Calcolo della similarità tra vettori
x,y: vettori non normalizzati in uno spazio a N dimensioni
x’, y’: vettori x, y ***normalizzati*** con una qualche strategia

sim(x,y) = prodotto interno
cossim(x,y) = similarità del coseno

Caso 1: Vettori **non normalizzati**

$d_1=(5,1,2)$
$d_2=(0,1,1)$
$d=(1,2,1)$

$\mathrm{sim}(d, d_1) = 5+2+2=9$
$\mathrm{sim}(d, d_2) = 0+2+1=3$

Risultato: d_1 è più simile a d che d_2

$$\mathrm{cossim}(d, d_1) = \frac{5+2+2}{\sqrt{6}\sqrt{30}} = \frac{9}{\sqrt{6}\sqrt{30}} = \frac{sim(d,d_1)}{|d||d_1|} = \frac{9}{\sqrt{180}} = 0{,}671$$

$$\mathrm{cossim}(d, d_2) = \frac{0+2+1}{\sqrt{6}\sqrt{2}} = \frac{3}{\sqrt{6}\sqrt{2}} = \frac{sim(d,d_1)}{|d||d_2|} = \frac{3}{\sqrt{12}} = 0{,}866$$

Risultato: d_2 è più simile a d che d_1

Caso 2: V***ettori normalizzati con strategia del coseno***

$d_1=(5,1,2)$ $\quad |d_1| = \sqrt{5^2+1^2+2^2} = \sqrt{30}$
$d_2=(0,1,1)$ $\quad |d_2| = \sqrt{0^2+1^2+1^2} = \sqrt{2}$
$d=(1,2,1)$ $\quad |d| = \sqrt{1^2+2^2+1^2} = \sqrt{6}$

$$d'_1 = \left(\frac{5}{\sqrt{30}}, \frac{1}{\sqrt{30}}, \frac{2}{\sqrt{30}}\right)$$
$$d'_2 = \left(\frac{0}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$$

$$d' = \left(\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$$

$$\mathrm{sim}(d', d'_1) = \frac{5}{\sqrt{6}\sqrt{30}} + \frac{2}{\sqrt{6}\sqrt{30}} + \frac{2}{\sqrt{6}\sqrt{30}} = \frac{9}{\sqrt{6}\sqrt{30}} = \frac{sim(d,d_1)}{|d||d_1|} = \cos sim(d,d_1) = \frac{9}{\sqrt{180}} = 0{,}671$$

$$\text{sim}(d',d'_2)=\frac{0}{\sqrt{6}\sqrt{2}}+\frac{2}{\sqrt{6}\sqrt{2}}+\frac{1}{\sqrt{6}\sqrt{2}}=\frac{3}{\sqrt{6}\sqrt{2}}=\frac{sim(d,d_2)}{|d||d_2|}=\cos sim(d,d_2)=\frac{3}{\sqrt{12}}=0{,}866$$

Risultato: d_2 è più simile a d che d_1

Caso 3: V***ettori normalizzati con strategia del peso massimo***

d_1=(5,1,2) peso max=5
d_2=(0,1,1) peso max=1
d=(1,2,1) peso max=2

$d'_1=(1,\frac{1}{5},\frac{2}{5})$ $|d_1|=\sqrt{1^2+\left(\frac{1}{5}\right)^2+\left(\frac{2}{5}\right)^2}=\sqrt{1+\frac{5}{25}}=\sqrt{\frac{6}{5}}$

$d'_2=(0,1,1)$ $|d_2|=\sqrt{0^2+1^2+1^2}=\sqrt{2}$

$d'=(\frac{1}{2},1,\frac{1}{2})$ $|d|=\sqrt{\left(\frac{1}{2}\right)^2+1^2+\left(\frac{1}{2}\right)^2}=\sqrt{1+\frac{1}{2}}=\sqrt{\frac{3}{2}}$

$$\text{sim}(d',d'_1)=\frac{1}{2}+\frac{1}{5}+\frac{1}{5}=\frac{9}{10}=0{,}90$$

$$\text{sim}(d',d'_2)=0+1+\frac{1}{2}=1{,}50$$

$$\text{cossim}(d',d'_1)=\frac{\frac{1}{2}+\frac{1}{5}+\frac{1}{5}}{\sqrt{\frac{6}{5}}\sqrt{\frac{3}{2}}}=\frac{\frac{9}{10}}{\sqrt{\frac{18}{10}}}=0{,}671$$

$$\text{cossim}(d',d'_2)=\frac{0+1+\frac{1}{2}}{\sqrt{2}\sqrt{\frac{3}{2}}}=\frac{\frac{3}{2}}{\sqrt{3}}=0{,}866$$

Risultato: d_2 è più simile a d che d_1

Caso 4: V***ettori normalizzati con strategia della somma delle coordinate***

d_1=(5,1,2) somma=8
d_2=(0,1,1) somma=2
d=(1,2,1) somma=4

$d'_1=(\frac{5}{8},\frac{1}{8},\frac{2}{8})$ $|d_1|=\sqrt{\left(\frac{5}{8}\right)^2+\left(\frac{1}{8}\right)^2+\left(\frac{2}{8}\right)^2}=\sqrt{\frac{30}{64}}=\sqrt{\frac{15}{32}}$

$d'_2=(0,\frac{1}{2},\frac{1}{2})$ $|d_2|=\sqrt{0^2+\left(\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^2}=\sqrt{\frac{1}{2}}$

$$d' = (1/4, 2/4, 1/4) \qquad |d| = \sqrt{(1/4)^2 + (2/4)^2 + (1/4)^2} = \sqrt{3/8}$$

$$\mathrm{sim}(d', d'_1) = 5/32 + 2/32 + 2/32 = 9/32 = 0{,}28$$

$$\mathrm{sim}(d', d'_2) = 0 + 2/8 + 2/8 = 0{,}50$$

$$\mathrm{cossim}(d', d'_1) = \frac{5/32 + 2/32 + 2/32}{\sqrt{15/32}\sqrt{3/8}} = \frac{9/32}{\sqrt{45/256}} = 0{,}671$$

$$\mathrm{cossim}(d', d'_2) = \frac{0 + 1/4 + 1/8}{\sqrt{1/2}\sqrt{3/8}} = \frac{3/8}{\sqrt{3/16}} = 0{,}866$$

Risultato: d_2 è più simile a d che d_1