Learning Analytics Tools

Ramkumar Rajendran

Educational Technology

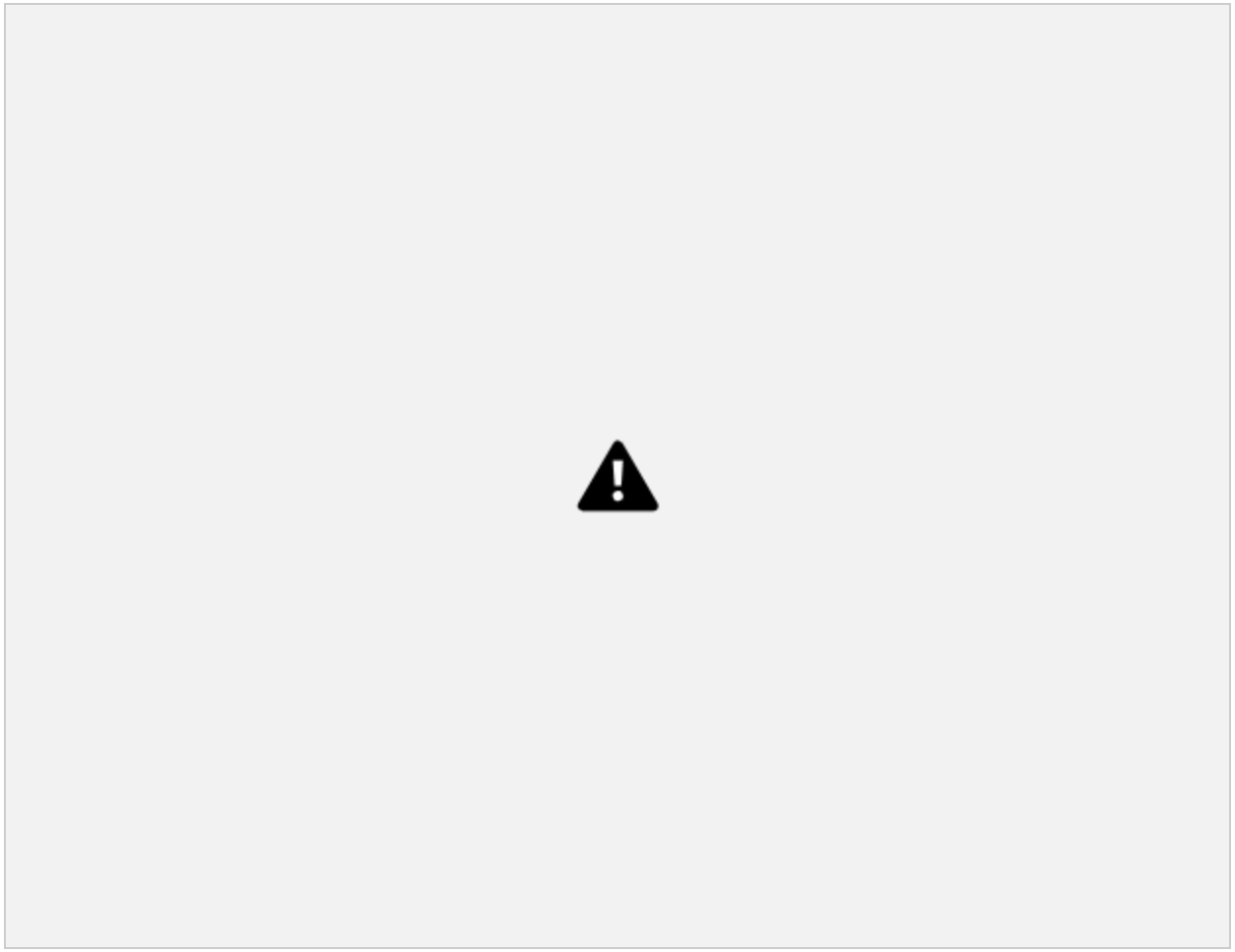
Indian Institute of Technology, Bombay

Lecture 36

Diagnostic Analytics - Clustering

Welcome back to learning analytics tools course. In this week we will continue discussing diagnostics analytics and in that, especially clustering. So, what is clustering? Clustering is grouping similar behaviour, the items which have similar behaviour in a cluster.

(Refer Slide Time: 00:34)

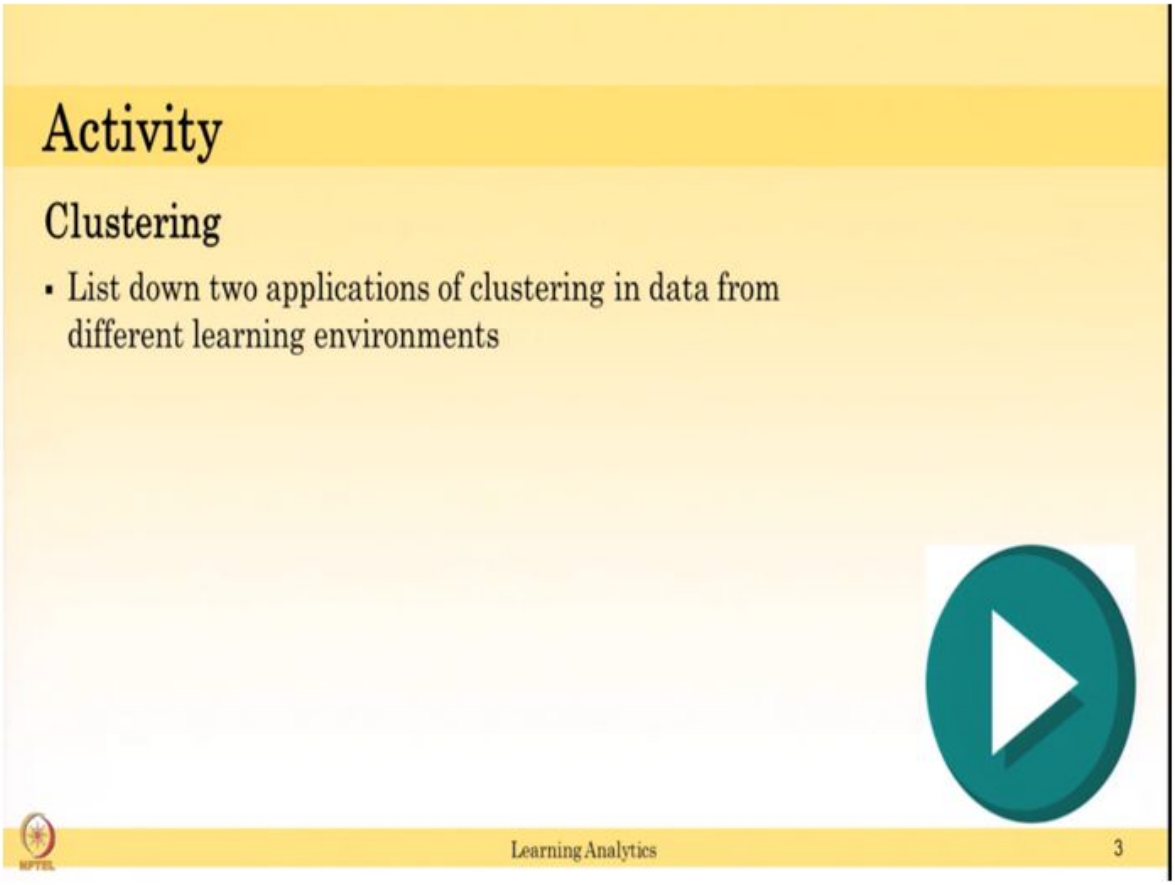


For example, if there are a hundred students taking the class and based on the behaviour, their background profile, you can group them - so that is called clustering. You are grouping the students based on their similarities in behaviour. Especially this behaviour is different from the other group.

So, if you have two groups, the student in group 1 will have different behaviour from group 2 that is how you create the clusters. In learning environments, we can use the student's interaction with the system, the numerical data to create clusters. So, you have to convert the data into frequency, the time spent those kinds of information and that can be used to cluster. Also, we can use the action sequences to clusters, but that is not discussed in this course.

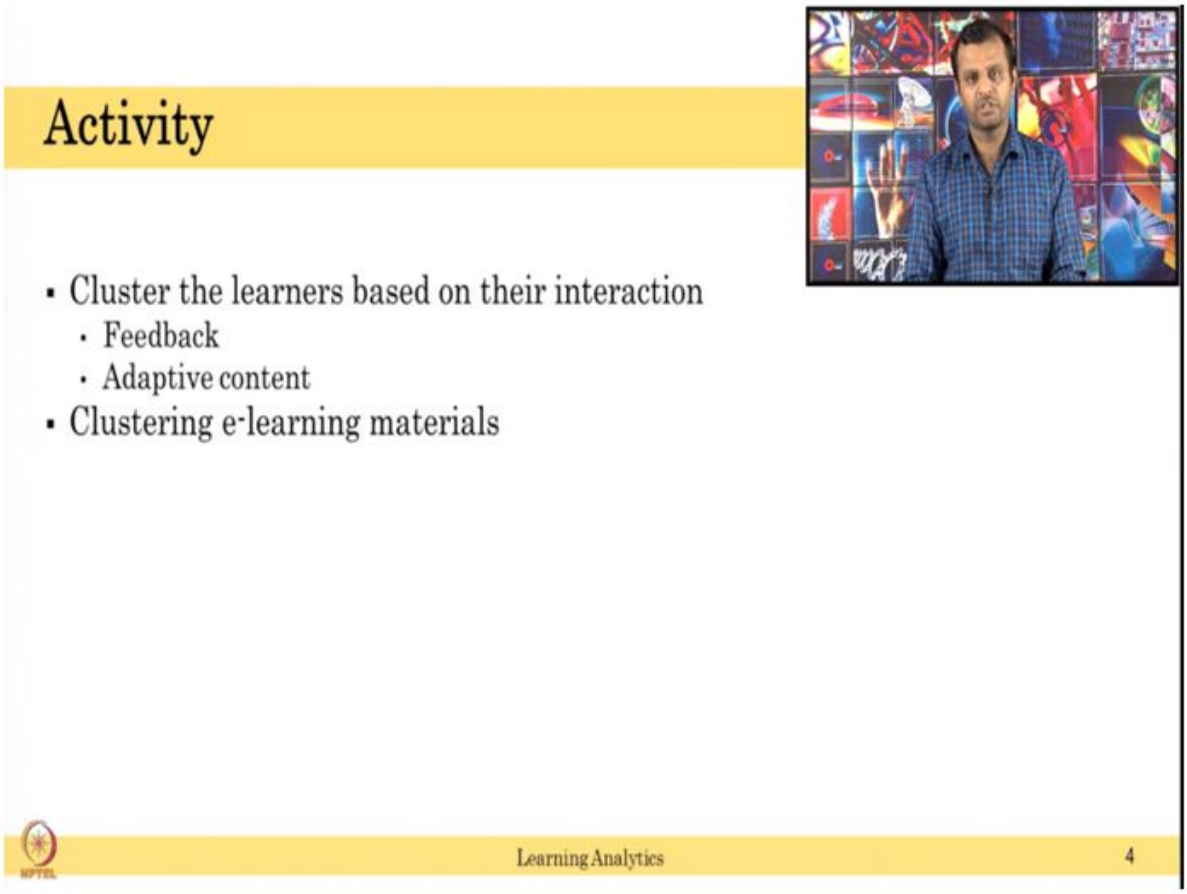
So, let us see how to use interaction behaviour in the terms of numerical values can be clustered. Before that let us discuss what is clustering and what are the famous clustering algorithms. There are two major categories of clustering algorithms. There are more, but these two are major. The one is based on Centroid, you might have seen the clusters creating, using centroids. For example, in this picture, you see there are three clusters. This picture is credited from Wikipedia so here is the link. In this week we will see an example algorithm for one centroid base and one example algorithm for the connectivity-based.

(Refer Slide Time: 02:04)



So, before we jump into what is clustering and the algorithms in clustering in the two main categories, I want you to think of why we need clustering for data from learning environments? Now, you can choose any learning environment like a classroom environment or MOOC, or TELE, anything. So, choose one of the learning environment and list down at least two applications of where clustering will be used, why we have to do clustering. The answer should not be - " I want to do cluster-based interaction with MOOC", but why? What is the information I want to use after creating clusters? Please pause this video after you list it down, resume to continue.

(Refer Slide Time: 02:49)



So, the clustering of learners based on the interaction, for example, students interacting with a MOOC or TELE, based on the interaction you can cluster then you can use that provide feedback. So instead of providing the same feedback to all learners, you know this learner vary by 3 or 4 clusters and you can give personalized feedback to certain cluster group.

You know some students might be spending much time in reading materials and taking quiz instead of watching any videos. For those learners, you might give, feedback saying that - why cannot you go and take a look at the videos. So, those kinds of feedback can go to those users when they were not able to give the answers to the quiz questions.

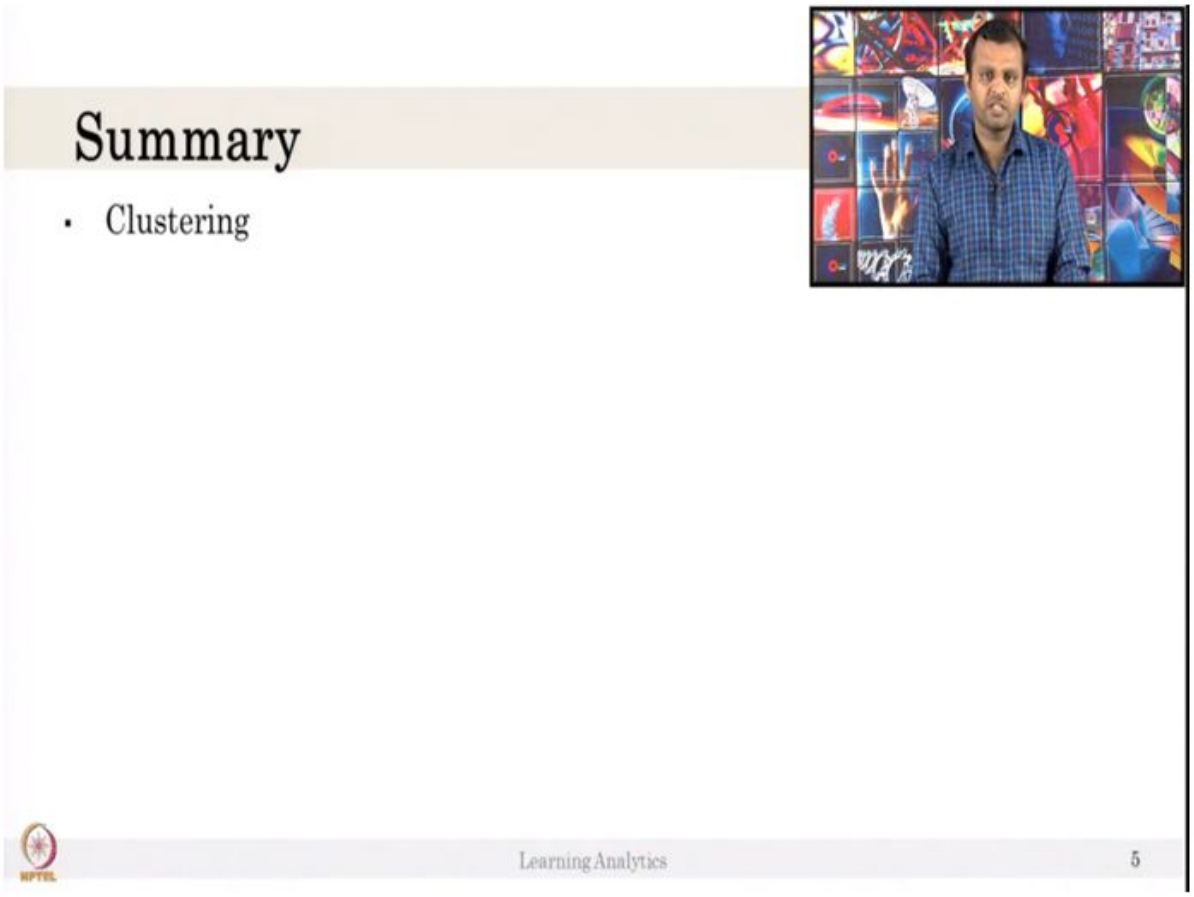
For some learners who are not really interested in reading or doing any simulation, they just watch videos and try to answer the question, for them you can say why cannot you try simulators or some read materials. Or some learners who don't have anything just directly go to answering

questions, try to answer them and complete the week's assignment and go for next week. For them, you can give a different kind of feedback.

So, you can provide feedback if you cluster the students based on the behaviour, the interaction with the system. Also, we can use this clustering to provide adaptive content based on the student's interaction, you might give a hint for a new video or new course material - that is also possible.

And there is another application of clustering for e-learning materials. There is a lot of e-learning content available, we can go and crawl all the e-Learning content and curate it and you can automatically classify them into topic wise or concept-wise. So, if you have those things the system can automatically pick their learning content based on the topic you are interested or concept you are interested in. So, that is also possible. So, clustering can be used to cluster the student's behaviour based on the interaction with the system and also content can be clustered or other thing can be clustered in the system.

(Refer Slide Time: 05:03)



So, in this video, we saw what is clustering and why it is needed for learning environments. In our next video, we will talk about one type of clustering algorithm and discuss detail. Thank you