Lesson 1 Quiz

测验, 4 个问题

1 point

1。

Which of the following statements are true? Select all that apply.

- [x] Clustering analysis is unsupervised learning since it does not require labeled training data.
- [ ] It is impossible to cluster objects in a data stream. We must have all the data objects that we need to cluster ready before clustering can be performed.
- [x] Clustering analysis has a wide range of applications in tasks such as data summarization, dynamic trend detection, multimedia analysis, and biological network analysis.
- [ ] When clustering, we want to put two dissimilar data objects into the same cluster.

---

1 point

2。

What are some common considerations and requirements for cluster analysis? Select all that apply.

- [ ] We must know the number of output clusters a *priori* for all clustering algorithms.
- [x] We need to be able to handle a mixture of different types of attributes (e.g.,numerical and categorical).
- [x] In order to perform cluster analysis, we need to have a similarity measure between data objects.

- [x] We need to consider how to incorporate user preference for cluster size and shape into the clustering algorithm.

1 point

3。
Which of the following statements are true? Select all that apply.

- [ ] There are no clustering algorithms that can handle time-series data since we always assume that the data points are temporally independent from each other.
- [x] Dimensionality reduction helps make high-dimensional clustering more feasible and scalable.
- [ ] Since cluster analysis is unsupervised learning, there's no way to incorporate user preference or guidance into the clustering process.
- [x] K-means is an example of a distance-based clustering method.

1 point

4。
If you need to choose between clustering and supervised learning for the following applications, for which ones would you choose clustering over supervised learning? Select all that apply.

- [ ] A hospital has collected the health measurements for a number of heathy people and the patients who are infected with disease X; it needs to determine whether a new person is infected with disease X after collecting his/her health data.
- [x] Find user communities in online social networks such as Facebook and Twitter.
- [ ]

Lesson 1 Quiz

测验, 4 个问题

Given a large number of emails, we already know whether each email is spam or not. Now we need to determine whether a newly incoming email is spam.

- [x] Given a large number of web pages, discover the latent topics discussed by those web pages.

- [x] 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

提交测验