

#杰卡德距离

1.定义

Jaccard距离用来度量两个集合之间的差异性，它是Jaccard的相似系数的补集，被定义为1减去Jaccard相似系数。

$$J_{\sigma} = 1 - J(A, B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}$$

###2.性质

若A、B两个集合都为空，则J(A,B)=1

$0 \leq J(A, B) \leq 1$

3.python中的实现

```
from numpy import *
import scipy.spatial.distance as dist
matV = mat([[1,1,1,1],[1,0,0,1]])
print (dist.pdist(matV,'jaccard'))
结果: [0.5]
```

应用:

给定两个n维二元向量A、B，A、B的每一维都只能是0或者1，利用Jaccard相似系数来计算二者的相似性