Week 6 Quiz

10 试题

1 point

1。
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?

- Topic modeling
- Sentiment analysis
- Latent Aspect Rating Analysis
- Contextual text mining

1 point

2。
Examine the objective function of NetPLSA in the lecture entitled **Contextual Text Mining: Mining Topics with Social Network Context**. Increasing λ will:

- Not affect the topic coverage of neighbor nodes
- Make neighbor nodes have more similar topic coverage
- Make neighbor nodes have less similar topic coverage

1 point

3。
You are given an undirected citation network composed of papers {p1,...,pn} as nodes, where a link between papers pi and pj 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.

- ○ Sentiment analysis
- ○ PLSA
- ● NetPLSA
- ○ CPLSA

---

1 point

4。
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?

- ● CPLSA
- ○ NetPLSA
- ○ Sentiment analysis

---

1 point

5。

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?

- M+K
- MK
- M
- K

---

1 point

6。
Which of the following is true?

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

---

1 point

7。
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?

- [ ] Contextual PLSA (CPLSA)
- [ ] Text clustering
- [x] Iterative topic modeling with time series supervision

---

1 point

8。
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?

- [ ] Text clustering
- [x] Contextual PLSA (CPLSA)
- [ ] Iterative Topic Modeling with Time Series Supervision

---

1 point

9。
Context can be used to (check all that apply):

- [x] Annotate topics
- [x] Partition text

---

1 point

10。
Which of the following statement of CPLSA is NOT correct?

- [ ]

- [ ] The EM algorithm can be used for optimization.
- [x] It models the joint probability of text and context.
- [ ] It enables contextual text mining.
- [ ] CPLSA is an extension of PLSA.

我（**伟臣 沈**）了解提交不是我自己完成的作业 将永远不会通过此课程或导致我的 Coursera 帐号被关闭。
了解荣誉准则的更多信息

提交测试