

UAS KONSEP APLIKASI DATA MAINING ALGORITMA KNN



Nama : Himatus Yulva Aprilia Sari

Jurusan :Sistem Infomasi

Nim :20.51.0018

**KEMENTRIAN RISET DAN TEKNOLOGI PENDIDIKAN TINGGI
SEKOLAH TINGGI MANAJEMEN INFORMATIKA KOMPUTER
PRADNYA PARAMITA
MALANG
2022**

3. Lakukan sebuah regresi dengan menggunakan Algoritma KNN dalam soal No. 3, dengan ketentuan sebagai berikut :

a. Umur : merupakan umur dari ibu hamil

b. Kelahiran ke- : merupakan kelahiran anak ke-... dari ibu tersebut

c. Waktu Kelahiran : 0 = sesuai dengan hari perkiraan lahir(HPL), 1 = prematur, 2 = melebihi hari perkiraan lahir

d. Tekanan_darah : 0 = untuk darah rendah, 1 = normal, 2 = tinggi

e. Kelainan Jantung : merupakan kelainan jantung untuk anak yang telah lahir 0 = normal, 1 = kelainan

f. Caesarian : 1 = Ya, 0 = Normal

Pertanyaan:

a. Berdasarkan data tersebut bagaimana perlakuan dengan kondisi Ibu hamil dengan Usia 30 Tahun, yang merupakan Kelahiran ke -1, dengan Waktu kelahiran sesuai dengan HPL, Memiliki tekanan darah Normal? Carilah KNN dengan menggunakan Key =8

JAWAB:

Data :

```
In [4]: df=pd.read_csv('C:/Users/Yulvavi/Documents/YULVA/Data Mining/dataset_soal_3.csv',delimiter=',')
```

```
In [5]: df
```

```
Out[5]:
```

	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

Dik = Usia = 30 th

Kelahiran =ke -1

Waktu kelahiran= 0

Tekanan darah Normal= 0

Ditanya: Carilah KNN dengan menggunakan Key = 8 ?

Gambar input data:

```
In [14]: 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))|
```

```
In [15]: df['dis'] = dis
df
```

Out[15]:

	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

```
In [16]: df.sort_values('dis')
```

Out[16]:

	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
...
41	19	1	0	1	0	1	11.000000
61	19	1	0	1	0	1	11.000000
25	18	1	0	1	0	0	12.000000
26	18	1	1	2	1	1	12.083046
70	17	1	0	0	0	1	13.038405

80 rows × 7 columns

Hasil dari key= 8

```
In [7]: y = df.sort_values('dis').head(8)
y
```

Out[7]:

	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
49	29	2	0	0	1	1	1.732051
77	29	2	1	2	0	1	2.000000
28	32	1	0	2	1	1	2.236068

```
In [8]: z = y["Caesarian"]
z
```

Out[8]:

27	0
38	0
67	0
54	1
59	1
49	1
77	1
28	1

Name: Caesarian, dtype: int64

Meneksport ke bentuk excel dengan Gambar berikut ini :

```
In [10]: df.to_excel('C:/Users/yulvavi/Outputs_UasNo3A.xls')
```

Jadi dari kesimpulan di atas hasil dari knn key=8 adalah 8 terdekat antara lain 0.00000, 1.00000, 1,414214, 1,414214, 1,732051, 1,732051, 2,00000 dan 2,236068 yang caesaria dengan normal ada 3 dan tidak normal ada 5. Rinciannya Antara lain :usia 30, 31, 29 itu caesarian =0 adalah normal dan usia 29,30,29,29, 32 caesarian=1 adalah Ya, maka kategori misterius tersebut adalah Tidak Normal (Ya)

b. Bagaimana Apabila Ibu hamil dengan Usia 29 Tahun, yang merupakan Kelahiran ke -2, dengan Waktu kelahiran sesuai dengan HPL, Memiliki tekanan darah Tinggi? Carilah KNN dengan menggunakan Key = 8

JAWAB:

Data :

```
In [4]: df=pd.read_csv('C:/Users/Yulvavi/Documents/YULVA/Data Mining/dataset_soal_3.csv',delimiter=',')
```

```
In [5]: df
```

Out[5]:

	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

Dik = Usia = 29 th

Kelahiran =ke -2

Waktu kelahiran= 0

Tekanan darah Tinggi= 2

Ditanya: Carilah KNN dengan menggunakan Key = 8 ?

Gambar input data:

```
In [22]: import math
dis = []
for i in range(80):
    dis.append(math.sqrt((float(df.iloc[i]['Usia'])-29)**2+
                        (float(df.iloc[i]['Kelahiran_ke-'])-2)**2+
                        (float(df.iloc[i]['Waktu_Kelahiran'])-0)**2+
                        (float(df.iloc[i]['Tekanan_darah'])-2)**2))
```

```
In [23]: df['dis'] = dis
df
```

```
Out[23]:
```

	Usia	Kelahiran_ke-	Waktu_Kelahiran	Tekanan_darah	Kelainan_jantung	Caesarian	dis
0	22	1	0	2	0	0	7.071068
1	26	2	0	1	0	1	3.162278
2	26	2	1	1	0	0	3.316625
3	28	1	0	2	0	0	1.414214
4	22	2	0	1	0	1	7.071068
...
75	27	2	1	1	0	0	2.449490
76	33	4	0	1	0	1	4.582576
77	29	2	1	2	0	1	1.000000
78	25	1	2	0	0	1	5.000000
79	24	2	2	1	0	0	5.477226

80 rows × 7 columns

```
In [24]: df.sort_values('dis')
```

```
In [24]: df.sort_values('dis')
```

```
Out[24]:
```

	Usia	Kelahiran_ke-	Waktu_Kelahiran	Tekanan_darah	Kelainan_jantung	Caesarian	dis
54	29	2	0	1	1	1	1.000000
77	29	2	1	2	0	1	1.000000
67	29	2	0	1	1	0	1.000000
59	30	2	1	2	1	1	1.414214
3	28	1	0	2	0	0	1.414214
...
61	19	1	0	1	0	1	10.099505
26	18	1	1	2	1	1	11.090537
31	40	1	0	1	1	1	11.090537
25	18	1	0	1	0	0	11.090537
70	17	1	0	0	0	1	12.206556

80 rows × 7 columns

Hasil dari key= 8

```
In [14]: y = df.sort_values('dis').head(8)
y
```

```
Out[14]:
```

	Usia	Kelahiran_ke-	Waktu_Kelahiran	Tekanan_darah	Kelainan_jantung	Caesarian	dis
54	29	2	0	1	1	1	1.000000
77	29	2	1	2	0	1	1.000000
67	29	2	0	1	1	0	1.000000
59	30	2	1	2	1	1	1.414214
3	28	1	0	2	0	0	1.414214
35	28	3	0	2	0	1	1.414214
8	28	2	0	1	0	0	1.414214
72	28	2	0	1	0	0	1.414214

```
In [15]: z = y["Caesarian"]
z
```

```
Out[15]: 54    1
77    1
67    0
59    1
3     0
35    1
8     0
72    0
Name: Caesarian, dtype: int64
```

Meneksport ke bentuk excel dengan Gambar berikut ini :

```
In [17]: df.to_excel('C:/Users/yulvavi/Outputs_UasNo3B.xls')
```

Jadi dari kesimpulan di atas hasil dari knn key=8 adalah 8 terdekat antara lain 1.000000, 1.000000, 1.000000, 1.414214, 1.414214, 1.414214, 1.414214 dan 1.414214 yang caesaria dengan normal ada 4 dan yang caisarian Ya ada 4 orang. Rinciannya Antara lain :usia 29, 29, 30 dan 28 Caisanya 1 = Ya(tidak normal) dan Usia 29, 28, 28 dan 28 Casisanya 0 = Normal maka kategori misterius tersebut adalah Ya (tidak Normal) dan Normal