Feedback — Week 4 Quiz

Help Center

You submitted this quiz on **Sat 11 Jul 2015 11:40 AM PDT**. You got a score of **10.00** out of **10.00**.

Question 1

Sentiment classification can be treated as a text categorization problem.

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

Question 2

Suppose we want to perform rating prediction using 10 features where the output of the classifier is a discrete prediction $r \in \{1, 2, 3, 4, 5\}$. Then, the number of parameters that need to be estimated using ordinal logistic regression is **less** than the number of parameters that need to be estimated if we use 4 independent logistic regression classifiers to perform prediction.

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

Question 3

Given a set of restaurant reviews along with the overall numeric rating of every restaurant, you are asked to infer the ratings of each of the restaurants on cleanliness, taste, and value. Which

of the following methods is the **most suitable** to solve such an inference problem?

Your Answer		Score	Explanation
Sentiment Analysis			
Latent Aspect Rating Analysis	✔	1.00	
Contextual Text Mining			
Topic Modeling			
Total		1.00 / 1.00	

Question 4

Examine the objective function of NetPLSA in the lecture entitled **Contextual Text Mining: Mining Topics with Social Network Context**. Increasing λ will:

Your Answer		Score	Explanation
Not affect the topic coverage of neighbor nodes			
Make neighbor nodes have less similar topic coverage			
Make neighbor nodes have more similar topic coverage	✔	1.00	
Total		1.00 / 1.00	

Question 5

You are given an undirected citation network composed of papers $\{p_1, \ldots, p_n\}$ as nodes, where a link between papers p_i and p_j means that one of the papers cited the other. Suppose you want to use the given data to discover the topics (research areas) of the papers. Which of the following methods is expected to work best?

Hint: Papers that have a citation relationship are more likely to belong to the same research area.

Your Answer	Score	Explanation

Sentiment Analysis		
CPLSA		
PLSA		
NetPLSA	✔	1.00
Total		1.00 / 1.00

Question 6

You are given a collection of news articles along with their publishing dates and want to reveal which topics have attracted increasing attention in a certain time period. Which of the following methods is most suitable for this task?

Your Answer		Score	Explanation
CPLSA	✔	1.00	
Sentiment Analysis			
NetPLSA			
Total		1.00 / 1.00	

Question 7

Suppose we are performing Latent Aspect Rating Analysis where the number of aspect segments is K and the number of words in each aspect segment is M. What is the total number of parameters for term sentiment weights, i.e., the β values, that have to be estimated?

Your Answer		Score	Explanation
K			
MK	✔	1.00	
M			
$M+K$			
Total		1.00 / 1.00	

Question 8

Which of the following is true?

Your Answer		Score	Explanation
○ Ordinal logistic regression trains $k-1$ independent classifiers, k being the number of classes.			
◉ Different types of features, such as POS tags and word n-grams, can be combined when performing sentiment analysis.	✔	1.00	
○ The objective function of NetPLSA does **not** try to make neighbor nodes have similar topic coverage.			
Total		1.00 / 1.00	

Question 9

Imagine a company is interested in understanding any factors related to their fluctuating sales of a new product in the past year. They collected the companion text data including the consumer reviews of the product from multiple websites with time stamps in the past year and hope to gain potential insights from such text data. Which of the following text mining techniques would you recommend to them?

Your Answer		Score	Explanation
◉ Iterative topic modeling with time series supervision	✔	1.00	
○ Text clustering			
○ Contextual PLSA (CPLSA)			
Total		1.00 / 1.00	

Question 10

The US government implemented a new health care policy in year 2010. Suppose the government is interested in understanding the impact of such a policy and how the policy has affected what people talk about in social media. For this purpose, we can collect social media text data such as forum posts and tweets with time stamps before 2010 and after 2010. Which of the following text mining techniques is most suitable for such a text mining task?

Your Answer		Score	Explanation
◉ Contextual PLSA (CPLSA)	✔	1.00	
○ Iterative Topic Modeling with Time Series Supervision			
○ Text clustering			
Total		1.00 / 1.00	