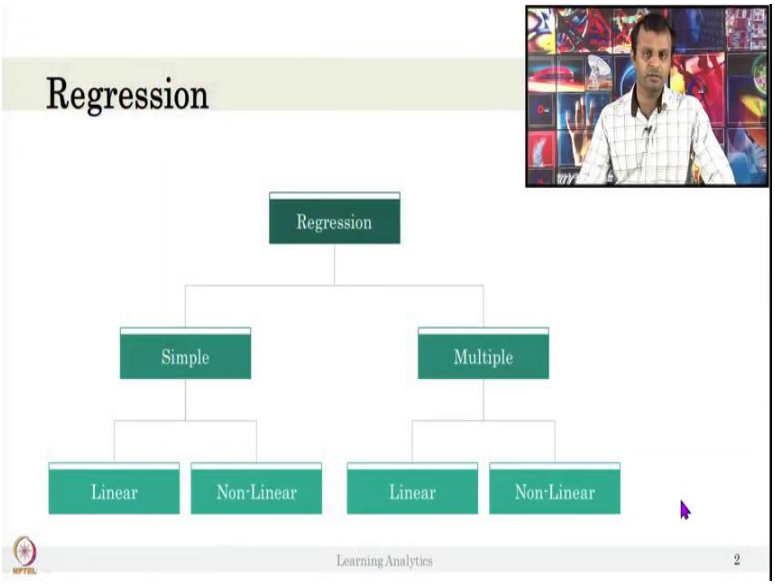
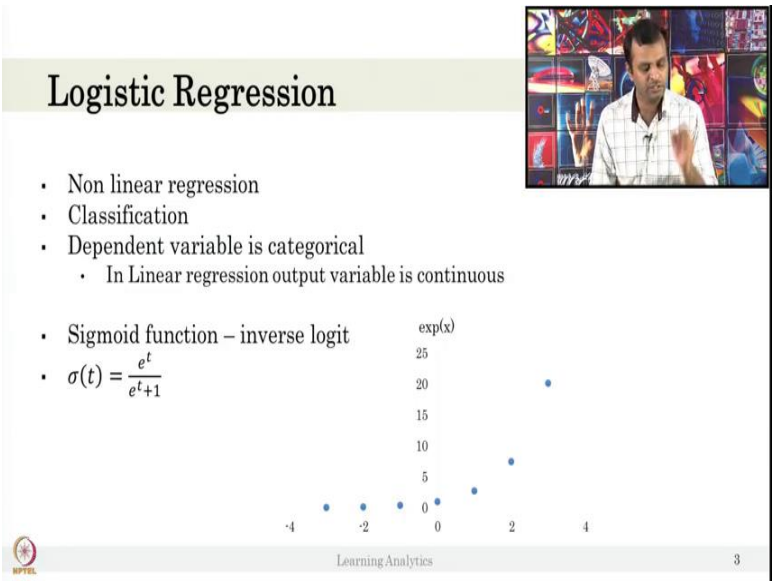
Learning Analytics Tools
Professor Ramkumar Rajendran
Department of Educational Technology
Indian Institute of Technology, Bombay
Lecture 43
Logistic Regression

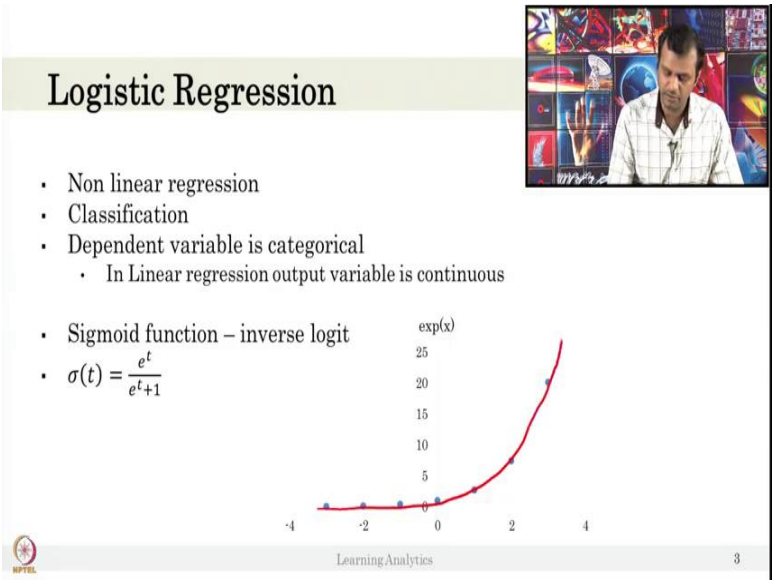
(Refer Slide Time: 0:21)



In this video, we will talk about logistic regression. This is a non-linear regression. There are many ways to do the non-linear regression like polynomial, multiple polynomial square, x square, x cube. But let us see the logistic regression in this video.

(Refer Slide Time: 0:32)





It is a non-linear regression, and it is actually not regression for continuity, but for the classification that is student pass the exam or not pass the exam.

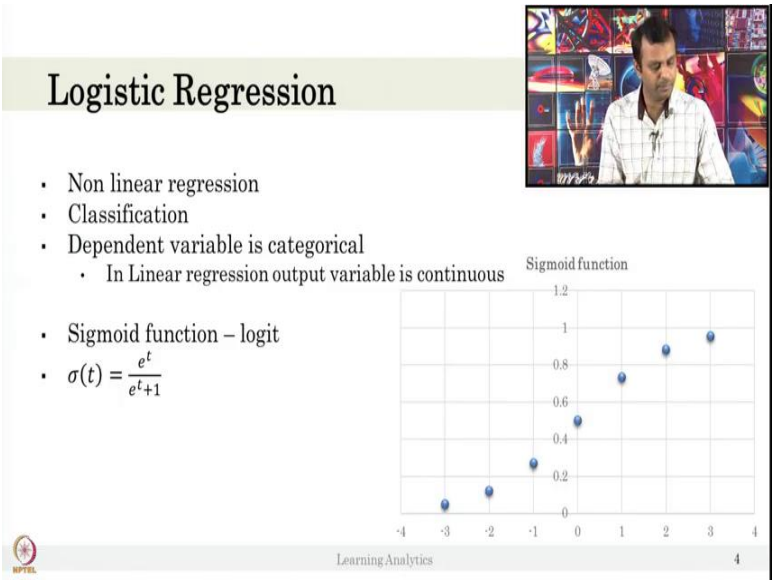
In a linear regression if you substitute the marks and attendance, you might be able to predict the students' performance and the final score say, 75.5, 79.6. But in logistic situation, it is not the case, we classify the student will pass the exam or not pass the exam. Now we saw the difference between classification and regression.

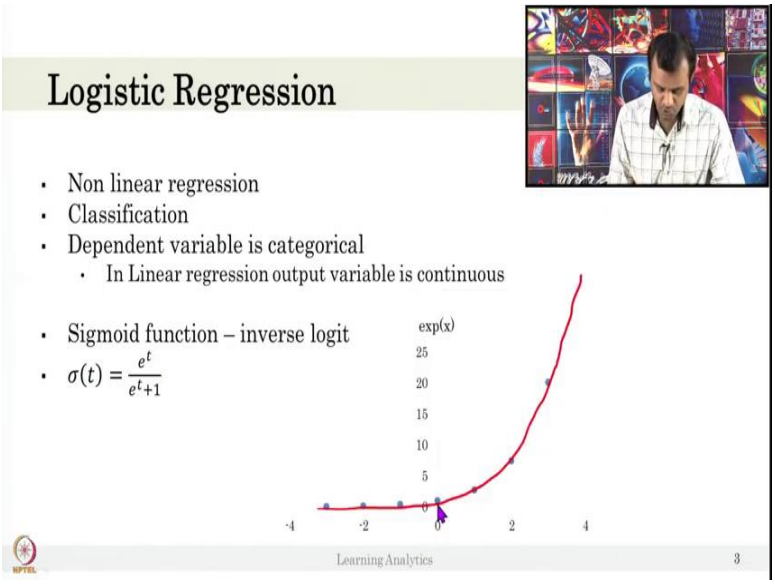
Basically, logistic regression is a Sigmoid function. What is Sigmoid function ? It is, exponential of t divide by 1 plus exponential of t. You are aware of exponential function and its variation.

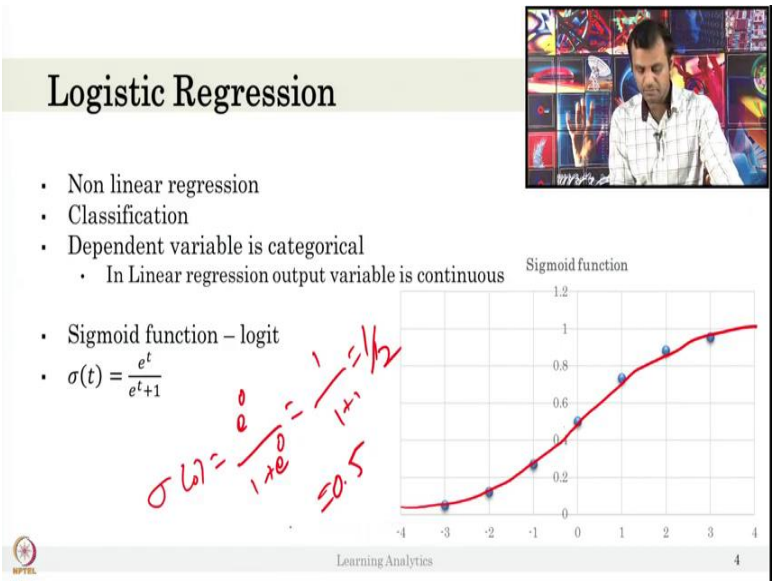
So if you have this exponential curve in this particular formula that is called Sigmoid function. Somebody know the log(odds)or logit might able to connect. This actually to inverse logic function or log odd.

Now in logit or log odd you have to consider not only the probability of getting the true value, also the non-true value, and it is actually one minus probability of true. So we do not want to go detail about logistic regression and logic or log odds do not worry about that.

(Refer Slide Time: 2:44)





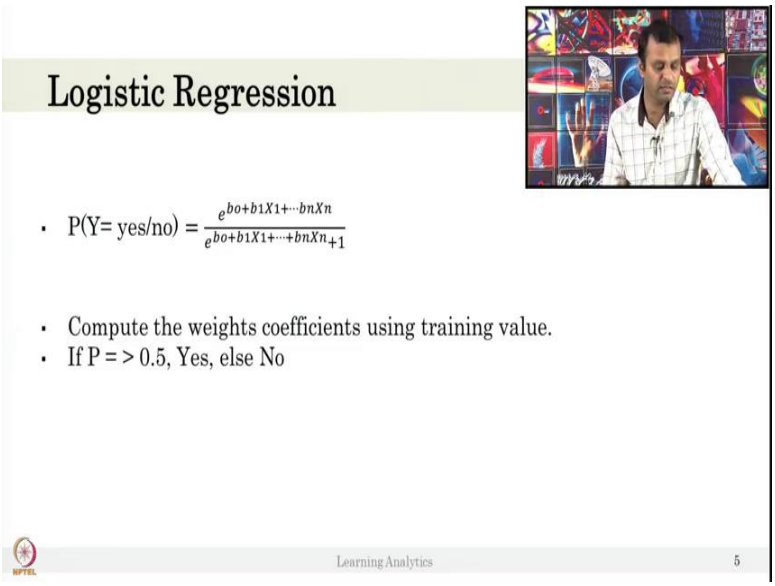


Let us try to understand what is logistic regression and what is a function. How the classification differ from regression?

So if I compute in this Sigmoid function value, consider the Sigmoid of 0 equal to e power 0 divided by 1 plus e power 0, 1 divided by 1 plus 1, 1 by 2 it is 0.5, so 0 will be 0.5. It value will be somewhere here.

This will be a 0, this maximum will be 1, which will never be 1 basically.

(Refer Slide Time: 4:00)



So if I want to put that in the regression , probability of Y being yes or no or probability of Y being pass or fail can be given as , P equal to e power x1b1 plus x2b2, x3b3... and b0 and divided by 1 plus x1b1 plus x2b2, x3b3... and b0, this is what we compute it in our regression formula.

You see the regression like y equal to mx plus c, here b1, bn is the weight, x1 is like independent variable and b0 is the intercept. So given the data, the model tries to fit this particular equation that is given a lot of training data, this model try to fit into this equation. After fitting this equation, we can use this equation to predict the future data, that is simple as this.

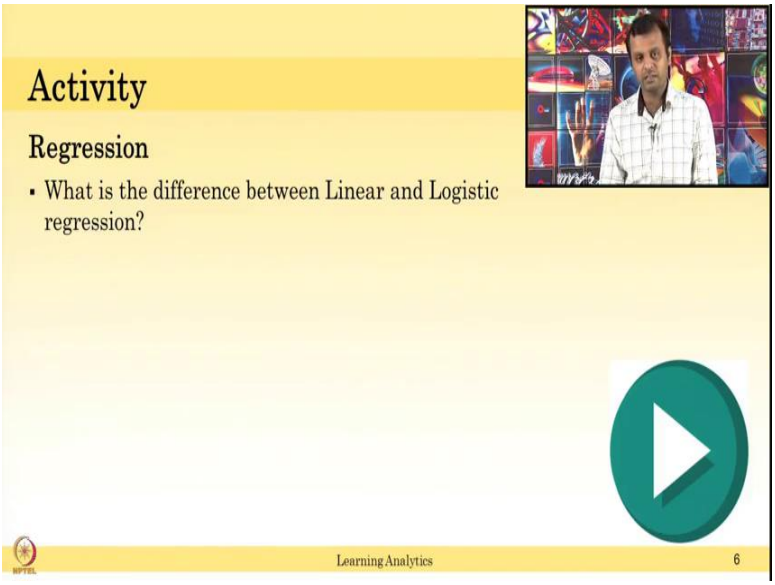
We will not go in detail, instead, this is the formula, there is a model which tries to fit the given data into this particular model. So it is similar to the linear regression where we get the weights weight 1, weight 2 or within weight n. An intercept value is the trained model or is the train model the while, y predict equal to intercept plus weight1 x1.

Similarly in logistic regression also you will get a weight b0, b1, b2, ..bn. Given that value, you can use that value to predict the future thing. For example, compute the weights coefficient using the training value and use that to predict the future values.

If the future value probability is greater than 0.5, you can put it as a yes. If it is less it will be 0, you can put it as no. But what about is 0.5? We saw that 0.5 in a Sigmoidal function value at a zeroth value. The 0 can be moved that is called the intercept is coming into the picture so above 0.5 you can group them into one particular classifier below 0.5 comes other classifier.

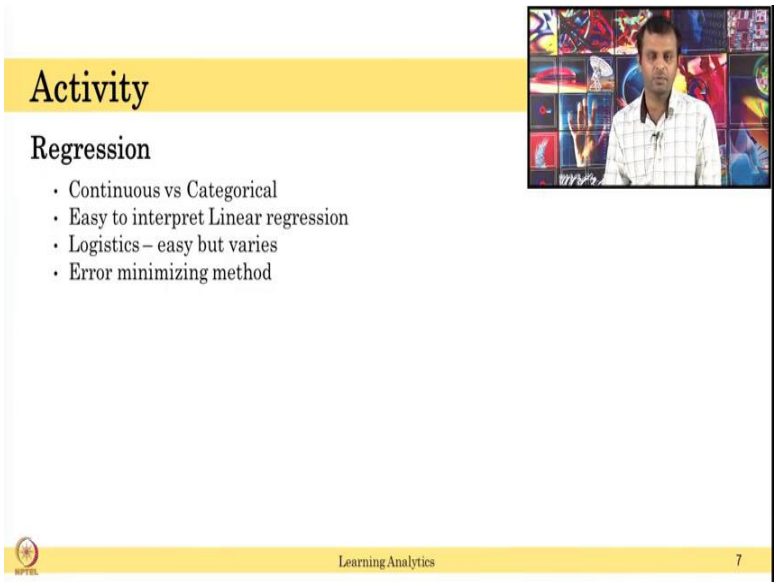
We compute the confusion matrix in machine learning,and the probability classifiers is used to predict the probability of outcome. And we can adjust the threshold of probability to improve the recall value and precision value, area under curve. If you remember that things that is how this probability is used.

(Refer Slide Time: 6:44)



So you knowing what is logistic regression and linear regression. In regression the plot is linear, and in a logistic regression the line is not linear. When the both model, you give input data, you get a train data in form of intercept weight 1 and weight 2 in both model but linear models fits simply into linear equation and a logistic equation it fits into a Sigmoid function, that is a difference.

(Refer Slide Time: 7:14)

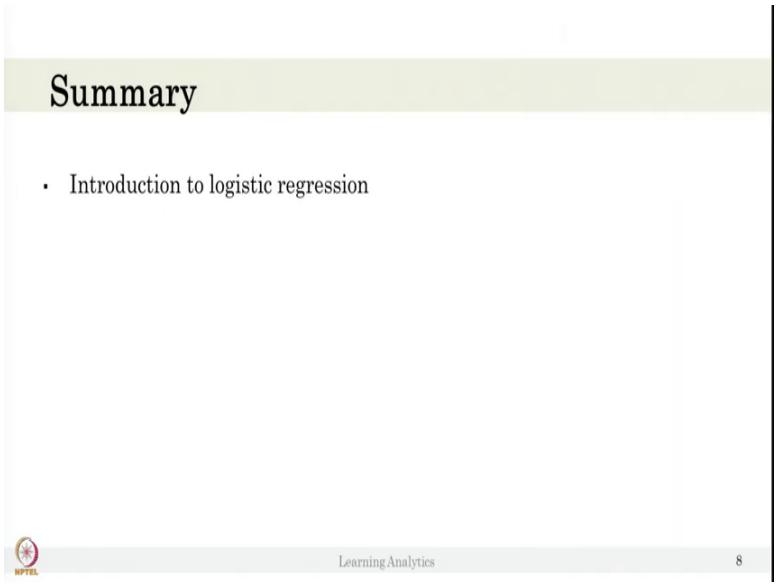


But what is the other difference between linear regression, logistic regression? Can you list down? After listing it down resume the video to continue. So it is simple, it is a continuous versus categorical data. In linear equation it is a continuous and in logistic regression it is a categorical yes or no pass or fail or multiple classes, not just two classes There are multiple ways to do the multi class classification.

Linear regression is easy to interpret, it is just a line you can say this is the weight of this particular value. It means keeping this all the other variables constant that's one as x1 is dependent on y1 in this particular scale, but in logistic regression also it is very easy because it also gives you the weight 1 weight 2, but it varies based on the functions you used the trained function used.

And error minimizing method in a linear regression we saw a least square, least mean square, we saw the difference between predictive value and actual value is used to predict the best objective function. But in logistic regression it uses ML method, maximum likelihood method to do that. So that is the difference. So we are not dealing with the training of a logistic regression in this video.

(Refer Slide Time: 8:35)



So in this video we introduced, what is logistic regression that is the purpose of this.

Also remember that most of the data science or data production happens today by using the logistic regression, although we have very sophisticated software or sophisticated classifiers for deep learning for better image classification or voice recognition. In a basic data classification of fraud detection or spam or anti-spam or lot of other things those regression is used. Thanks for watching this video. Thank you.