HOMEWORK 2 | TASK 3

CSCI-548

Sarvesh Parab
sparab@usc.edu

Performance Evaluation:

- **'Person' Entity**
    1. True Positives $(\mathcal{TP})$: 14
    2. False Positives $(\mathcal{FP})$: 1
    3. False Negatives $(\mathcal{FN})$: 3

    Precision $\mathcal{P} = \dfrac{\mathcal{TP}}{(\mathcal{TP} + \mathcal{FP})} = \dfrac{14}{15} = 0.933$

    Recall $\mathcal{R} = \dfrac{\mathcal{TP}}{(\mathcal{TP} + \mathcal{FN})} = \dfrac{14}{17} = 0.823529$

    F1 measure $= \dfrac{2\mathcal{PR}}{(\mathcal{P} + \mathcal{R})} = 0.875$

- **'Organization' Entity**
    1. True Positives $(\mathcal{TP})$: 9
    2. False Positives $(\mathcal{FP})$: 2
    3. False Negatives $(\mathcal{FN})$: 33

    Precision $\mathcal{P} = \dfrac{\mathcal{TP}}{(\mathcal{TP} + \mathcal{FP})} = \dfrac{9}{11} = 0.8181$

    Recall $\mathcal{R} = \dfrac{\mathcal{TP}}{(\mathcal{TP} + \mathcal{FN})} = \dfrac{9}{42} = 0.2142$

    F1 measure $= \dfrac{2\mathcal{PR}}{(\mathcal{P} + \mathcal{R})} = 0.339623$

- **'Organization' Entity (Considering 'Company' equivalent to 'Organization')**
    1. True Positives $(\mathcal{TP})$: 34
    2. False Positives $(\mathcal{FP})$: 2
    3. False Negatives $(\mathcal{FN})$: 8

    Precision $\mathcal{P} = \dfrac{\mathcal{TP}}{(\mathcal{TP} + \mathcal{FP})} = \dfrac{34}{36} = 0.944$

    Recall $\mathcal{R} = \dfrac{\mathcal{TP}}{(\mathcal{TP} + \mathcal{FN})} = \dfrac{34}{42} = 0.80952$

    F1 measure $= \dfrac{2\mathcal{PR}}{(\mathcal{P} + \mathcal{R})} = 0.871795$

NOTE: Calais tags terms like 'Microsoft' or 'Apple' as "Company" and not an "Organization". Thus I have considered both cases:
1. when I assume that those terms were wrongly tagged as "Company" and they should have been tagged as "Organization" and the tagged matched terms are taken as 'False Negatives'.
   They are marked as FN* in the CSV file and considered as FN
2. when I assume that "Company" and "Organization" are equivalent and the tagged matching terms are taken as 'True Positives'
   They are marked as FN* in the CSV file and considered as TP