'
```

3. Selanjutnya adalah membaca file excel menggunakan library pandas. Dimana nama file excel yang di gunakan adalah ('D:/Aminurachma Aisyah Nilatika/STIMATA/Semester 6/Data Mining/UAS/datamining/Uas/dataset_soal No. 2.xls') dengan syntax berikut ini :

```
In [12]: df = pd.read_csv('D:/Aminurachma Aisyah Nilatika/STIMATA/Semester 6/Data Mining/UAS/datamining/Uas/dataset_soal No. 4.txt',  
delimiter=',')
```

```
In [4]: df
```

Dengan output :

Out[4]:

	Usia	Kelahiran_ke-	Waktu_Kelahiran	Tekanan_darah	Kelainan_jantung	Caesarian
0	22	1	0	2	0	0
1	26	2	0	1	0	1
2	26	2	1	1	0	0
3	28	1	0	2	0	0
4	22	2	0	1	0	1
...
75	27	2	1	1	0	0
76	33	4	0	1	0	1
77	29	2	1	2	0	1
78	25	1	2	0	0	1
79	24	2	2	1	0	0

80 rows × 6 columns

4. Selanjutnya lakukan perhitungan Algoritma KNN dengan kondisi seperti Nomer 4A yaitu Ibu hamil dengan Usia 30 Tahun, yang merupakan Kelahiran ke -1, dengan Waktu kelahiran sesuai dengan HPL, Memiliki tekanan darah Normal, dengan syntax berikut ini :

```
In [19]: import math
dis = []
for i in range(80):
    dis.append(math.sqrt((float(df.iloc[i]['Usia'])-30)**2+
                        (float(df.iloc[i]['Kelahiran_ke-'])- 1)**2+
                        (float(df.iloc[i]['Waktu_Kelahiran'])-0)**2+
                        (float(df.iloc[i]['Tekanan_darah'])-1)**2))
```

Dan menampilkan hasil dari dis

```
In [20]: df['dis'] = dis
df
```

Dengan output :

Out[20]:

	Usia	Kelahiran_ke-	Waktu_Kelahiran	Tekanan_darah	Kelainan_jantung	Caesarian	dis
0	22	1	0	2	0	0	8.062258
1	26	2	0	1	0	1	4.123106
2	26	2	1	1	0	0	4.242641
3	28	1	0	2	0	0	2.236068
4	22	2	0	1	0	1	8.062258
...
75	27	2	1	1	0	0	3.316625
76	33	4	0	1	0	1	4.242641
77	29	2	1	2	0	1	2.000000
78	25	1	2	0	0	1	5.477226
79	24	2	2	1	0	0	6.403124

80 rows × 7 columns

5. Selajutnya, Library kita menyortir hasil dari dis dari yang terkecil dan mengambil hanya 5 teratas untuk mencari yang terdekat dengan kondisi yang sudah ada, dengan syntax :

```
In [22]: y = df.sort_values('dis').head(5)
y
```

Dengan output :

Out[22]:

	Usia	Kelahiran_ke-	Waktu_Kelahiran	Tekanan_darah	Kelainan_jantung	Caesarian	dis
27	30	1	0	1	0	0	0.000000
38	31	1	0	1	0	0	1.000000
67	29	2	0	1	1	0	1.414214
54	29	2	0	1	1	1	1.414214
59	30	2	1	2	1	1	1.732051

Lalu kita menampilkan hanya bagian Qty nya saja dengan syntax berikut:

```
In [23]: z = y["Caesarian"]
z
```

Dengan output :

Out[23]:

27	0
38	0
67	0
54	1
59	1

Name: Caesarian, dtype: int64

6. Langkah terakhir yaitu mencari rata rata berapa Qty atau banyak roti yang harus dibuat dengan cara berikut ini :

```
In [24]: np.mean(z)
```

Dengan output :

Out[24]: 0.4

Kesimpulan

Algoritma KNN untuk soal dengan kondisi Ibu hamil dengan Usia 30 Tahun, yang merupakan Kelahiran ke -1, dengan Waktu kelahiran sesuai dengan HPL, Memiliki tekanan darah Normal maka di temukan bahwa ibu hamil tersebut akan di berlakukan 0,4 atau mendekati 0 maka akan bersalin secara normal.