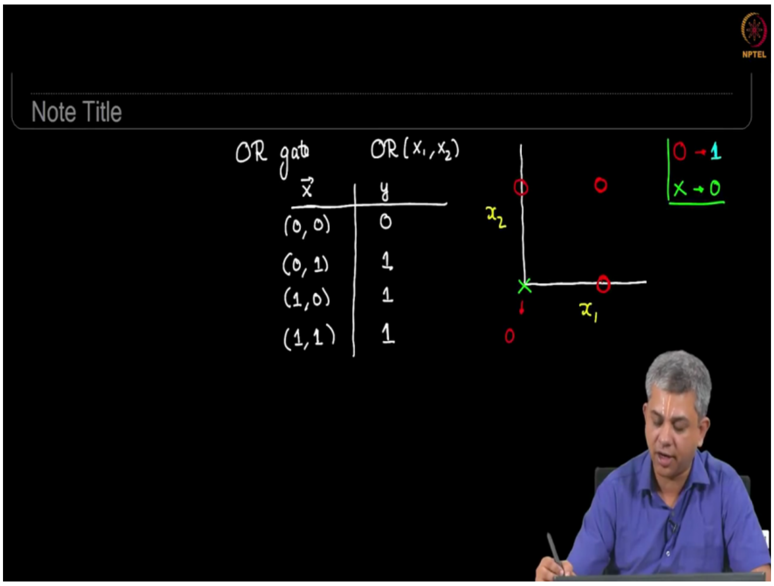
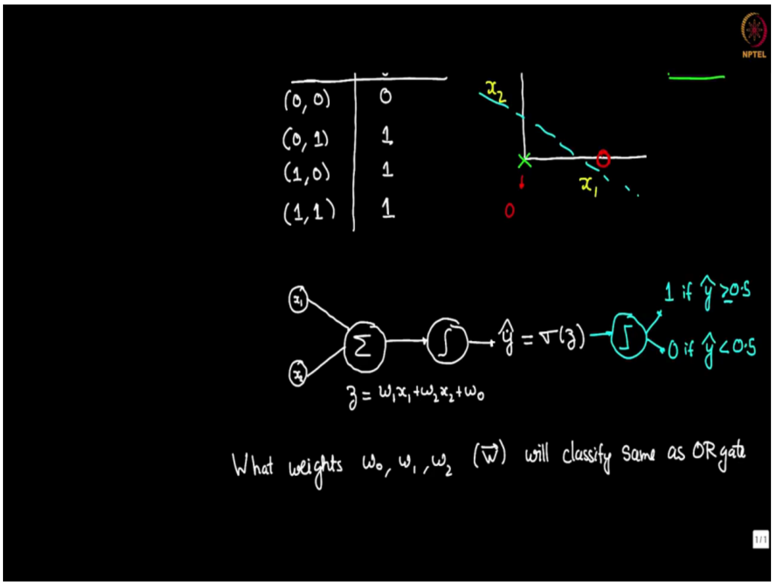
Machine Learning for Engineering and Science Applications
Professor Dr Balaji Srinivasan
Indian Institute of Technology, Madras
Department of Mechanical Engineering
OR Gate Via Classification

(Refer Slide Time: 1:11)



In this video we will be looking at a simple example of logistic regression by trying to represent OR Gate as a classification problem, so let me show you what I mean here. We all know what and OR Gate is, it takes into input let us call them X1 and X2, and X1 and X2 are always 0 OR 1 so let us call this X factor so 0 OR 0 you see 0, 0 OR 1 gives you 1, so even if one of the inputs is 1 we get a 1 so it is a simple logic gate. Suppose we represent this as a figure, let us take X1 here X2 on this axis, so 00 gives a 0 and the others give us 1 so let X represent 0 and let the circle the present 1 ok.

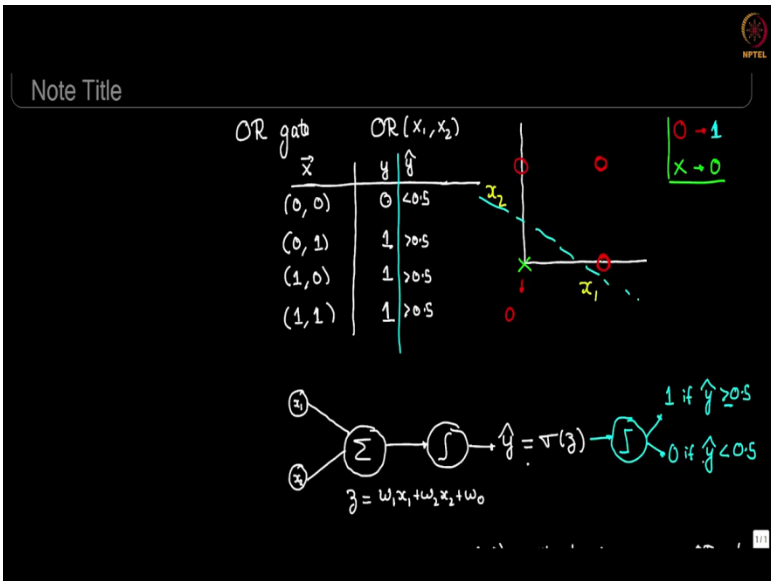
(Refer Slide Time: 4:17)



What we want to do is to find out an algorithm OR logistic regression which will classify at least these 4 points correctly, you can see this as a binary classification problem where **0 is** O is one class and X is another class ok. Now intuitively we can see that if I draw a line somewhere in the middle here, one side of the line will be classified correctly as X and the other side of the line will be classified correctly as O but we will try and do this mathematically. So remember, we can represent our logistic regression as a simple neutral diagram as follows, you have X1, you have X2, these 2 combined here gives submission Z is equal to W1 X1 + W2 X2 + biased unit W0 followed by a Sigma and our prediction Y hat is Sigma of Z.

Finally we classify after this as our prediction is 1 is Y hat is greater than 0.5 and it is 0 if Y hat is less than 0.5, we can just arbitrarily decide that the equal to sign goes to 1 ok. So the question we are asking is, what weights W0, W1, W2 all put together basically W vector will classify the same as the OR gate. If we do that, we have a essentially represented the OR Gate as a simple neural network okay OR as a simple logistic regression network ok so let us try these values here.

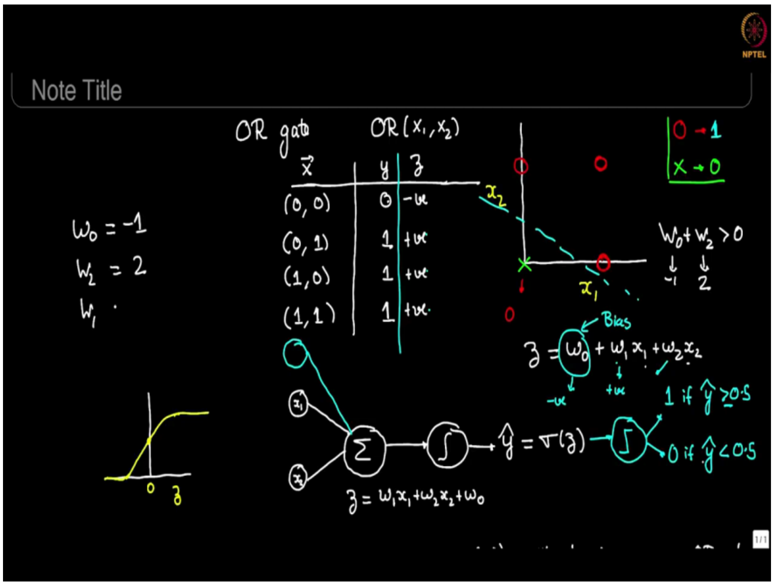
(Refer Slide Time: 5:05)



Now if I want this, let us back calculate this whole value and see if we can come up with something so if I want my classification as 0, it means Y hat has to be less than 0.5 ok. Let me write that here, Y hat has to be less than 0.5 here it has to be greater than 0.5, greater than 0.5, greater than 0.5.

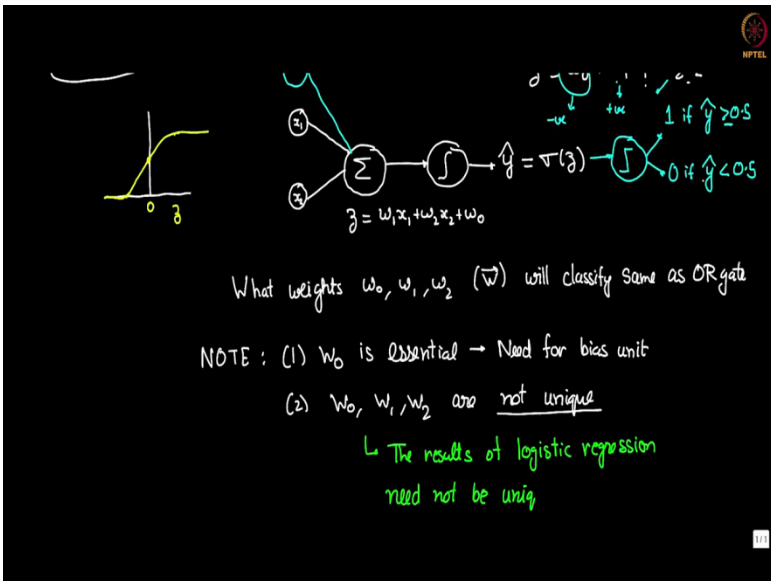
Y hat is Sigma of Z, we remember that the Sigma function works this way, if this is Z and this is Sigma of Z then at 0 your Sigma of Z is 0.5, whenever Z is greater than 0, Sigma is greater than 0.5, whenever that is less than 0, Sigma is less than 0.5 so now let us now find out what Z has to satisfy. If Sigma of Z has to be less than 0.5 then we know that Z has to be negative similarly to for these 3, Z has to be positive ok. Now we also know what Z is, Z is the W0 + W1 X1 plus W2 X2 so let us consider this case, both X1 and X 2 are 0 we want Z to be negative which means automatically that W0 has to be negative. Also notice that without W0 we could not have made this possible at all.

(Refer Slide Time: 7:29)



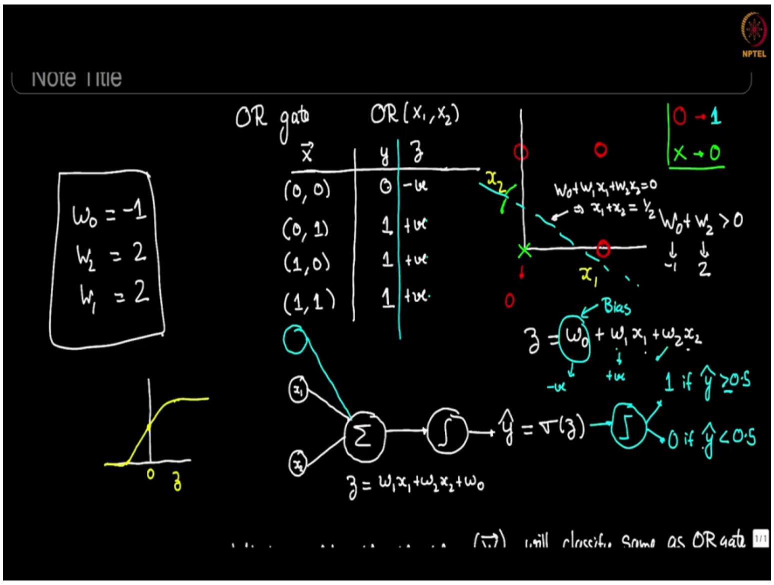
If you just have W1X1 plus W2 X2 that is a simple linear combination of X1 and X2 without this assign without this biased term, you cannot make this case work out at all ok. So we know now that W0 has to be negative. We also know that W1 and W2 have to be positive to make these cases work, so what is one set of weights which will work? Let us take a simple example, let us take W0 is equal to - 1, we come to this example it tells us that W2 X2 is active, X2 is 1, I now know that W0 plus W2 has to be greater than 0, W0 is already -1, I can easily assign W2 equal to let us say 2. Similarly you can argue that W1 can be made 2, so we have come to this set of weights.

(Refer Slide Time: 8:01)



Now notice something, we will come to the physical meaning of this weight soon but note, 1st of course the point I said earlier is essential, you cannot make this go away which tells you the need for the biased unit. Second, notice that W0, W1, W2 are not unique, this is very important. In general, the result of logistic regression need not be unique. In fact you can see this geometrically true, this classifying line here whichever line we draw arbitrarily this has a lot of give, you can go back and forth and still make this work, you can make this inclined in several ways.

(Refer Slide Time: 10:04)



Also this lack of uniqueness is due to 2 things, you can even multiply W0, W1, W2 by a constant and even that would not be unique okay, but that is a trivial sort of non-uniqueness, more actual sort of non-uniqueness is the fact that this line can move back and forth, it can also be translated as well as rotated a little bit and still the classification will work, so these 2 are important point for us to notice. Now the line that we have got is W 0 + W1 X1 + W2 X2 is Z. Remember, the interpretation of the classification line is that this is the line Z equals 0 which would be the line W0 plus W1 X1 plus W2 X2 equal to 0, which is the line if you write down the values X1 plus X2 equal to half because W0 is -1 divided by W1 and W2.

So this is exactly the line that bisects these 2 okay, so that is the arbitrary set of values that we have given here. You could also give the value for example W0 = -1, W2 = 3, W1 = 3 in which case the classification line would be inclined at a slightly different angle ok. So this would be it would be still parallel to this but will be slightly here ok so that would be the line W0 X1 plus X2 = 1 by 3. So this tells you how logistic regression can be replicated, OR gate can be replicated using logistic regression. In the next video we will see how the other gates can also be replicated using logistic regression.