Like Hamming, Jaccard is intended for boolean data, but tests for "not equal".

$$d(u,v) = \frac{|\{j|u_j \neq v_j\}|}{|\{j|u_j \neq 0 \; or \; v_j \neq 0\}|}$$

HCluster Vector Dissimilarity Calculation Methods for Boolean Data

The remaining metrics are for boolean data. Any data is accepted and each value is converted to a boolean value by testing for u_j != 0.

The following definitions are used:

$$a = |\{j \mid u_j \wedge v_j\}|$$

$$b = |\{j \mid u_j \wedge (\neg v_j)\}|$$

$$c = |\{j \mid (\neg u_j) \wedge v_j\}|$$

$$d = |\{j \mid (\neg u_j) \wedge (\neg v_j)\}|$$

D is the vector length, $D = a + b + c + d$.

***dm* = Yule**

$$d\,(u,v) = \frac{2bc}{ad + bc}$$

***dm* = Dice**

$$d\,(u,v) = \frac{b+c}{2a+b+c}\;,\; d\,(0,0) = 0$$

***dm* = RogersTanimoto**

$$d\,(u,v) = \frac{2\,(b+c)}{b+c+D}$$

***dm* = RusselRao**

$$d\,(u,v) = \frac{b+c+d}{D}$$

***dm* = SokalSneath**

$$d\,(u,v) = \frac{2(b+c)}{a+2(b+c)}\;,\; d(0,0) = 0$$

***dm* = Kulsinski**

$$d\,(u,v) = \frac{1}{2} \cdot \left(\frac{b}{a+b} + \frac{c}{a+c}\right)$$