

Will

# Euclidian Distance

| 1     |    |       |      |     |          |     |         |          |       |       |        | Key |
|-------|----|-------|------|-----|----------|-----|---------|----------|-------|-------|--------|-----|
| Sorta | ID | Money | Free | For | Gambling | Fun | Machine | Learning | Spam  | ED    | Sorted |     |
| 3     | 1  | 3     | 0    | 0   | 0        | 0   | 0       | 0        | true  | 3     | 6      |     |
| 5     | 2  | 1     | 2    | 1   | 1        | 1   | 0       | 0        | true  | 2.828 | 4      | ✓   |
| 6     | 3  | 0     | 0    | 1   | 1        | 1   | 0       | 0        | true  | 1.732 | 1      | ✓   |
| 2     | 4  | 0     | 0    | 1   | 0        | 3   | 1       | 1        | false | 3.317 | 5      |     |
| 1     | 5  | 0     | 1    | 0   | 0        | 0   | 1       | 1        | false | 1.732 | 2      | ✓   |
| 4     | 6  | 0     | 1    | 1   | 0        | 0   | 1       | 1        | true  | 2     | 3      | ✓   |

Exclude

$$ED 1 = \sqrt{(3-0)^2 + (0-0)^2 \times 6} = \sqrt{9} = 3$$

$$ED 2 = \sqrt{(1-0)^2 + (2-0)^2 + (1-0)^2 \times 3 + (0-0)^2 \times 2} = \sqrt{8} = 2.828$$

$$ED 3 = \sqrt{(0-0)^2 \times 2 + (0-0)^2 \times 3 + (0-0)^2} = \sqrt{3} = 1.732$$

$$ED 4 = \sqrt{(0-0)^2 \times 2 + (1-0)^2 + (0-0)^2 + (3-0)^2 + (1-0)^2 \times 2} = \sqrt{11} = 3.317$$

$$ED 5 = \sqrt{(0-0)^2 + (0-1)^2 \times 3 + (0-0)^2 \times 3} = \sqrt{3} = 1.732$$

$$ED 6 = \sqrt{(1-0)^2 \times 4 + (0-0)^2 \times 3} = \sqrt{4} = 2$$

if  $k = 3$

$$M_k(q) = \arg \max_{l \in \text{levels}(t)} \sum_{i=1}^k x_i(t)$$

$$M_k(q) = \text{majority}((\text{true} \times 2) + (\text{false} \times 1))$$

3 nearest

meaning

$\left. \begin{array}{l} 2 = \text{true} \\ 3 = \text{true} \\ 5 = \text{false} \end{array} \right\} 2$  so  $\rightarrow$  the model will predict Spam = true for target level

Will find cos @ pt

186 409 1587

Based on 38-39 stars in K nearest

②

| id | money | free | or | grabbing | fun | media | learning | game  | Dr |
|----|-------|------|----|----------|-----|-------|----------|-------|----|
| 1  | 3     | 0    | 0  | 0        | 0   | 0     | 0        | true  |    |
| 2  | 1     | 2    | 1  | 1        | 1   | 0     | 0        | true  |    |
| 3  | 0     | 0    | 1  | 1        | 1   | 0     | 0        | true  |    |
| 4  | 0     | 0    | 1  | 0        | 0   | 1     | 1        | false |    |
| 5  | 0     | 1    | 0  | 0        | 0   | 1     | 1        | false |    |
| 6  | 0     | 1    | 1  | 0        | 0   | 1     | 1        | false |    |

$$ED1=3, ED2=2.828, ED3=1.732, ED4=3.317$$

$$ED5=1.732, ED6=2$$

cosine similarity For 3-NN = cos @

$$= \frac{A \cdot B}{\|A\| \|B\|} = \frac{A \cdot B}{\sqrt{A \cdot A} \sqrt{B \cdot B}} = \frac{A \cdot B}{\sqrt{A \cdot A} \sqrt{B \cdot B}}$$

Similarity(id1, id2) 3-NN to id6 = 2, 3, 5 =  $\frac{A \cdot B}{\|A\| \|B\|} = \frac{A \cdot B}{\sqrt{A \cdot A} \sqrt{B \cdot B}}$

id6 · id2  $(0 \cdot 1) + (1 \cdot 2) + (1 \cdot 1) + (0 \cdot 1) + (0 \cdot 0) + (1 \cdot 0) + (1 \cdot 0) = 3$

$$\sqrt{A \cdot A} = \sqrt{0+1+1+0+0+1+1} = \sqrt{4}$$

$$\sqrt{B \cdot B} = \sqrt{1+2+1+1+0+0+0} = \sqrt{5}$$

$$\frac{A \cdot B}{\|A\| \|B\|} = \frac{3}{\sqrt{4} \sqrt{5}} = 0.67082, A = id6, B = id2$$

id6 · id3  $(0 \cdot 0) + (1 \cdot 0) + (1 \cdot 1) + (0 \cdot 1) + (0 \cdot 1) + (1 \cdot 0) + (1 \cdot 0) = 1$

id6 · id3  $\sqrt{A \cdot A} = \sqrt{0+1+1+0+0+1+1} = \sqrt{4}$ ,  $\sqrt{B \cdot B} = \sqrt{0+0+1+1+1+0+0} = \sqrt{3}$

$$\frac{A \cdot B}{\|A\| \|B\|} = \frac{1}{\sqrt{4} \sqrt{3}} = 0.288675, A = id6, B = id3$$

id6 · id5  $(0 \cdot 0) + (1 \cdot 1) + (1 \cdot 0) + (0 \cdot 0) + (0 \cdot 0) + (1 \cdot 1) + (1 \cdot 1) = 3$

id6 · id5  $\sqrt{A \cdot A} = \sqrt{0+1+1+0+0+1+1} = \sqrt{4}$ ,  $\sqrt{B \cdot B} = \sqrt{0+1+0+0+0+1+1} = \sqrt{3}$

$$\frac{A \cdot B}{\|A\| \|B\|} = \frac{3}{\sqrt{4} \sqrt{3}} = 0.866025, A = id6, B = id5$$

id6 · id1  $(0 \cdot 3) + (1 \cdot 0) + (1 \cdot 0) + (0 \cdot 0) + (0 \cdot 0) + (1 \cdot 0) + (1 \cdot 0) = 0$

id6 · id1  $\sqrt{A \cdot A} = \sqrt{0+1+1+0+0+1+1} = \sqrt{4}$ ,  $\sqrt{B \cdot B} = \sqrt{3+0+0+0+0+0+0} = \sqrt{3}$

$$\frac{A \cdot B}{\|A\| \|B\|} = \frac{0}{\sqrt{4} \sqrt{3}} = 0, A = id6, B = id1$$



cos  $\theta$  p.2

2.

$$id6-id4 \quad (0 \cdot 0) + (1 \cdot 0) + (1 \cdot 1) + (0 \cdot 0) + (0 \cdot 3) + (1 \cdot 1) + (1 \cdot 1) = 3$$

$$A \cdot B \quad \sqrt{0^2+1^2+1^2+0^2+1+1} = \sqrt{4} \quad , \quad \sqrt{0^2+0^2+1+0+3+1+1} = \sqrt{6}$$

$$\frac{A \cdot B}{\|A\| \cdot \|B\|} = \frac{3}{\sqrt{4} \cdot \sqrt{6}} = 0.866025 \quad A=id6 \quad B=id4$$

Similarity(id<sub>i</sub>, id<sub>j</sub>)

|                          |                 |
|--------------------------|-----------------|
| (6,2) = 0.67082 = true   | = id 2 = NN     |
| (6,3) = 0.28865 = true   | = id 3 = Not NN |
| (6,5) = 0.866025 = false | = id 5 = NN     |
| (6,1) = 0 = true         | = id 1 = Not NN |
| (6,4) = 0.866025 = false | = id 4 = NN     |

3-NN

ids { 2 = true - 1  
5 = False  
4 = false } So target level with cosine similarity would be false