

Quiz 1: Coursework

The Preprocessing Step

Please enter your name: *

HM

Preprocessing the data

Let us consider the following corpus

Raw Corpus

$\mathcal{D}_1 =$ Neural Networks are awesome

$\mathcal{D}_2 =$ LSTMs are Sequential Neural Networks

$\mathcal{D}_3 =$ Attention Models are awesome

The word2idx dictionary associated with the Raw Corpus is the following dictionary:

Word2idx = { Neural : 1,
 Networks : 2,
 are : 3,
 awesome : 4,
 LSTMs : 5,
Sequential : 6,
 Attention : 7,
 Models : 8 }

What is the one hot vector representing the word "Sequential" ?

1 point

- 6
- [0, 0, 0, 0, 0, 1, 0, 0]
- [0, 0, 0, 0, 0, 6, 0, 0]

Using the word2idx dictionary, which answer represents the processed raw corpus into lists of integers, as explained in the Coursework 1 point

(a)

$$\begin{aligned} \mathcal{D}_1 &= [0, 0, 0, 1] \\ \mathcal{D}_2 &= [0, 1, 0, 0, 0] \\ \mathcal{D}_3 &= [0, 0, 1, 0] \end{aligned}$$

(b)

$$\begin{aligned} \mathcal{D}_1 &= [1, 2, 3, 4] \\ \mathcal{D}_2 &= [5, 3, 6, 1, 2] \\ \mathcal{D}_3 &= [7, 8, 3, 4] \end{aligned}$$

(c)

$$\begin{aligned} \mathcal{D}_1 &= [1, 1, 1, 1, 0, 0, 0, 0] \\ \mathcal{D}_2 &= [1, 1, 1, 0, 1, 1, 0, 0] \\ \mathcal{D}_3 &= [0, 0, 1, 1, 0, 0, 1, 1] \end{aligned}$$

 (a) (b) (c)

Introducing the Context Words

Let us consider a context size = 2 in the rest of the section

What are the context words associated with the center word 5? 1 point

$$\mathcal{D}_2 = [5, 3, 6, 1, 2]$$

 [3, 6, 1] [1, 2, 3, 6] [3, 6]

What are the context words associated with the center word 3 ?

1 point

$$\mathcal{D}_2 = [5 , \textcircled{3} , 6 , 1 , 2]$$

- [5, 6]
- [5, 6, 1]
- [6, 1, 2]

What are the context words associated with the center word 6 ?

1 point

$$\mathcal{D}_2 = [5 , 3 , \textcircled{6} , 1 , 2]$$

- [3, 6]
- [3, 6, 2]
- [5, 3, 1, 2]

What are the context words associated with the center word 1 ?

1 point

$$\mathcal{D}_2 = [5 , 3 , 6 , \textcircled{1} , 2]$$

- [3, 6, 2]
- [6, 2]
- [2]

What are the context words associated with the center word 2 ?

1 point

$$\mathcal{D}_2 = [5 , 3 , 6 , 1 , \textcircled{2}]$$

- [6, 1]
- [3, 6, 1]
- [5, 3, 6, 1]

What is the sum of the total number of context words associated with each center word in document \mathcal{D}_2 ?

1 point

- 12
- 9
- 14

What is the sum of the total number of context words associated with each center word in document D_2 , as a function of n_2 (the length of document D_2)? (with the assumption $n_2 \geq 4$). 2 points

$$\mathcal{D}_2 = [w_2^1, w_2^2, \dots, w_2^t, \dots, w_2^{n_2-1}, w_2^{n_2}]$$

- $10 + 4*(n_2 - 4)$
- $2*n_2 + 4$
- $3*n_2 - 1$

Any question ?

This content is neither created nor endorsed by Google.

Google Forms