Concept Quiz Over Week 9 Material

Due Dec 3 at 11:59pm **Points** 1 **Questions** 10
Available Nov 29 at 12am - Dec 3 at 11:59pm **Time Limit** None

Score for this survey: **1** out of 1
Submitted Nov 30 at 11:15pm
This attempt took 53 minutes.

Question 1

Which of the following are tasks in unsupervised learning?

ou Answered ☑ Clustering

ou Answered ☑ Density Estimation

ou Answered ☑ Dimensionality Reduction

ou Answered ☑ Classification

ou Answered ☑ Regression

Clustering, density estimation, and dimensionality reduction are all tasks in unsupervised learning. Regression and classification are examples of supervised learning because labels are used when training the models.

Question 2

There are two main categories of clustering algorithms -- hierarchical and partitional. Partitional clustering algorithms like k-means assign each datapoint to one of a finite set of clusters. Hierarchical algorithms like Hierarchical Agglomerative Clustering (HAC) iteratively merge or divide clusters in order to produce a hierarchy where each point belongs to a set of clusters.

ou Answered

- (●) True
- ( ) False

True.

Question 3

Centroid initialization is not important to k-means because k-means always converges to the global minima of sum-of-squared error.

- ( ) True

ou Answered

- (●) False

False. The k-mean algorithm is not guaranteed to reach the global minima of SSE, just a local one. As such, k-means is sensitive to which initialization.

Question 4

What is the stopping criteria for the k-means algorithm?

Your Answer:

if it is converged which means that it stops when the centroids of the clusters do not change significantly between iterations.

k-means is guaranteed to converge within a finite number of steps. Convergence occurs when the assignments don't change across an iteration.

Note that many implementations also consider a maximum number of iterations or the a percentage-assignments-changed criteria.

Question 5

The k-means algorithm is derived by attempting to minimize the sum-of-squared error between centroids and assigned datapoints using coordinate descent.

- True
- False

ou Answered

True. See the derivation in the class slides.

Question 6

Why is it a bad idea to choose *k* that minimizes SSE on a dataset?

Your Answer:

because k is largely heuristic as larger k produces lower SSE and unsupervised learning has no validation set because it has no labels

SSE decreases with larger *k* and thus doesn't tell us much about the quality of the clustering.

Question 7

Match the algorithm with its description

ou Answered

k-Means — A partitional clustering

ou Answered

Gaussian Mixture Model (GMM) — A probabilistic model

ou Answered

Hierarchical Agglomerative Clustering (HAC) — A clustering algorithm

ou Answered

Principal Component Analysis (PCA) — A dimensionality redu

k-Means - A partitional clustering algorithm that assigns each point to one cluster by iteratively updated centroids.

Gaussian Mixture Model (GMM) - A probabilistic model of data that assumes the data was generated by a set of Gaussian distributions. Parameters for these distributions are found via MLE using the Expectation-Maximization algorithm.

Hierarchical Agglomerative Clustering (HAC) - A clustering algorithm that builds a hierarchy of clusters by iteratively merging the closest clusters according to some "link function" that measures cluster similarity.

Principal Component Analysis (PCA) - A dimensionality reduction that finds linear subspaces that capture the variance of high dimensional data.

Question 8

Which of the following are true statements about Principal Component Analysis (PCA)?

ou Answered

- [x] PCA is derived by maximizing the variance of data after being projected to a linear subspace.

ou Answered

- [x] The solution to the PCA objective reduces to an Eigenvector decomposition.

- [ ] PCA works well for data with non-linear subspaces.

ou Answered

- [x] PCA transforms high dimensional data into low dimensional representations.

ou Answered

PCA ensures points belonging to different classes are well-separated in the projection.

False Statements:

PCA ensures points belonging to different classes are well-separated in the projection. -- PCA is an unsupervised approach and does not consider class labels.

PCA works well for data with non-linear subspaces. -- Standard PCA only considers linear subspaces.

Question 9

Complete-link HAC is able to generate very long clusters because it only considers the minimum distance between points in two clusters for merging.

ou Answered

- True
- False

False. Complete-link considers the maximum distance between points in two clusters. The question would describe Single-Link.

Question 10

The height of a joint in a Dendrogram reflects _____,

ou Answered

distance between the two merged clusters.

The height of a joint is the distance between two merged clusters.

Survey Score: **1** out of 1