

**UAS KONSEP APLIKASI DATA MAINING**  
**REGRESI DENGAN MENGGUNAKAN ALGORITMA KNN**  
**(SOAL NOMOR 2B)**



**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 [3]: df = pd.read_excel('D:/Aminurachma Aisyah Nilatika/STIMATA/Semester 6/Data Mining/UAS/datamining/Uas/dataset_soal No. 2.xls')
```

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

Dengan output :

```
Out[4]:
```

	Category	Weather V-1	Holiday V-2	Game V-3	Qty
0	A	5	1	0	250
1	B	3	1	1	200
2	C	1	1	0	75
3	D	4	1	1	400
4	E	4	0	0	150
5	F	2	0	0	50

4. Selanjutnya lakukan perhitungan Algoritma KNN dengan kondisi seperti Nomer 2b yaitu Cuaca baik dengan nilai 4, Weekend, dan Game =1, dengan syntax berikut ini :

```
In [5]: import math
dis = []
for i in range(6):
    dis.append(math.sqrt((float(df.iloc[i]['Weather V-1'])-4)**2
                        +(float(df.iloc[i]['Holiday V-2'])- 1)**2
                        +(float(df.iloc[i]['Game V-3'])-1)**2))
```

Dan menampilkan hasil dari dis

```
In [6]: df['dis'] = dis
df
```

Dengan output :

Out[6]:

	Category	Weather V-1	Holiday V-2	Game V-3	Qty	dis
0	A	5	1	0	250	1.414214
1	B	3	1	1	200	1.000000
2	C	1	1	0	75	3.162278
3	D	4	1	1	400	0.000000
4	E	4	0	0	150	1.414214
5	F	2	0	0	50	2.449490

5. Selajutnya, Library kita menyortir hasil dari dis dari yang terkecil dan mengambil hanya 4 teratas untuk mencari yang terdekat dengan kondisi yang sudah ada, dengan syntax :

```
In [7]: df.sort_values('dis').head(4)
```

Dengan output :

Out[7]:

	Category	Weather V-1	Holiday V-2	Game V-3	Qty	dis
3	D	4	1	1	400	0.000000
1	B	3	1	1	200	1.000000
0	A	5	1	0	250	1.414214
4	E	4	0	0	150	1.414214

Lalu kita masukkan hasil output tersebut ke variable y dengan syntax berikut ini :

```
In [10]: y = df.sort_values('dis').head(4)
```

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

```
In [9]: z = y["Qty"]  
z
```

Dengan output :

```
Out[9]: 3    400  
        1    200  
        0    250  
        4    150  
        Name: Qty, dtype: int64
```

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

```
In [11]: np.mean(z)
```

Dengan output :

```
Out[11]: 250.0
```

### **Kesimpulan**

Algoritma KNN untuk soal dengan kondisi Cuaca baik dengan nilai 4, Weekend, dan Game = 1 maka di temukan bahwa Jumlah roti yang harus dibuat adalah sebanyak 250 roti.