Feedback — Week 1 Quiz

Help Center

You submitted this quiz on **Fri 10 Jul 2015 11:06 PM PDT**. You got a score of **9.00** out of **10.00**.

Question 1

A paradigmatic relation is a relation between two words that tend to *co-occur* with each other, while a syntagmatic relation is between two words that tend to occur in *similar* context.

Your Answer		Score	Explanation
False	✔	1.00	
True			
Total		1.00 / 1.00	

Question 2

In a collection of English news articles, which word do you expect to have a higher IDF?

Your Answer		Score	Explanation
“the”			
“learning”	✔	1.00	
Total		1.00 / 1.00	

Question 3

Let X_w be a random variable denoting whether word w occurs in a text document in a collection of English news articles. Which random variable do you expect to have a **lower** entropy?

Your Answer	Score	Explanation

◉ $H(X_{the})$	✔	1.00	
○ $H(X_{learning})$			
Total		1.00 / 1.00	

Question 4

Suppose the pseudo-document representations for the contexts of the terms A and B in the vector space model are given as follows:

d_A = (0.10, 0.50, 0.00, 0.40, 0.00, 0.00)

d_B = (0.20, 0.40, 0.30, 0.00, 0.10, 0.00)

What is the EOWC similarity score?

Your Answer		Score	Explanation
◉ 0.22	✔	1.00	
○ 1			
○ 0.02			
○ 0.20			
Total		1.00 / 1.00	

Question 5

Let X be a binary random variable. Which of the following is **not** true?

Your Answer		Score	Explanation
○ $H(X) \leq 1$			
○ If P(X=0)=1 then H(X) = 0			
◉ If P(X=1)=1 then H(X) = 1	✔	1.00	
Total		1.00 / 1.00	

Question 6

For any two binary random variables X and Y, which of the following does **not** always hold?

Your Answer		Score	Explanation
○ If X=Y then H(X|Y) = 0			
◉ H(X|Y) = H(Y|X)	✔	1.00	
○ H(X|Y) $\geq$ 0			
○ H(X|Y) $\leq$ H(X)			
Total		1.00 / 1.00	

Question 7

Let X_{text}, X_{mining}, and X_{the} be binary random variables associated with the words “text”, “mining”, and “the” respectively. Assume that the probabilities of the random variables are estimated based on a large corpus. Then we should expect $I(X_{text} ; X_{mining}) > I(X_{text} ; X_{the})$.

Your Answer		Score	Explanation
◉ True	✔	1.00	
○ False			
Total		1.00 / 1.00	

Question 8

You are given two binary random variables X and Y such that H(X|Y) = H(X). What can you conclude about X and Y?

Your Answer		Score	Explanation
◉ X and Y are dependent.	✖	0.00	
○ We cannot tell whether X and Y are dependent or not.			

○ X and Y are independent.	
Total	0.00 / 1.00

Question 9

The table below shows whether each word wi occurred in document D_i.

	w_1	w_2	w_3
D1	1	0	0
D2	1	1	1
D3	1	1	1
D4	1	0	0

Assuming that $P(X_{wi} = 1)$ and $P(X_{wi} = 1, X_{wj} = 1)$ are estimated by $Count(w_i)/4$ and $Count(w_i, w_j)/4$, respectively, which of the following is true?

Hint: The question can be answered without explicitly calculating the mutual information.

Your Answer		Score	Explanation
◉ $I(X_{w1}; X_{w2}) = 0$ and $I(X_{w2}; X_{w3}) = 1$	✔	1.00	
○ $I(X_{w1}; X_{w2}) = 1$ and $I(X_{w2}; X_{w3}) = 0$			
○ $I(X_{w1}; X_{w2}) = 1$ and $I(X_{w2}; X_{w3}) = 1$			
○ $I(X_{w1}; X_{w2}) = 1/2$ and $I(X_{w2}; X_{w3}) = 1/2$			
Total		1.00 / 1.00	

Question 10

Assume we want to mine the top syntagmatically related words to the word “computer" by ranking all the candidate words in descending order of their likelihood of being syntagmatically

related to "computer". Then, the top terms will be the same whether we rank each word w using conditional entropy $H(X_{computer}|X_w)$ or mutual information $I(X_{computer}; X_w)$.

Your Answer		Score	Explanation
True	✔	1.00	
False			
Total		1.00 / 1.00	