Entity Linking Evaluation

Lecture 17, Oct 29, 2019

In this exercise, you will evaluate the performance of entity linking systems.

We provide you with two sample texts and the ground truth, and you need to evaluate the output of an entity linking system.

Text 1

Ground truth:

Mention	**Entity**
mpla	⟨wikipedia:MPLA⟩
angola	⟨wikipedia:Angola⟩
1992 elections	⟨wikipedia:Angolan_presidential_election,_1992⟩
multiparty democracy	⟨wikipedia:Multi-party_system⟩

System output:

Mention	**Entity**
1992 elections	⟨wikipedia:Philippine_general_election,_1992⟩
angola	⟨wikipedia:Angola⟩
multiparty democracy	⟨wikipedia:multiparty_democracy⟩

Text 2

Ground truth:

Mention	**Entity**
sweet potato	⟨wikipedia:Sweet_potato⟩
bell pepper	⟨wikipedia:Bell_pepper⟩
tomato	⟨wikipedia:Tomato⟩

System output:

Task 1

Given the above ground truth and system results, answer the following questions:

1. What is Precision for Text 1?
2. What is Precision for Text 2?
3. What is P_{macro}?
4. What is Recall for Text 1?
5. What is Recall for Text 2?
6. What is R_{macro}?
7. What is the $F1$ score?

Micro-averaged evaluation metrics are computed as:

$$P_{micro} = \sum_{d \in D} \frac{|\hat{A} \cap A|}{A} / |D| \qquad R_{micro} = \sum_{d \in D} \frac{|\hat{A} \cap A|}{\hat{A}} / |D| \qquad F1 = \frac{2 \times P \times R}{P + R}$$

where $\hat{A}$ denotes ground truth annotations, A is system generated annotations, and $|D|$ is the total number of documents (i.e. 2 in this exercise).

Mention	Entity
potato	⟨wikipedia:Sweet_potato⟩
perennials	⟨wikipedia:Perennial_plant⟩

Task 2

Given the above ground truth and system results, answer the following questions:

1. What is P_{micro}?
2. What is R_{micro}?
3. What is $F1$ score?

Micro-averaged evaluation metrics are computed as:

$$P_{micro} = \frac{|\hat{A} \cap A|}{A} \qquad R_{micro} = \frac{|\hat{A} \cap A|}{\hat{A}} \qquad F1 = \frac{2 \times P \times R}{P + R}$$

where $\hat{A}$ denotes ground truth annotations and A is system generated annotations.