

Ungraded: Introduction to Unsupervised Learning

Ungraded: Introduction to Unsupervised Learning
Practice Assignment • 10 min

Exit

Your grade: 100%

Your latest: 100% • Your highest: 100% • To pass you need at least 66%. We keep your highest score.

Next item →

1. Which statement about unsupervised algorithms is TRUE?

1 / 1 point

☐ Unsupervised algorithms are relevant when we have outcomes we are trying to predict.

☒ Unsupervised algorithms are relevant when we don't have the outcomes we are trying to predict and when we want to break down our data set into smaller groups.

Correct! They are helpful to find structures within our data set and when we want to partition our data set into smaller pieces for a better performance.

☐ Unsupervised algorithms are typically used to forecast time related patterns like stock market trends or sales forecasts.

☐ Unsupervised algorithms are relevant in cases that require explainability, for example comparing parameters from one model to another.

2. What is one of the real-world solutions to fix the problems of the curse dimensionality?

1 / 1 point

☐ Increase the size of the data set

☐ Use more computational power

☒ Reduce the dimension of the data set.

Correct! By doing dimensionality reduction we can improve both the performance and the interpretability of this grouping.

☐ Balance the classes of a data set

3. Which statement is a common use of Dimension Reduction in the real world?

1 / 1 point

☒ Image tracking

Correct! This is an example of reduce data to the primary factors.

☐ Explaining the relation between the amount of alcohol consumption and diabetes.

☐ Deep Learning

☐ Predicting whether a customer will return to a store to make a major purchase.

Like

Dislike

Report an issue

Ungraded: K Means Clustering

Your grade: 100%

Your latest: 100% • Your highest: 100% • To pass you need at least 66%. We keep your highest score.

Next item →

1. Which of the following statements best describes the iterative part of the K-means algorithm? 1 / 1 point
- ☐ The k-means algorithm assigns a number of clusters at random.
 - ☒ The k-means algorithm adjusts the centroids to the new mean of each cluster, and then it keeps repeating this process until no example is assigned to another cluster.
- Correct! Once the cluster centroids don't move anymore, it means that we have reached convergence.
- ☐ The k-means algorithm iteratively deletes outliers.
 - ☐ The k-means algorithm iteratively calculates the distance from each point to the centroid of each cluster.
2. Which statement describes better "the smarter initialization of K-mean clusters"? 1 / 1 point
- ☐ "Draw a line between the data points to create 2 big clusters."
 - ☐ "After we find our centroids, we calculate the distance between all our data points."
 - ☒ "Pick one random point, as initial point, and for the second point, instead of picking it randomly, we prioritize by assigning the probability of the distance."
- Correct! This one defines it and remember: The smarter initialization of K-mean clusters is called, K-means ++, and it helps to avoid getting stuck at these local optima. This is the default implementation of the K-means.
- ☐ "We start by having two centroids as far as possible between each other."
3. What happens with our second cluster centroid when we use the probability formula? 1 / 1 point
- ☐ When we use the probability formula, we put less weight on the points that are far away. So, our second cluster centroid is likely going to be closer.
 - ☒ When we use the probability formula, we put more weight on the points that are far away. So, our second cluster centroid is likely going to be more distant.
- Correct! This happens because it will take a larger proportion of the total distance square of all our points.
- ☐ When we use the probability formula, we put more weight on the lighter centroids, because it will take more computational power to draw our clusters. So, the second cluster centroid is likely going to be less distant.
 - ☐ When we use the probability formula, we put less weight on the points that are far away. So, our second cluster centroid is likely going to be more distant.

[Like](#) [Dislike](#) [Report an issue](#)

Graded: Module 1 Quiz

Your grade: 80%

Your latest: 80% • Your highest: 80% • To pass you need at least 60%. We keep your highest score.

Next item →

1. What is the implication of a small standard deviation of the clusters?

1 point

- ☒ A small standard deviation of the clusters defines the size of the clusters.
- ☐ The standard deviation of the cluster defines how tightly around each one of the centroids are. With a small standard deviation, the points will be closer to the centroids.
- ☐ The standard deviation of the cluster defines how tightly around each one of the centroids are. With a small standard deviation, we can't find any centroids.
- ☐ A small standard deviation of the clusters means that the centroids are not close enough to each other.

✖ Incorrect
Incorrect. Please review the lesson *K-means Initialization* lesson.

2. After we plot our elbow and we find the inflection point, what does that point indicate to us?

1 / 1 point

- ☒ The ideal number of clusters.
- ☐ The data points we need to form a cluster
- ☐ How we can reduce our number of clusters.
- ☐ Whether we need to remove outliers.

✔ Correct
Correct! You can find more information in the *K-means Initialization* lesson.

3. What is one of the most suitable ways to choose K when the number of clusters is unclear?

1 / 1 point

- ☐ You can start by choosing a random number of clusters.
- ☒ By evaluating Clustering performance such as Inertia and Distortion.
- ☐ By increasing the number of clusters calculating the square root.
- ☐ You can start by using a k nearest neighbor method.

✔ Correct
Correct! Both are measures of entropy part cluster. You can find more information in the *Selecting the Right Number of Clusters in K-Means* lesson.

4. Which statement describes correctly the use of distortion and inertia?

1 / 1 point

- ☐ When the sum of the points equals a prime number, use inertia, and when the sum of the points equals a pair number, use distortion.
- ☐ When we can calculate a number of clusters higher than 10, we use distortion, when we calculate a number of clusters smaller than 10, we use inertia.
- ☐ When outliers are a concern use inertia, otherwise use distortion.
- ☒ When the similarity of the points in the cluster are more important, you should use distortion, and if you are more concerned that clusters have similar numbers of points, then you should use inertia.

✔ Correct
Correct! This statement describes best how we can choose between distortion and inertia.

5. Which method is commonly used to select the right number of clusters?

1 / 1 point

- ☒ The elbow method.
- ☐ The ROC curve.
- ☐ The perfect Square Method
- ☐ The Sum of Square Method

✔ Correct
Correct! The method consists of plotting the interpreted variation as a function of the number of clusters, and selecting the elbow of the curve as the number of clusters to use.