

Introduction to Learning Analytics
Prof. Ramkumar Rajendran
Interdisciplinary Programme in Educational Technology
Indian Institute of Technology, Bombay

Lecture 14
Predictive Learning Analytics

In this LED, we will talk about Predictive Learning Analytics. Although we mentioned that in week 4 only we will see predictive learning analytics. We realized that in week 4 there are too many activities and we will be talking about two different models. So, we thought of moving the last LED to about predictive learning analytics in week 3. What is predictive analytics? We briefly touched about this in a week one.

(Refer Slide Time: 00:41)



Predictive Analytics

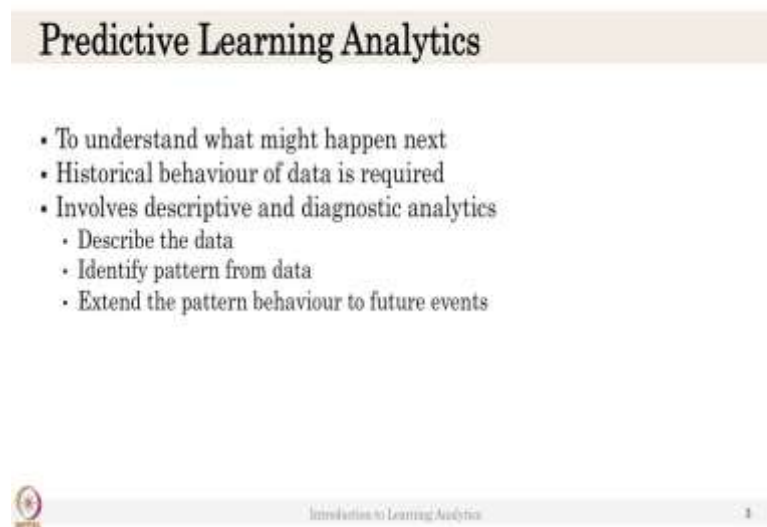
- Analyze current and historical data to predict future events
- Machine Learning and Data Mining Tools
- Examples:
 - Weather prediction
 - Fraud Detection
 - Anti-Spammers
- Domain
 - Finance
 - Health Sector
 - Telecom
 - others

Introduction to Learning Analytics

Predictive analytics that is to analyze the current and historical data to predict the future events. So, in this predictive analytics, we will involve machine learning and data mining tools in their learning analytics. Predictive analytics is very common, it is updating different domains; for example, finance, health sector, telecom and others other domains. For example, the weather prediction is by done by predictive analytics. It collects a lot of historical data over the period of last year, last 10 years then it predicts what will happen next, whether it will be sunny or rainy and next day. Also for fraud deduction in a bank transactions can be done using predictive analytics.


Another example is in a day to day life is having anti spammers in email. In an email, we collect historical data about what of the words the spam email contains, then we create a model, then we will predict whether the next email is spam or not. Although it is not predicting the future event, but it predicts the incoming email before the learner or the user detects it is a spam.

(Refer Slide Time: 01:49)



Predictive Learning Analytics

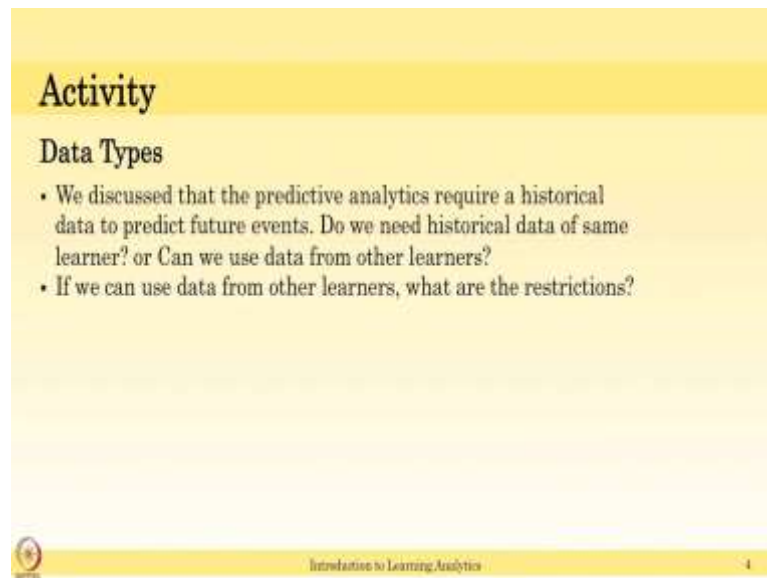
- To understand what might happen next
- Historical behaviour of data is required
- Involves descriptive and diagnostic analytics
 - Describe the data
 - Identify pattern from data
 - Extend the pattern behaviour to future events

 Introduction to Learning Analytics 3

Predictive learning analytics. So, what is predictive analytics in for learning. In predictive learning analytics, it is to understand what might happen next in the students' interaction to the students, where the students able to answer the question correct or student needs more help in the next slide or the student we need will pass the exam or student needs hint in the next question, those things will be predicted.

In order to do the predictive analytics or to predict the future event of the students or learners' interaction, we need to record the historical behavior of the student. From the historical data we can create a model then extrapolate that model to predict the future events. Predictive analytics involves descriptive and diagnostic analytics also; for example, first we need to describe the data what are the data, then we need to identify the pattern that is the descriptive analytics or diagnostic analytics, then we need to extend or extrapolate that model to predict the future events. Let us see the activity.

(Refer Slide Time: 02:51)



Activity

Data Types

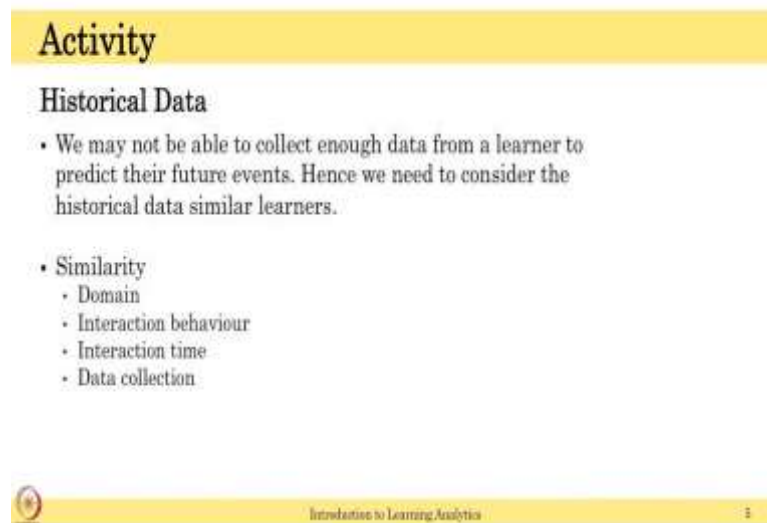
- We discussed that the predictive analytics require a historical data to predict future events. Do we need historical data of same learner? or Can we use data from other learners?
- If we can use data from other learners, what are the restrictions?

Introduction to Learning Analytics 4

In this activity, we discussed that the predictive analytics needs historical data to predict the future events; however, do we need historical data for same learner or we can use data from other learner. For example, as I mentioned in weather prediction we realized that in order to predict the temperature tomorrow or in order to predict the climate tomorrow, we need to understand what was the climate in last 1 month or 1 year or last 10 years in different seasons, but it is about the same climate and same area right. So, in learning analytics in order to predict the future events, do we need a data from the same student or can we use the data from other learners.

If you say we can use the data from other learners what are the restrictions, what are the assumptions. Please think about this question, write down your answers. After writing down the answer we can resume this video.

(Refer Slide Time: 03:55)



Activity

Historical Data

- We may not be able to collect enough data from a learner to predict their future events. Hence we need to consider the historical data similar learners.
- Similarity
 - Domain
 - Interaction behaviour
 - Interaction time
 - Data collection

Introduction to Learning Analytics

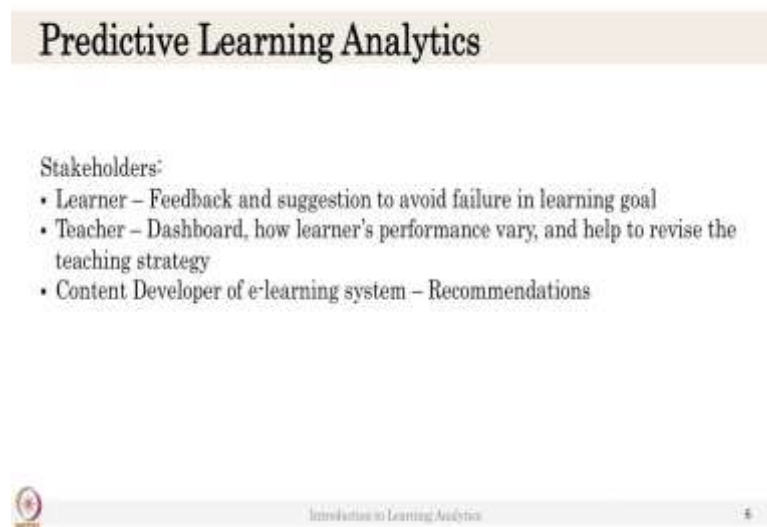
You might have realized that we may not be able to collect enough data of a student in order to predict the future events, because a student will be newly starting the learning environment. In order to give a recommendation or prediction based on that you might need to wait for a 1 week or a 1 month, so it is not possible right. We need to give a personalization or adaptive feedback hint messages immediately after they join the learning environment program.

In order to do that, we need to consider the data from other students or from the similar learning environment. So, if I say similar data, what are the similar data, what are the restrictions in order to collect the similar data. For example, if we see domain. Is the student the data we collected is in the same domain, is in a math or the same concept we need to think about it. Also is the same learning environment, we cannot apply the data collected from learning environment x on learning environment y. So, we need to collect data from the same learning environment. Also the interaction behavior should be similar. For example, in learning environment x first question you might have different menus or different interaction be given along for the students, in second environment you might have a additional features that is not allowed.

You need to collect data which has a similar domain, similar interaction behavior also in similar learning environment if you are talking about technology enhanced learning environment. For example, if it is a MOOC or a classroom environment we need to

provide a personalized learning content based on the similar user data, taken this course in last semester. In a MOOC, the student will be taking the course for first time, but we have a same set of students or similar kind of students who have taken the course in MOOC. All the data we collected from the MOOC platform will be applied used to predict the future events of the current students. So, we can collect data from the students who were taken this course in the previous semester or previous year or previous events and create a model and use that to predict the future events of the current students.


(Refer Slide Time: 06:07)



Predictive Learning Analytics

Stakeholders:

- Learner – Feedback and suggestion to avoid failure in learning goal
- Teacher – Dashboard, how learner's performance vary, and help to revise the teaching strategy
- Content Developer of e-learning system – Recommendations

 Introduction to Learning Analytics 5

So, who are the stakeholders and why were you doing this predictive learning analytics, what is this, what is in for stakeholders for predictive learning analytics. For learner it is very important, because based on this predictive learning analytics model we are going to provide the suggestions or feedback to improve the learning. So, it is very important for learner as a stakeholder. The second stakeholder we know is teacher; for teacher it is very important to know what will happen to student next; whether student will be able to solve that particular problem or will he pass the exam or the student will be able to complete the given task. In order to help if they are not able to complete task, the teacher has to know when to intervene and what to intervene.

So, the predictive analytics should give a insight, saying that the particular student in your class might have a problem, because his learning activities or interactions with the systems or environment is not enough. It says that the student might not be able to complete

the problem given to them. For example, a teacher can know that in the class out of 10 students 3 or 4 students may not submit the assignment in next month. So, what can we do? So, teacher can think of other different strategies and make motivation to students and ask them to submit assignments on time or they can extend their assignment deadline. So, similar activities are helpful for teacher.

So, predictive learning analytics is very important for teachers as the stakeholders, to understand which student will complete the course, also to understand what is the problem in the current learning content. We talked about other stakeholder; that is institutes, academics, academic analytics, stakeholders or the institute heads or district heads. They also will be interested in about predictive learning analytics. For example, what will happen, which course will be taken by most students in next semester, which course will not be taken by most students in this college. So, they can decide whether to run a course next year or not.

The other important stakeholder in a learning analytics is content developers of the e-learning systems. If the learning analytics or predictive analytics says that this students will have a problem, if their interaction behavior is kind of certain pattern then you might need to give a this particular recommendations. These information will help the students to create the automatic agents. In a e-learning systems, we can have a agents which can interact with students based on the learner behavior. In order to design this agent, modeling the students' behavior from the historical data is very important.