

SEARCH



RESOURCES



CONCEPTS

- ✓ 11. Screencast: Solution
- ✓ 12. Video: How Does K-Means W...
- ✓ 13. Screencast + Text: How Does ...
- ✓ 14. Quiz: How Does K-Means Wor...
- ✓ 15. Video: Is that the Optimal Sol...
- ✓ 16. Video: Feature Scaling
- ✓ 17. Video: Feature Scaling Example
- ✓ 18. Notebook: Feature Scaling Ex...
- ✓ 19. Notebook: Feature Scaling
- ✓ 20. Screencast: Solution
- ✓ 21. Video: Outro
- ✓ 22. Text: Recap

Clustering Recap

We just covered a bunch of information! Here is a quick recap!

I. Clustering

You learned about clustering, a popular method for unsupervised machine learning. There are three ways to identify clusters in your dataset.

1. **Visual Inspection** of your data.
2. **Pre-conceived** ideas of the number of clusters.
3. **The elbow method**, which compares the average distance of each point to different numbers of centers.

II. K-Means

You saw the k-means algorithm for clustering data, which has 3 steps:

1. Randomly place k-centroids amongst your data.

Then repeat the following two steps until convergence (the centroids don't change).

2. Look at the distance from each centroid to each point. Assign each point to the nearest centroid.
3. Move the centroid to the center of the points assigned to it.

III. Concerns with K-Means

Finally, we discussed some concerns with the k-means algorithm. These concerns are:

1. **Concern:** The random placement of the centroids may lead to non-optimal clusters.

Solution: Run the algorithm multiple times and choose the centroids that create the minimum total distance of the points to the centroids.

2. **Concern:** Depending on the scale of the features, you may end up with different results.

Solution: Scale the features using Standardizing, which will create features with a standard deviation of 1 before running the k-means algorithm.



Mentor Help

Ask a mentor on our Q&A platform