FOWLKES-MALLOWS

SCORE

Evaluate clusters when ground-truth is available. Used to evaluate pairs of clusters. Higher score is better. Ranges between 0 and 1

$$FMI = \frac{TP}{\sqrt{(TP + FP)(TP + FN)}}$$

Fowlkes-Mallows Index

True Positive

False Positive

False Negative

TP: # of pairs of observations are part of the same cluster in both y and $\hat{y}$.

FP: # of pairs of observations that are part of the same cluster in y but part of different clusters in $\hat{y}$.

FN: # of pairs of observations that are not part of the same cluster in y but part of the same cluster in $\hat{y}$