

CSE519 Quiz 24

Total points 4/7

This quiz is set to analyze your basic understanding about the last lecture of the course. Be sure to answer all questions carefully because this will be graded.

The respondent's email address (**aditya.choudhary@stonybrook.edu**) was recorded on submission of this form.

Section score 4/7

✓ Which of the following is/are true for K-means clustering? * 1/1



It is difficult to train on categorical data since there is no mean for centroids



It doesn't need the number of clusters to be specified in advance



It cannot handle outliers



It can handle nested clusters



✗ Which of the following options suggest ways to handle the vertices with out degree equal to zero in PagaRank algorithm? 0/1

*

- ☒ We can make them randomly link to other vertices ✗
- ☒ We can make such a vertices link to themselves ✗
- ☐ We can create a new vertex (god vertex) which links to every page and make every other page link to it
- ☒ We can remove such vertices ✓

Correct answer

- ☒ We can remove such vertices
- ☒ We can create a new vertex (god vertex) which links to every page and make every other page link to it

✗ In which of the following types of graphs does the PageRank algorithm fail to converge? * 0/1

- ☐ An acyclic graph
- ☒ A bipartite graph ✓
- ☐ A complete bipartite graph
- ☐ A regular graph

Correct answer

- ☒ A bipartite graph
- ☒ A complete bipartite graph



✓ How many clusters do you see in the figure below? *

1/1



- ☐ 3
- ☒ The answer depends on our interpretation/senses
- ☐ 8
- ☐ 9

✓

✓ Which of the following can be the case(s) where clustering can be helpful? *

1/1

- ☐ It helps us in normalizing the data set
- ☒ It helps us in reducing the size of the dataset
- ☐ It helps us in solving a supervised learning problem
- ☒ It can help us form new hypothesis
- ☒ It can help us find outliers in a data set

✓

✓

✓



✗ Which of the following option(s) holds true in case of Grid Files? * 0/1

- ☒ Most of the cells are generally empty, making it inefficient ✓
- ☒ One of the limitations is that all points can be locally concentrated in just one cell making it hard to understand clusters ✓
- ☒ Things get difficult when the number of dimensions increases. ✓
- ☒ The idea is to cluster points by using a hash function such that $h(a)=h(b)$ ✗

Correct answer

- ☒ One of the limitations is that all points can be locally concentrated in just one cell making it hard to understand clusters
- ☒ Most of the cells are generally empty, making it inefficient
- ☒ Things get difficult when the number of dimensions increases.

✓ Which of the following option(s) is/are true regarding unsupervised learning? * 1/1

- ☒ Clustering is an example of unsupervised learning ✓
- ☒ PageRank is an example of unsupervised learning ✓
- ☐ In unsupervised learning, the target variables are present in the training data set.
- ☒ In unsupervised learning, the target variables are not present in the training data set. ✓

This form was created inside of Stony Brook University.

Google Forms

