X

Week 9 : Assignment 9

The due date for submitting this assignment has passed.

Due on 2024-09-25, 23:59 IST.

Assignment submitted on 2024-09-25, 23:10 IST

1) Which techniques can be utilized to assess the quality of the word embeddings generated by our model? ***1 point***

- Evaluating semantic relatedness
- Identifying synonyms
- Solving analogy problems
- All of the above

Yes, the answer is correct.
Score: 1

Accepted Answers:
All of the above

2) Let X be the co-occurrence matrix such that the (i, j)-th entry of X captures the PMI between the i-th and j-th word in the corpus. Every row of X corresponds to the representation of the i-th word in the corpus. Suppose each row of X is normalized (i.e., the L_2 norm of each row is 1) then the (i, j)-th entry of XX^T captures the: ***1 point***

- PMI between word i and word j
- Euclidean distance between word i and word j
- Probability that word i
- Cosine similarity between word i

Yes, the answer is correct.

Score: 1

Accepted Answers:
Cosine similarity between word i

3) Consider the following corpus: "human machine interface for computer applications. **1 point** user opinion of computer system response time. user interface management system. system engineering for improved response time". What is the size of the vocabulary of the above corpus?

- 13
- 14
- 15
- 16

Yes, the answer is correct.
Score: 1

Accepted Answers:
15

4) At the input layer of a continuous bag of words model, we multiply a one-hot vector **1 point** $x \in \mathbb{R}^{|V|}$ with the parameter matrix $W \in \mathbb{R}^{k \times |V|}$. What does each column of W correspond to?

- the representation of the i-th word in the vocabulary
- the i-th eigen vector of the co-occurrence matrix

Yes, the answer is correct.
Score: 1

Accepted Answers:
the representation of the i-th word in the vocabulary

5) Suppose that we use the continuous bag of words (CBOW) model to find vector representations of words. Suppose further that we use a context window of size 3 (that is, given the 3 context words, predict the target word $P(w_t|(w_i, w_j, w_k))$). The size of word vectors (vector representation of words) is chosen to be 100 and the vocabulary contains 10,000 words. The input to the network is the one-hot encoding (also called 1-of-V encoding) of word(s). How many parameters (weights), excluding bias, are there in W_{word}? Enter the answer in thousands. For example, if your answer is 50,000, then just enter 50.

1000

Yes, the answer is correct.
Score: 1

Accepted Answers:
(Type: Numeric) 1000

1 point

6) Let $count(w, c)$ be the number of times the words w and c appear together in the **1 point** corpus (i.e., occur within a window of few words around each other). Further, let $count(w)$ and $count(c)$ be the total number of times the word w and c appear in the corpus respectively and let N be the total number of words in the corpus. The PMI between w and c is then given

$\log \frac{count(w,c)*count(w)}{N*count(c)}$

$\log \frac{count(w,c)*count(c)}{N*count(w)}$

$\log \frac{count(w,c)*N}{count(w)*count(c)}$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$\log \frac{count(w,c)*N}{count(w)*count(c)}$

7) Which of the following is an advantage of the CBOW model compared to the Skip-gram model? ***1 point***

- It is faster to train
- It requires less memory
- It performs better on rare words
- All of the above

No, the answer is incorrect.
Score: 0
Accepted Answers:
It is faster to train

8) Which of the following is true about the input representation in the CBOW model? ***1 point***

- Each word is represented as a one-hot vector
- Each word is represented as a continuous vector
- Each word is represented as a sequence of one-hot vectors
- Each word is represented as a sequence of continuous vectors

Yes, the answer is correct.
Score: 1
Accepted Answers:
Each word is represented as a one-hot vector

9) Which of the following is an advantage of using the skip-gram method over the bag-of-words approach? ***1 point***

- The skip-gram method is faster to train
- The skip-gram method performs better on rare words
- The bag-of-words approach is more accurate
- The bag-of-words approach is better for short texts

Yes, the answer is correct.
Score: 1
Accepted Answers:
The skip-gram method performs better on rare words

10) We add incorrect pairs into our corpus to maximize the probability of words that occur in the same context and minimize the probability of words that occur in different con... This technique is called- ***1 point***

- Hierarchical softmax
- Contrastive estimation
- Negative sampling
- Glove representations

No, the answer is incorrect.
Score: 0

Accepted Answers:
Negative sampling