NLP week 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

Option C) is correct. Mention identify -> candidate selction -> reference disambiguation.

Question 2

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

$$c1 \quad 189$$
$$c2 \quad 273$$
$$c3 \quad 87$$
$$c4 \quad 53$$
$$No\ Link \quad 162$$

Calculate the keyphraseness of "world".

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

0.7879 Option B)

Question 3

What is the commonness of (s, c2) in the above question?

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

Option C) is correct.

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

Option B) Mention Identify.

Question 5

Which of the following problem exists in bootstrapping technique for Information extraction are:

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

Option A)

Question 6

Which of the following is an advantage of unsupervised relation extraction:

  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

D) is not correct
C) is not correct as it doesn't need any training data.
B) is not correct AS IT IS NOT AN ADVANTAGE
A) is not correct AS IT IS NOT AN ADVANTAGE
Option E) is correct.

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

Multi-choice multi correct
Option A), B) and C) will be correct.

Hearst Lexico syntactic patterns ⇒ It is rule based kind of approach

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

Option C) It can handle noisy data better.

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

Option B) Recall.

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

Option c) semi-supervised approach.