

Unit 3 - Week 1

Course outline
How to access the portal
Week 0 : Assignment 0
Week 1
<ul style="list-style-type: none"> Lecture 1 : Introduction to the Course Lecture 2 : What Do We Do in NLP Lecture 3: Why is NLP hard Lecture 4: Empirical Laws Lecture 5: Text Processing: Basics Week 1: Lecture Materials
<input type="radio"/> Quiz : Assignment 1
<input type="radio"/> Feedback for Week 1
Week 2
Week 3
Week 4
Week 5
Week 6
Week 7
Week 8
Week 9
Week 10
Week 11
Week 12
DOWNLOAD VIDEOS
Solutions

Assignment 1

The due date for submitting this assignment has passed.
As per our records you have not submitted this assignment.

Due on 2019-08-14, 23:59 IST.

1) **Which of the following doesn't require application of Natural Language Processing algorithms?** 1 point

- Classifying spam emails from good ones
- Classifying images of scanned documents as "hand-written" or "printed" documents
- Automatically generating captions for images
- Building a sentiment analyzer for tweets on Twitter.

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
2.

2) **One of the main challenges of NLP is _____** 1 point

- Ambiguity of sentences
- Handling tokenization
- Usage of non-standard English
- New senses of a word

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
1.

3) **Which one is not an example of neologisms:** 1 point

- Cryptocurrency
- Blogging
- Friendship
- Googling

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
3.

4) **Ambiguity can appear in which of the following steps / tasks ?** 1 point

- Tokenization
- Language Understanding
- Sentence Segmentation
- All of these

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
4.

5) **Which of the following sentence contains a ditransitive verb usage?** 1 point

- Maureen gave Dan the pencil.
- She lied.
- That pumpkin pie smells delicious.
- The dog chased the cats.

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
1.

6) **If first corpus has $TTR_1 = 0.013$ and second corpus has $TTR_2 = 0.13$, where TTR_1 and TTR_2 represents type/token ratio in first and second corpus respectively, then** 1 point

- First corpus has more tendency to use different words.
- Second corpus has more tendency to use different words.
- Both a and b
- None of these

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
2.

7) **Which of the following is/are true for English Language?** 1 point

- Lemmatization works only on inflectional morphemes and Stemming works only on derivational morphemes.
- The outputs of lemmatization and stemming for the same word might differ.
- Output of lemmatization are always real words
- Output of stemming are always real words

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
2.
3.

8) **Which of the following are instances of stemming? (as per Porter Stemmer)** 1 point

- are -> be
- plays -> play
- saw -> s
- university -> univers

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
2.
4.

9) **As per Zipf's law, the correct statement about a corpus is:** 1 point

- 10th most common word will occur with 10 times the frequency of the 100th most common word.
- 100th most common word will occur with 10 times the frequency of the 10th most common word
- Frequency of a word is directly proportional to its position in the ranked list.
- None of these

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
1.

10) **Which one is not related to the concept of decision tree algorithm:** 1 point

- PCA
- ID3
- Random Forest
- C4.5

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
1.

11) **Word segmentation is mostly used when:** 1 point

- Hyphens are present
- Multiple alphabets intermingled
- Long sentences
- No space between words

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
4.

12) **What is the valid range of type-token ratio of any text corpus?** 1 point

- $TTR \in (0, 1]$ (excluding zero)
- $TTR \in [0, 1]$
- $TTR \in [-1, 1]$
- $TTR \in [0, +\infty]$ (any non-negative number)

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
1.

13) **Find the type-token ratio for following sentence,** 1 point

But what are thoughts? Well, we all have them. They are variously described as ideas, notions, concepts, impressions, perceptions, views, beliefs, opinions, values, and so on. At times they are brief, coming and going in an instant.

- 1.0
- 37/33
- 33/37
- No space between words

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
3.

14) **In the sentence, "In Delhi I took my hat off. But I can't put it back on.", total number of word tokens and word types are:** 1 point

- 14, 13
- 13, 14
- 15, 14
- 14, 15

1.
 2.
 3.
 4.

No, the answer is incorrect.
Score: 0

Accepted Answers:
1.