

# Natural Language Processing

## Assignment 10

Type of Question: MCQ

**Number of Questions: 10**

**Total Marks:(10×1)= 10**

=====

**Question 1.** Common Steps of entity linking are -

- A) Reference Disambiguation -> Candidate Selection
- B) Reference Disambiguation -> Candidate Selection -> Mention Identify
- C) Mention Identify -> Candidate Selection -> Reference Disambiguation
- D) All of the above

**[Marks 1]**

**Answer: C**

**Solution:** Theory. Slide 7; Lecture 1, Week 10 (Entity Linking 1).

=====

**Question 2.** The text span s="world" occurs in 764 different Wikipedia articles.

- c1 189
- c2 273
- c3 87
- c4 53
- No Link 162

Calculate the keyphraseness of "world". **[Marks 1]**

**Answer:**

- A) 0.232
- B) 0.788
- C) 0.688
- D) 0.976

**Solution: B)**

$$CF(s_i) / CF(s) = 189 + 273 + 87 + 53 / 764 = 602 / 764 = 0.788$$

=====

**Question 3.** What is the commonness of (s, c2) in the above question?

**[Marks 1]**

**Answer:**

- A) 0.765
- B) 0.389
- C) 0.453
- D) 0.910

**Solution:** C)  $273/(189 + 273 + 87 + 53) = 273/602 = 0.453$

=====

**Question 4. Higher value of keyphraseness represents higher probability of:**

- A) an article to get selected as linkable candidate
- B) a phrase to get detected as a mention
- C) an article to get disambiguated from other candidates
- D) None of the above

**Solution: B)** a phrase to get detected as a mention. Lecture Slide.

=====

**Question 5.** Which of the following problem exists in bootstrapping technique for Information extraction are: **[Marks 1]**

- A) Sensitiveness towards the seed set
- B) High precision
- C) Less manual intervention
- D) All of the above

**Answer: A**

**Solution:** Theory

=====

**Question 6.** Which of the following is an advantage of unsupervised relation extraction: **[Marks 1]**

- A) Can work efficiently with small amount of hand-labeled data
- B) Not easily generalizable to different relations
- C) Need training data.
- D) Always perform better than supervised techniques.
- E) None of the above

**Answer: E**

**Solution:** Theory

=====

**Question 7.** Which of the following is/are a Hearst's Lexico Syntactic Patterns for automatic acquisition of hyponyms -

- A) X or other Y
- B) X and other Y
- C) Y including X
- D) X but not Y

**[Marks 1]**

**Answer:** A, B, C

**Solution:** Theory. Lecture Video 48 : Information Extraction - Introduction, Week 10, Lecture 3 - Slide 15/18.

=====

**Question 8.** Advantage of Distant supervision over bootstrapping method

- A) Need more data
- B) Less human effort
- C) Can handle noisy data better
- D) No Advantage

**[Marks 1]**

**Answer:** C

**Solution:** Theory

=====

**Question 9.** Consider a dataset with a very low number of relations - all of which are very important. For a relation extraction task on that dataset, which of the following is the most useful metric

- A) Precision
- B) Recall
- C) Accuracy
- D) F1-Score

**[Marks 1]**

**Answer:** B

**Solution:** Theory

=====

**Question 10.** Bootstrapping can be considered as

- A) Supervised Approach
- B) Unsupervised Approach
- C) Semi-supervised Approach
- D) All of the above
- E) None of the above

**[Marks 1]**

**Answer:** C

**Solution:** Theory. Slide => Relation Extraction, Slide - 2/17, Week4, Lecture 10.

=====