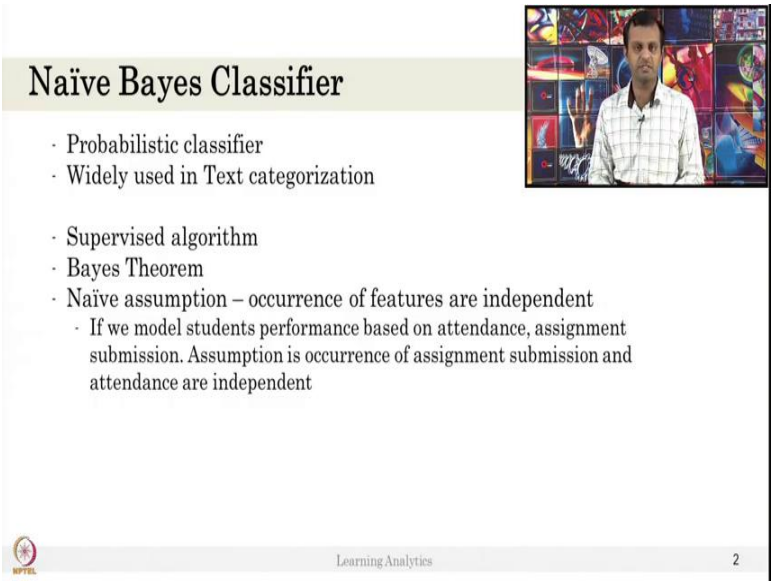
Learning Analytics Tools
Professor Ramkumar Rajendran
Department of Educational Technology
Indian Institute of Technology, Bombay
Lecture 45
Predictive Analysis - 2

Welcome back to Learning Analytics Tools course. This week will continue on to predictive analytics.

(Refer Slide Time: 0:29)

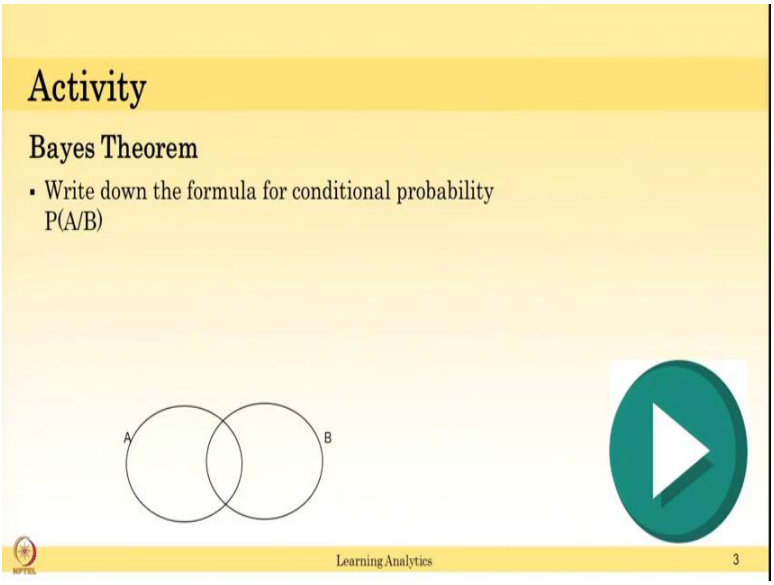


So we will discuss Naive Bayes classifier and classification algorithm in this week. So Naive Bayes classifier is a probabilistic classifier, the output will be probability of getting passes in examination with probability 0.7 or 0.8 or 0.2 or 0.3. And it is very widely used in text categorization.

It is one of the supervised algorithm because we have to give the label also, not only the X, we also have to give the y. It like a classification, whether you have to predict pass or fail, or you have to predict whether the student will continue in the course or not, something like that. So it is based on Bayes theorem that is why it is called Naïve Bayes. So what is Naïve? That is a naive assumption here that is very important in the Bayes Theorem.

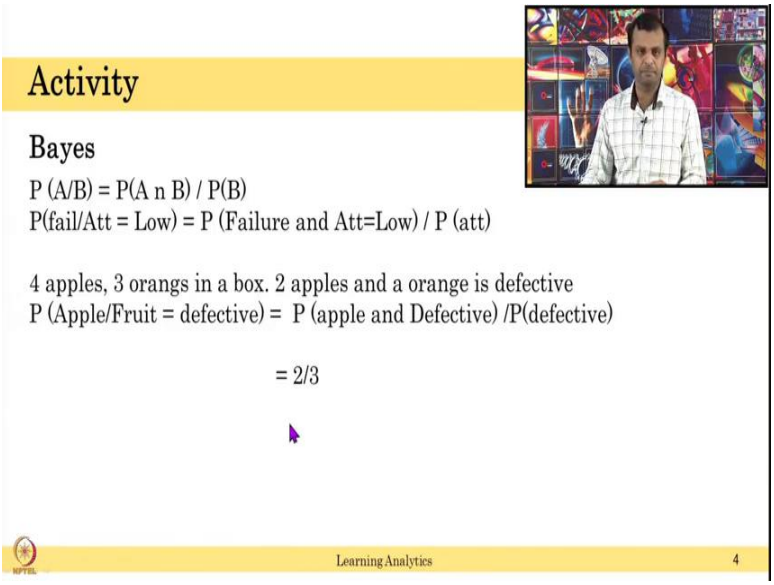
Bayes Theorem is very simple, you might have used it in a probability classes, but this Naïve Bayes classifier assumes that the features occurring in this are independent. Suppose if you try to model the student's performance based on their attendance, assignment or submission or some engagement in the class, there is something you have to assume that these occurrences are independent, although it is not possible. As you know, the student has to be attending the class to improve the engagement, but the Naïve Bayes makes this assumption. So you can think of other features like in MOOC a student logging in, student interacting with the forum, student watching video these are independent because not every student is going to interact with forum or something like that.

(Refer Slide Time: 2:12)



So let us start with the activity before we jump into what is Naïve Bayes theorem. Let us see what is conditional probability. Just to brush up your memory so there are two dependent events, a shown in a diagram, what is the probability of A given B that is what is probability of A occurring given B already occurred. Can you write down the formula for conditional probability.

(Refer Slide Time: 2:52)



So Bayes theorem starts with this conditional theory that is probability of A given B is probability of A and B occurring together, given probability of B occurred like probability of complete B.

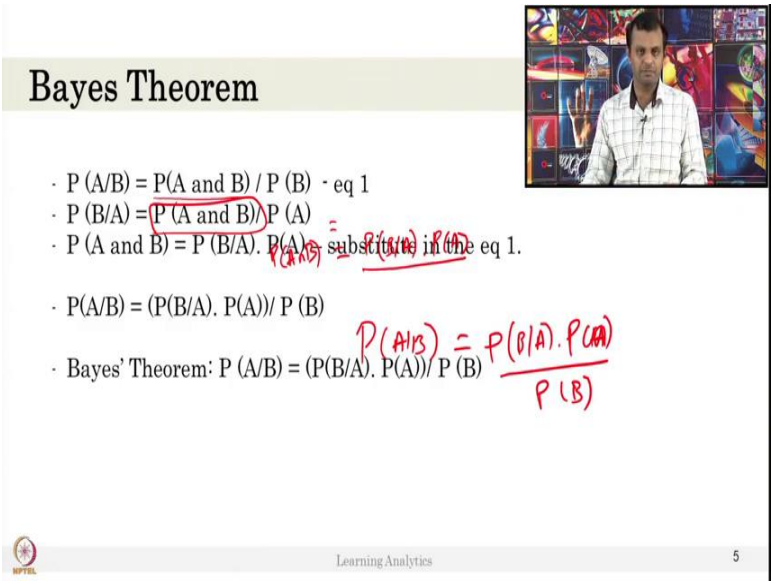
For example, if I want to know probability of student who fail the exam if attention is low. So that means probability of failure and low attention has to occur that is the student who are having low attention also failed an exam, there are lot of other that might have not failed given that the probability of attention is low. So let us see, this in a detailed about conditional probability.

Let look at the example, this is standard probability, there is a box which contains four apples and three oranges, out of these two apples are defective and one orange is defective. Then what is the probability if we pick a defective fruit that fruit be the apple, to calculate it you can use conditional probability, also since this is simple problem, you can answer directly. Let us see what is the probability of choosing apple and it is being defective and probability of defective fruit.

For this problem, I am just going to completely ignore everything I am just going to simply apply, defective apple are two and the total defective fruit is three that is there are three fruits

available defective, how many of them can be defective apples, two out of three. So it is very simple probability.

(Refer Slide Time: 4:49)

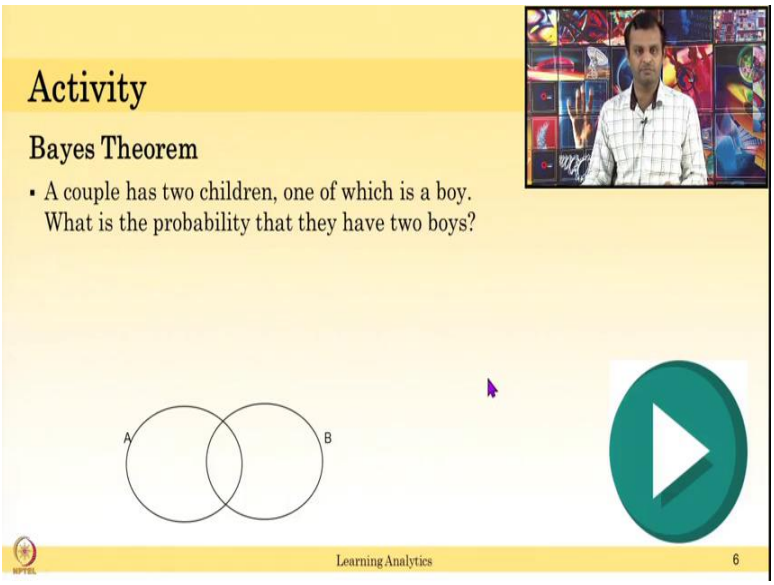


Let us look at the Bayes theorem from the conditional probability. Probability of A given B is probability of A and B wrt probability of B, similarly probability of B by A, probability of A and B wrt probability of A.

If I want to replace this term with the other terms, suppose consider this probability of B by A, this term can give probability of B by A into probability of A, this is equal to probability of of A and B.

So you can replace this particular statement by this value, so that is probability of A by B equal to this P of A and B, and this can be replaced by this particular term probability of B by A into probability of A and the whole things can be divided by this P of B, that is very simple. This is called the Bayes theorem. So Bayes theorem is very simple, it is just two conditional probability. It is probability of that A will occur given B already occurred. This simple Bayes theorem can be derived from the conditional probability.

(Refer Slide Time: 7:12)

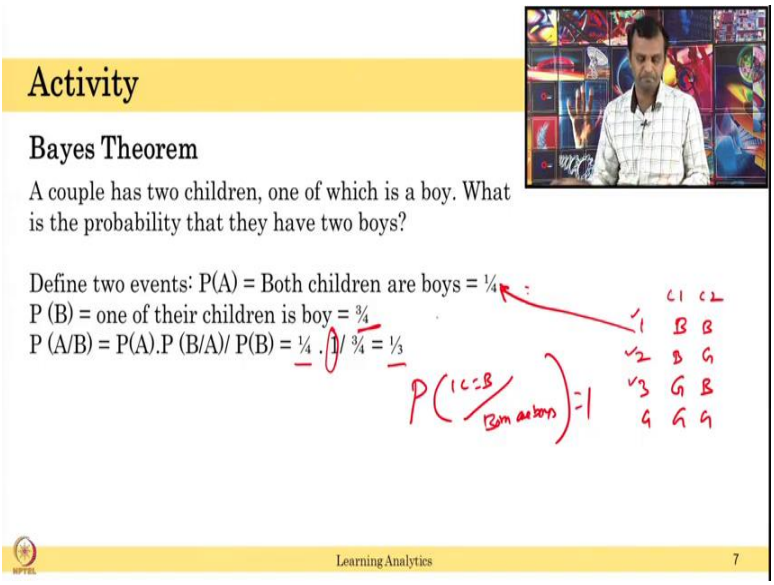


So now even this particular question also can be solved without using Bayes theorem but I want to apply a Bayes theorem in this.

Consider a couple has two children, they are not twins. One of whom is boy, what is the probability that they have two boys? They have two children boys or girls, anything is possible, so boys, boys, boys, girls, girls, boys, boys, girls something like that. Consider one of which is boys say first child is boy. What is the probability that they have two boys? Second child also boy, something like that.

Please write down the answer after writing it down, resume to continue.

(Refer Slide Time: 7:58)



So, what is the probability of both children are boys given this one children is boy. So that is probability of A given B, you can simply apply the Bayes theorem, probability of A into probability of B by A into, so how these numbers came? This numbers are very simple, right? First possibility is both are boys. Second possible outcome is, one is boy, first children is boy, second child is girl, third is first child is girl second child is boy, fourth is both child's are girl. So the outcome the given space is only this four events possible. So given there are four possible outcomes in that being both are boys is one, that is why it is 1 by 4.

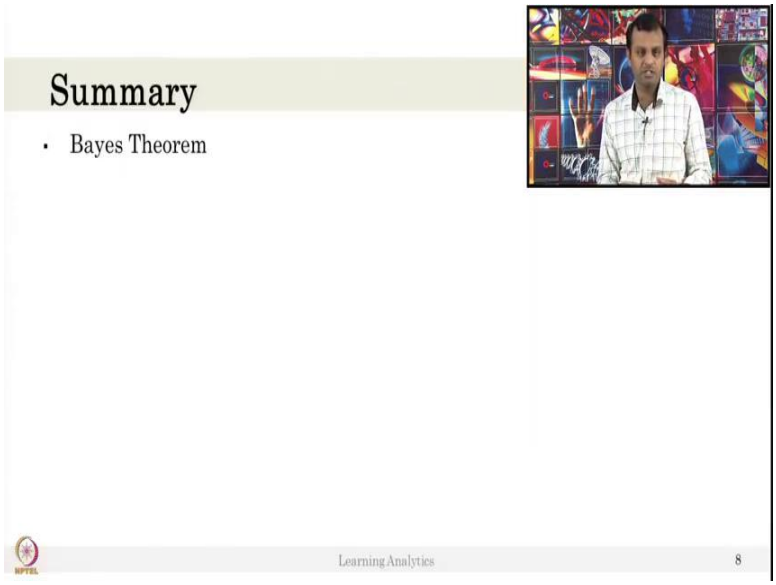
Also probability of being one of their child being boy it is like here at least one child is boy, this one child is a boy and here also one child is boy. At least one child is boy out of four the three times that one child can be boy, it is simple. So that is answer here.So one, two, three, three chance that one at least one of them will be boy.

So now we have to compute this, so probability of A, A is you know probability of A is 1 by 4, true. The tricky part you know the tricky part is this one. Probability of B given A what is B? B is probability of one of them being boy given probability of both are children are boys. So the probability is 1

But I took the simple problems to explain the theorems here because that you can understand step by step. So you can solve without applying the Bayes theorem that is fine but this is how

you can apply Bayes theorem on this kind of problems. Please go ahead and solve more problems related to this probability based problems in the Internet and understand how it works, that there is a lot to understand Naïve Bayes theorem.

(Refer Slide Time: 11:23)



Naïve Bayes theorem is just this Bayes theorem only. So we saw what is Bayes theorem not a Naïve Bayes classifier in this slide, but it is a Bayes theorem how it is constructed from the conditional probability and I request you to go to solve much more probability problems using Bayes theorem. Thank you.