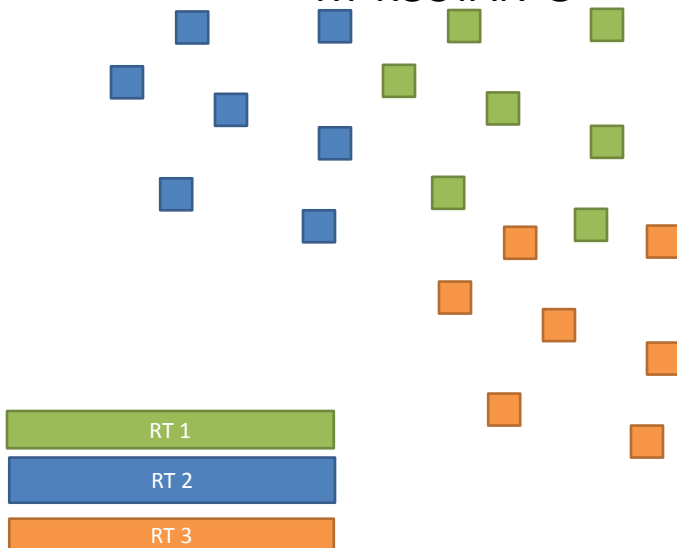


K Nearest Neighbour

Pertemuan 4
Metode Kuantitatif

Analogi sederhana

RT KOSTAN 😊



Analogi sederhana RT KOSTAN 😊

RT BERAPAKAH KOSTAN INI?

RT 1

RT 2

RT 3

PENENTUAN RT, DILAKUKAN DENGAN MELIHAT TETANGGA TERDEKAT

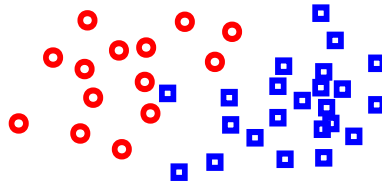
RT BERAPAKAH KOSTAN INI?

RT 1

RT 2

RT 3

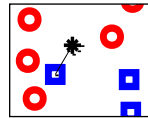
Nearest Neighbour Rule



Tersedia beberapa data (x,y) yang terklasifikasi menjadi 2 kelas

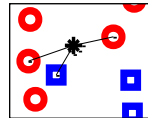
Diberikan query point q , titik ini akan ditentukan masuk ke kelas mana

$k = 1$



Menghitung sejumlah k tetangga terdekatnya dan menentukan kelas dari query point tsb dengan majority vote

$k = 3$



Algoritma KNN

- Tentukan k
- Hitung jarak data baru dengan semua data training

Ukuran jarak yang digunakan:

Euclidean distance

$$d(\mathbf{p}, \mathbf{q}) = d(\mathbf{q}, \mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}$$

Algoritma KNN (lanjutan..)

- Urutkan jarak tersebut dari yang terdekat
- Periksa kelas k tetangga terdekat
- Kelas data baru = Kelas mayoritas k tetangga terdekatnya

DATA KERTAS TISU SPESIAL

X1 = Acid Durability (second)	X2 = Strength (kg/m ²)	Class
7	7	Bad
7	4	Bad
3	4	Good
1	4	Good

Data baru: $X_1 = 3$, $X_2 = 7$, dengan $k = 3$ tentukan kelas dari data baru tersebut !

- $K = 3$, artinya jumlah tetangga yang dilihat adalah 3 tetangga terdekat
- Ada 2 kelas yaitu good dan bad
- X_1 menunjukkan acid durability
- X_2 menunjukkan strenght

Ayo berhitung

- Hitung jarak titik sampel dengan data training menggunakan *euclidean distance formula*

$$d(\text{data 1}, \text{data baru}) = \sqrt{(7-3)^2 + (7-7)^2} = 4$$

$$d(\text{data 2}, \text{data baru}) = \sqrt{(7-3)^2 + (4-7)^2} = 5$$

$$d(\text{data 3}, \text{data baru}) = \sqrt{(3-3)^2 + (4-7)^2} = 3$$

$$d(\text{data 4}, \text{data baru}) = \sqrt{(1-3)^2 + (4-7)^2} = 3.605$$

Urutkan hasilnya

$$\begin{aligned}
 d(\text{data 1}, \text{data baru}) &= \sqrt{(7-3)^2 + (7-7)^2} = 4 \longrightarrow \boxed{3} \\
 d(\text{data 2}, \text{data baru}) &= \sqrt{(7-3)^2 + (4-7)^2} = 5 \longrightarrow \boxed{4} \\
 d(\text{data 3}, \text{data baru}) &= \sqrt{(3-3)^2 + (4-7)^2} = 3 \longrightarrow \boxed{1} \\
 d(\text{data 4}, \text{data baru}) &= \sqrt{(1-3)^2 + (4-7)^2} = 3.605 \longrightarrow \boxed{2}
 \end{aligned}$$

$K = 3$, berarti ambil 3 tetangga terdekat

$$d(\text{data 1}, \text{data baru}) = \sqrt{(7-3)^2 + (7-7)^2} = 4$$

$$d(\text{data 3}, \text{data baru}) = \sqrt{(3-3)^2 + (4-7)^2} = 3$$

$$d(\text{data 4}, \text{data baru}) = \sqrt{(1-3)^2 + (4-7)^2} = 3.605$$

PERIKSA KELAS DARI 3 TETANGGA TERSEBUT

$$d(\text{data 1}, \text{data baru}) = \sqrt{(7-3)^2 + (7-7)^2} = 4 \quad \text{BAD}$$

$$d(\text{data 3}, \text{data baru}) = \sqrt{(3-3)^2 + (4-7)^2} = 3 \quad \text{GOOD}$$

$$d(\text{data 4}, \text{data baru}) = \sqrt{(1-3)^2 + (4-7)^2} = 3.605 \quad \text{GOOD}$$

Data baru: $X_1 = 3$, $X_2 = 7$ MASUK
KE KELAS

GOOD

KNN di Matlab

- $Class = knnclassify(Sample, Training, Group)$
 $Class = knnclassify(Sample, Training, Group, k)$
- Sample : Matrik yang setiap barisnya merupakan titik-titik kueri
- Training : Data yang diberikan sebagai calon neighbour dari kueri
- Group : Matrik hasil klasifikasi dari *Training*
- k : Berapa nearest neighbour yang akan dihitung. Defaultnya 1. k=3 berarti 3 Nearest Neighbour

BUAT MATRIKS UNTUK MENYIMPAN DATA LATIH

```
>> training = [7 7;7 4;3 4;1 4]
```

```
training =
```

```
    7    7  
    7    4  
    3    4  
    1    4
```


BUAT MATRIKS UNTUK MENYIMPAN DATA SAMPEL

```
>> sampel = [3 7]
```

```
sampel =
```

```
    3    7
```

MATRIKS UNTUK MENYIMPAN KELAS DARI DATA LATIH (0 = BAD, 1 = GOOD)

```
>> group = [0;0;1;1]
```

```
group =
```

```
    0
```

```
    0
```

```
    1
```

```
    1
```

Lakukan klasifikasi dengan perintah
`knnclassify(Sample, Training, Group, k)`

```
>> class = knnclassify(sampel, training, group, 3)
class =
     1
```

Terima Kasih