

**UAS KONSEP APLIKASI DATA MAINING ALGORITMA KNN**



**Nama : Himatus Yulvi A.S**

**Jurusan :Sistem Infomasi**

**Nim :17.51.0005**

**KEMENTRIAN RISET DAN TEKNOLOGI PENDIDIKAN TINGGI  
SEKOLAH TINGGI MANAJEMEN INFORMATIKA KOMPUTER  
PRADNYA PARAMITA  
MALANG  
2020**

4. Lakukan sebuah regresi dengan menggunakan Algoritma KNN dalam soal No. 4, 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 = 5

JAWAB:

Data :

```
In [12]: df=pd.read_csv('C:/Users/Yulvavi/Documents/KULIAH/Himatus Yulvi/Sistem Data maining/dataset_soal_4.csv',delimiter=',')
```

```
In [13]: df
```

```
Out[13]:
```

	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 = 5 ?

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	<a href="#">13.038405</a>

80 rows × 7 columns

Hasil dari key= 5

```
In [17]: y = df.sort_values('dis').head(5)
y
```

```
Out[17]:
```

	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

```
In [18]: z = y["Caesarian"]
z
```

```
Out[18]: 27    0
38    0
67    0
54    1
59    1
Name: Caesarian, dtype: int64
```

Meneksport ke bentuk excel dengan Gambar berikut ini :

```
In [21]: df.to_excel('C:/Users/yulvavi/Outputs_KNN_Nomer4a.xls')
```

**Jadi dari kesimpulan** di atas hasil dari knn key=5 adalah 5 terdekat antara lain 0.00000, 1.00000, 1,414214, 1.414214 dan 1.732051 yang caesaria dengan normal ada 3 dan yang caisarian Ya ada 2 orang. Rinciannya Antara lain :usia 30, 31, 29 itu caesarian =0 adalah normal dan usia 29 dan 30 Caisanya 1 = Ya( tidak normal) maka kategori misterius tersebut adalah Normal

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 = 5

JAWAB:

Data :

```
In [12]: df=pd.read_csv('C:/Users/Yulvavi/Documents/KULIAH/Himatus Yulvi/Sistem Data maining/dataset_soal_4.csv',delimiter=',')
```

```
In [13]: df
```

```
Out[13]:
```

	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 = 5 ?

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	0	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	0	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= 5

```
In [25]: y = df.sort_values('dis').head(5)
y
```

```
Out[25]:
```

	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

```
In [26]: z = y["Caesarian"]
z
```

```
Out[26]: 54    1
77    1
67    0
59    1
3     0
Name: Caesarian, dtype: int64
```

```
In [27]: np.mean(z)
```

```
Out[27]: 0.6
```

Meneksport ke bentuk excel dengan Gambar berikut ini :

```
In [28]: df.to_excel('C:/Users/yulvavi/Outputs_KNN_Nomer4b.xls')
```

**Jadi dari kesimpulan** di atas hasil dari knn key=5 adalah 5 terdekat antara lain 1.000000, 1.000000, 1.000000, 1.414214 dan 1.414214 yang caesaria dengan normal ada 2 dan yang caisarian Ya ada 3 orang. Rinciannya Antara lain :usia 29 san 28 itu caesarian =0 adalah normal dan usia 29 ,29 dan 30 Caisanya 1 = Ya( tidak normal) maka kategori misterius tersebut adalah Ya ( tidak Normal)