

# Pairwise Measures: Jaccard Coefficient and Rand Statistic

- ❑ **Jaccard coefficient:** Fraction of true positive point pairs, but after ignoring the true negatives (thus asymmetric)

- ❑  $Jaccard = TP / (TP + FN + FP)$  [i.e., denominator ignores  $TN$ ]

- ❑ Perfect clustering:  $Jaccard = 1$

- ❑ **Rand Statistic:**

- ❑  $Rand = (TP + TN) / N$

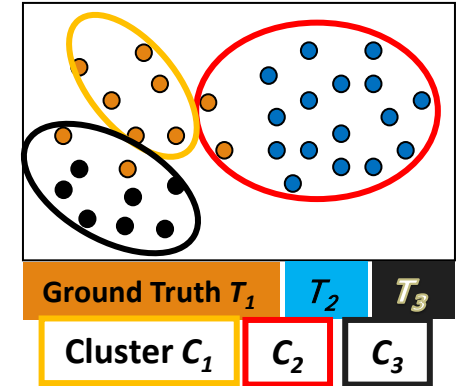
- ❑ Symmetric; perfect clustering:  $Rand = 1$

- ❑ **Fowlkes-Mallow Measure:**

- ❑ Geometric mean of precision and recall

$$FM = \sqrt{prec \times recall} = \frac{TP}{\sqrt{(TP + FN)(TP + FP)}}$$

- ❑ Using the above formulas, one can calculate all the measures for the green table (leave as an exercise)



| $C \backslash T$ | $T_1$ | $T_2$ | $T_3$ | Sum |
|------------------|-------|-------|-------|-----|
| $C_1$            | 0     | 20    | 30    | 50  |
| $C_2$            | 0     | 20    | 5     | 25  |
| $C_3$            | 25    | 0     | 0     | 25  |
| $m_j$            | 25    | 40    | 35    | 100 |