

DAILY ASSESSMENT FORMAT

Course:	Machine learning with Python	Name:	Bindu.N.R
Link:	https://cognitiveclass.ai/courses	USN:	4AL17EC101
Org By :	IBM	Semester & Section:	6-B
Github Repository:	bindunr-python	Date:	19/06/2020

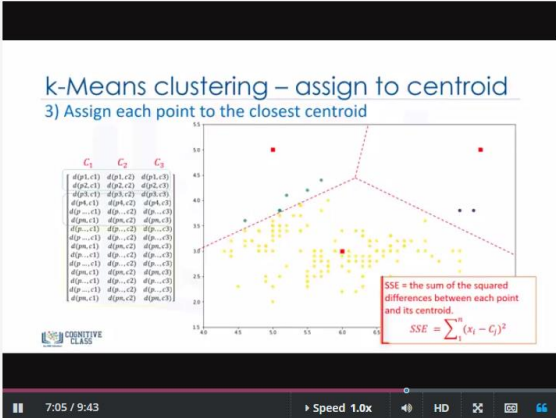
Progress on 19-06-2020

• Topic Completed Today

K-Means Clustering (9:43)

Bookmark this page

K-Means Clustering (9:43)



k-Means clustering – assign to centroid

3) Assign each point to the closest centroid

SSE = the sum of the squared differences between each point and its centroid.

$$SSE = \sum_{i=1}^n (x_i - C_j)^2$$

nearest centroid to data points.

Finding the closest centroids for each data point, we assign each data point to that cluster.

In other words, all the customers will fall to a cluster, based on their distance from centroids. We can easily say that it does not result in good clusters, because the centroids were chosen randomly from the first.

Indeed, the model would have a high error. Here, error is the total distance of each point from its centroid. It can be shown as within-cluster sum of squares

error. Intuitively, we try to reduce this error. It means we should shape clusters in such a way that the total distance of all members of a cluster from its centroid be minimized. Now, the question is, "How we can turn it into better clusters, with less error?"

Okay, we move centroids. In the next step, each cluster center will be updated to be the mean for data points in its cluster. Indeed, each centroid moves

DBSCAN Clustering (6:57)
Bookmark this page

DBSCAN Clustering (6:57)

So, what is a core point? A data point is a core point if, within R-neighborhood of the point, there are at least M points. For example, as there are 6 points in the 2-centimeter neighbor of the red point, we mark this point as a core point.

Ok, what happens if it's NOT a core point? Let's look at another point. Is this point a core point? No.

As you can see, there are only 5 points in this neighborhood, including the yellow point.

So, what kind of point is this one? In fact, it is a "border" point.

What is a border point? A data point is a BORDER point if:

- Its neighborhood contains less than M data points, or
- It is reachable from some core point. Here, Reachability means it is within R-distance from a core point. It means that even though the yellow point is within the 2-centimeter neighborhood of the red point, it is not by itself a core point because it does not have at least 6 points

• Progress Report

ML0101ENV3 Progress | Cognitive

cognitiveclass.ai
Cognitive Class: ML0101ENV3
Machine Learning with Python
Learning Paths
Courses
Help
Bindu_N_R

Course
Discussion
Wiki
Resources
Progress

Course Progress for Student 'Bindu_N_R' (bindu6433@gmail.com)

Component	Progress (%)
RQ 01	100%
RQ 02	100%
RQ 03	100%
RQ 04	100%
RQ 05	0%
RQ Avg	~80%
Final	0%
Total	40%

Welcome!

Welcome! (3:15)
No problem scores in this section