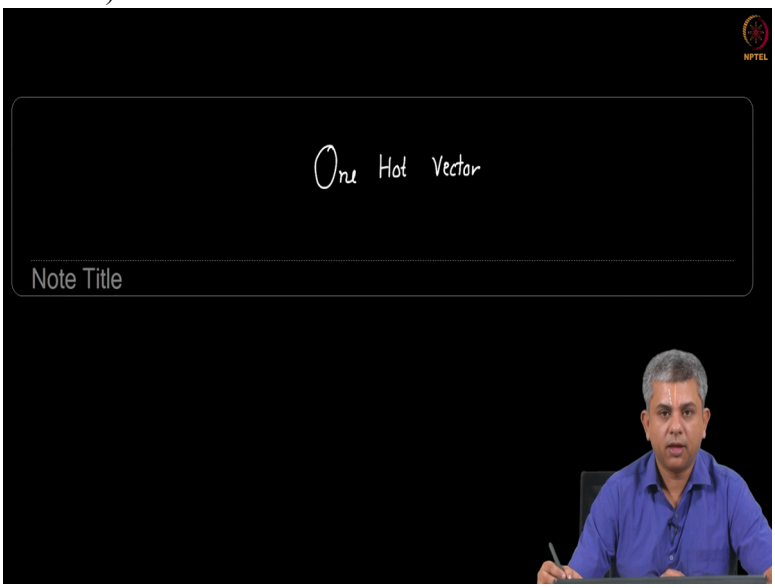
Machine Learning for Engineering and Science Applications
Professor Doctor Balaji Srinivasan
Department of Mechanical Engineering
Indian Institute of Technology Madras
Multinomial Classification One Hot Vector

(Refer Slide Time: 00:14)



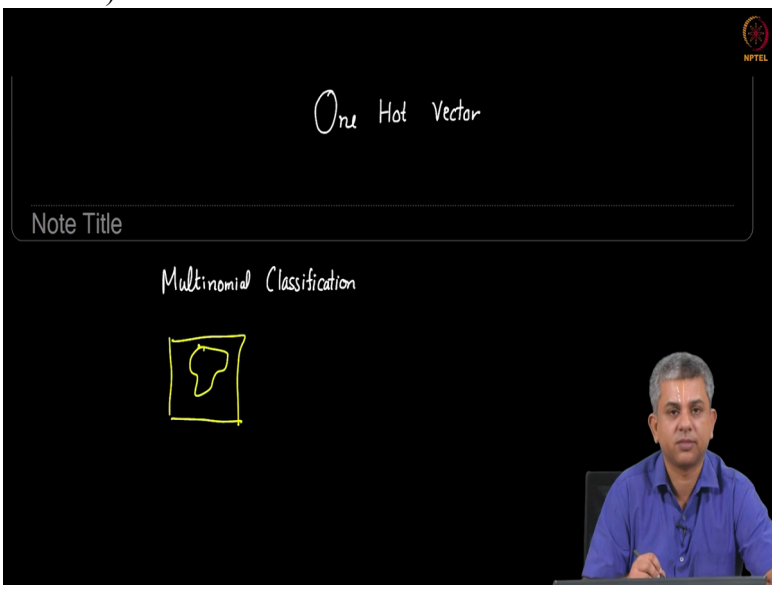
In this video we will be looking at the One Hot Vector representation. Remember that the One Hot Vector representation is used for a case where you have more than 2 classes. You can also use it for the case when you have just 2 classes but

(Refer Slide Time: 00:36)

it is generally overkill to use a One Hot Vector.

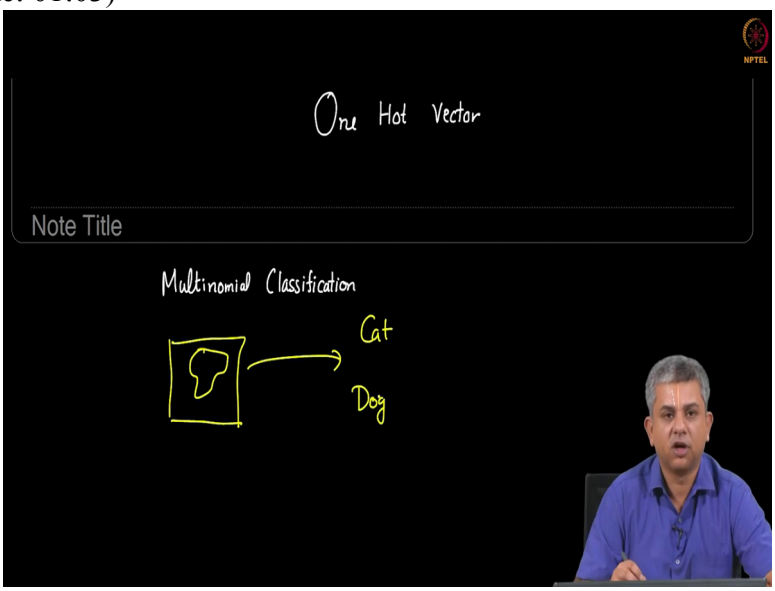
Let us start with a binary classification task and then we will go to a multinomial classification task and see how a One Hot Vector works. So let us say you have 2 classes. You have an image. And this image is that of a cat or a dog.

(Refer Slide Time: 00:55)



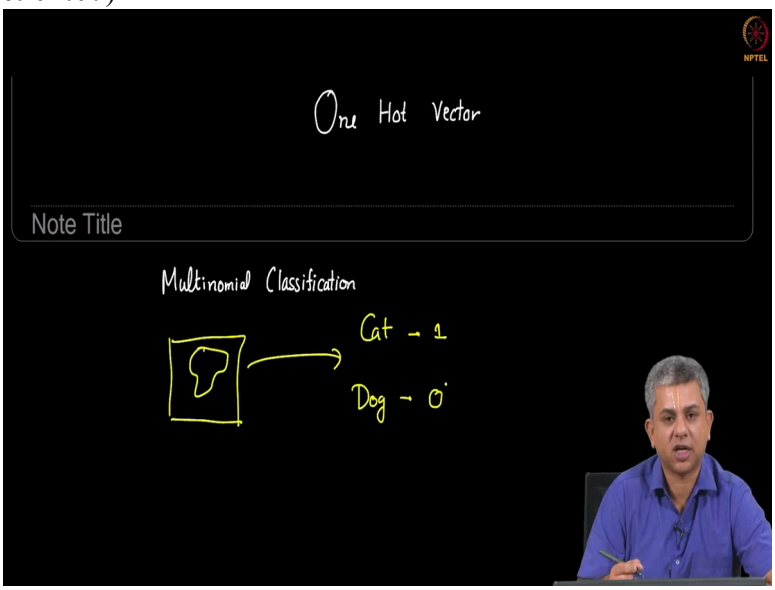
And I want to represent the output. The output says either cat or it says dog.

(Refer Slide Time: 01:03)



So earlier we would simply call cat as class 1

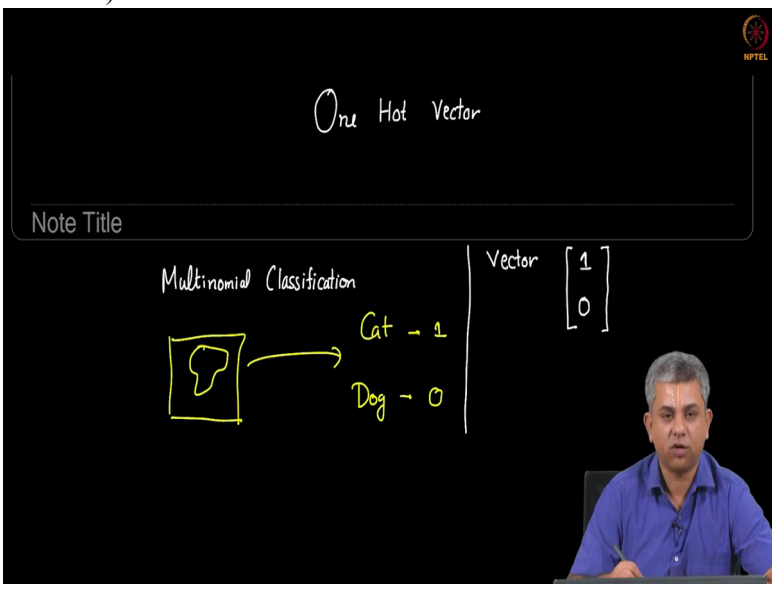
(Refer Slide Time: 01:07)



and dog as class 0 or vice versa. Either way it is a simple number. Now in order to generalize for the multiple class cases we are going to represent our output instead as a vector.

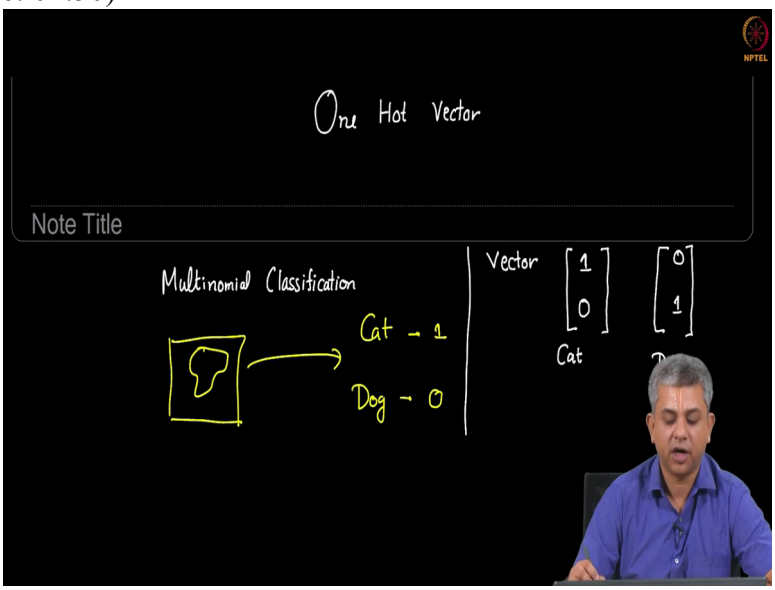
So the vector is going to have 2 elements, Ok. So in the ideal case instead of calling cat simply 1, I would call it 1 0,

(Refer Slide Time: 01:38)



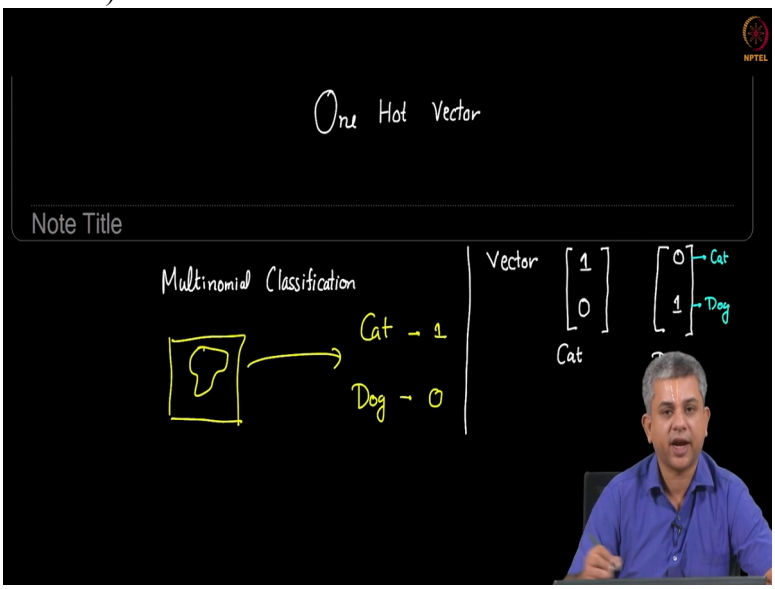
this would represent cat. And I could call dog as 0 1. This would represent

(Refer Slide Time: 01:50)



a dog. Essentially first element represents the cat;

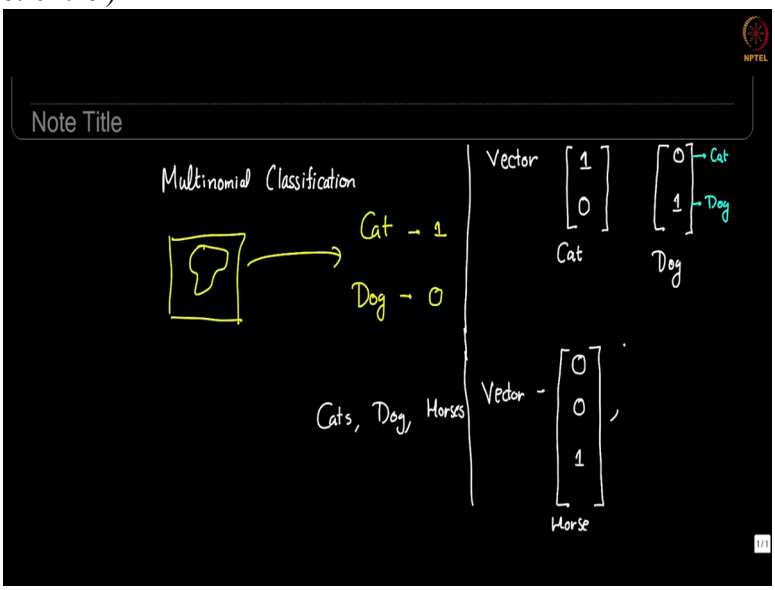
(Refer Slide Time: 01:58)



the second element represents the dog, Ok.

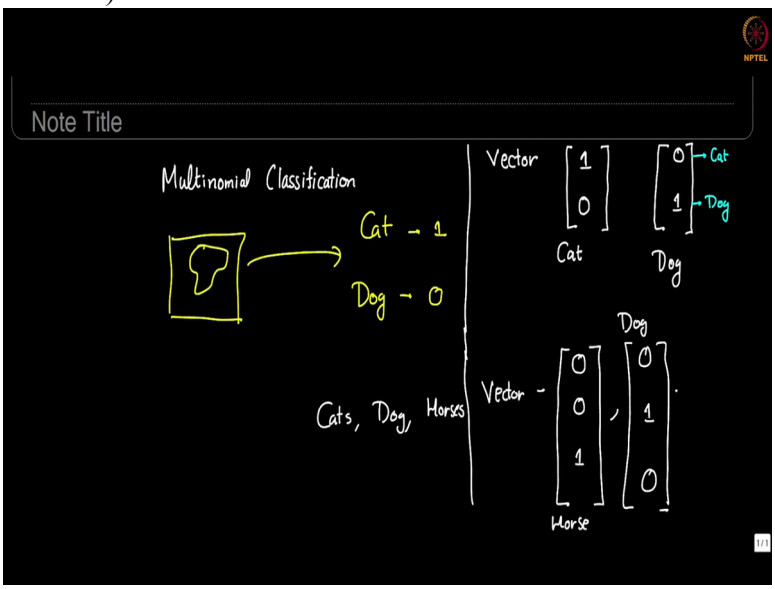
Now how does this help? Let us go one more step. Let us say we have 3 classes. Let us say we have cats, dogs and horses. Once again the One Hot Vector then looks like 0 0 1 for a horse,

(Refer Slide Time: 02:29)



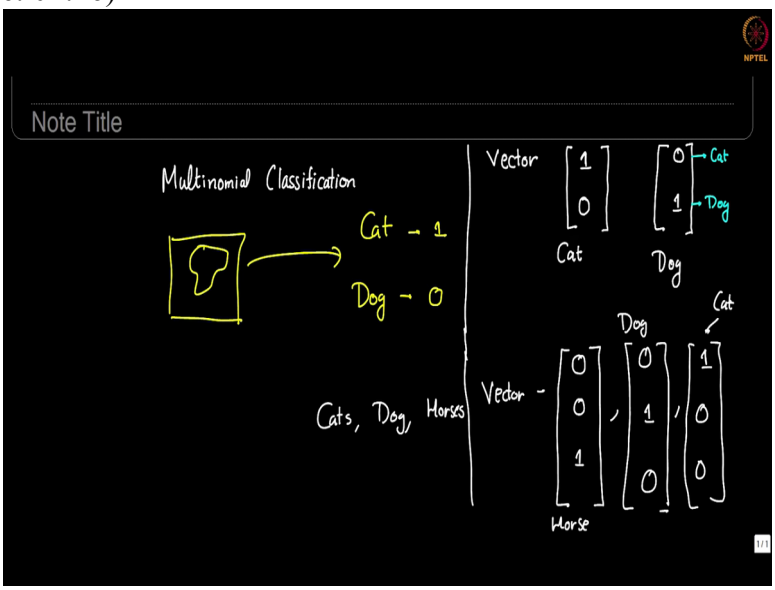
0 1 0 for a dog. This is of course

(Refer Slide Time: 02:38)



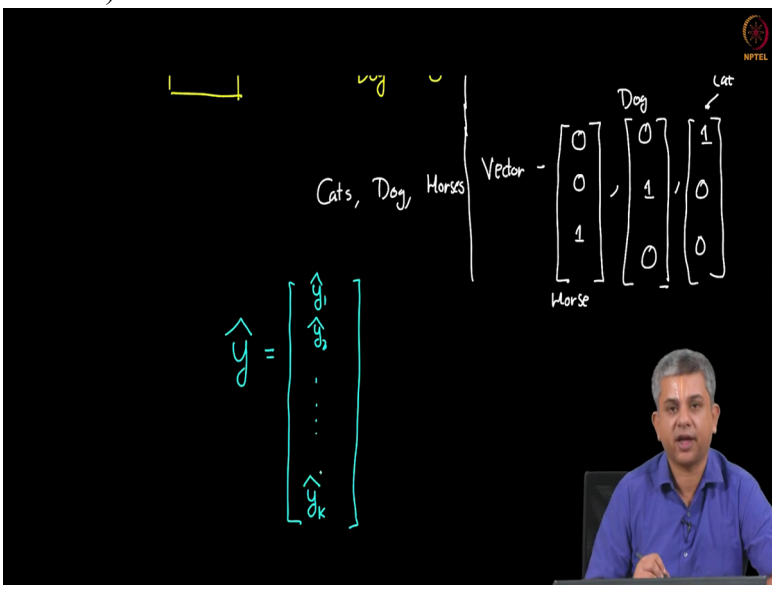
arbitrary. You just have to make consistent choices, 1 0 0 for a cat.

(Refer Slide Time: 02:46)



In general if our prediction is y hat, y hat is going to have k elements, y hat 1, y hat 2 so on and so forth up until

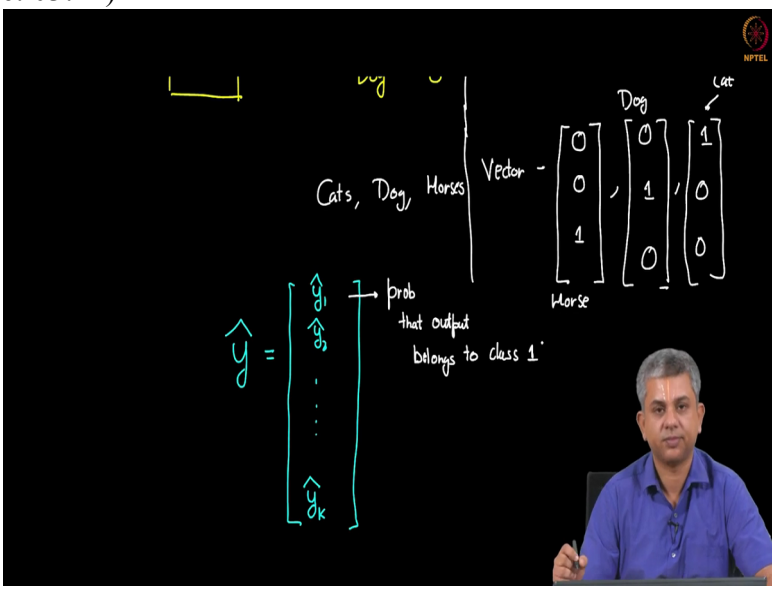
(Refer Slide Time: 03:07)



y hat k. So k was 3 in our cats dogs horses example. What is the meaning of each of these terms? Let us think about this carefully.

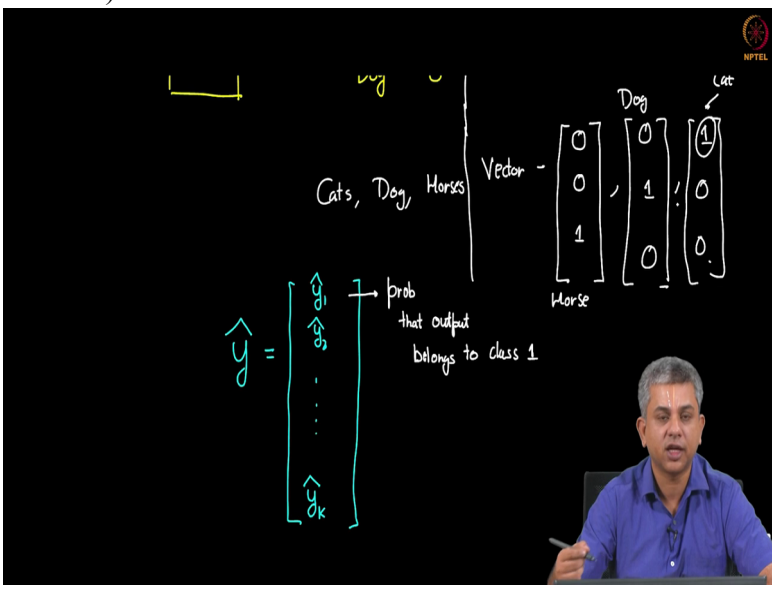
So in my cats dogs horses example this is probability that my output belongs to class 1, Ok.

(Refer Slide Time: 03:41)



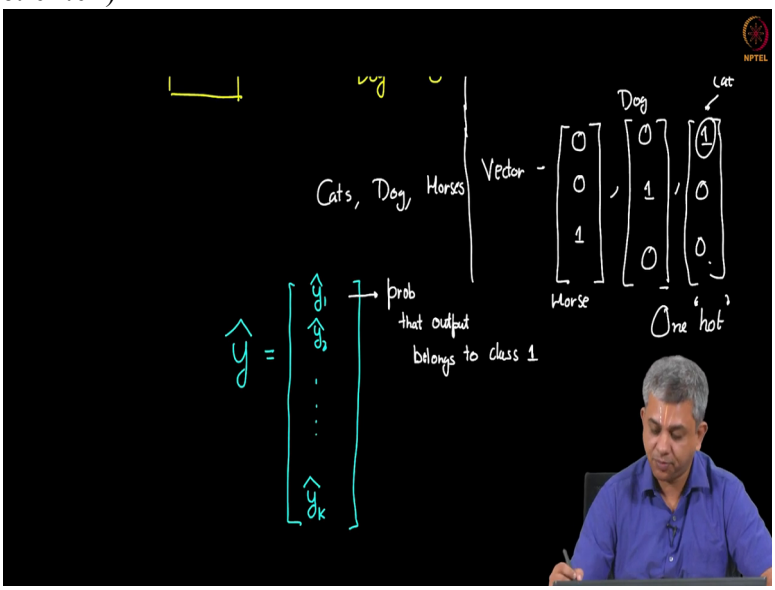
Now my ideal output would be simply 1 0 0, 0 1 0 or 0 0 1, which is why it is called One Hot, that is only one element is 1,

(Refer Slide Time: 03:52)



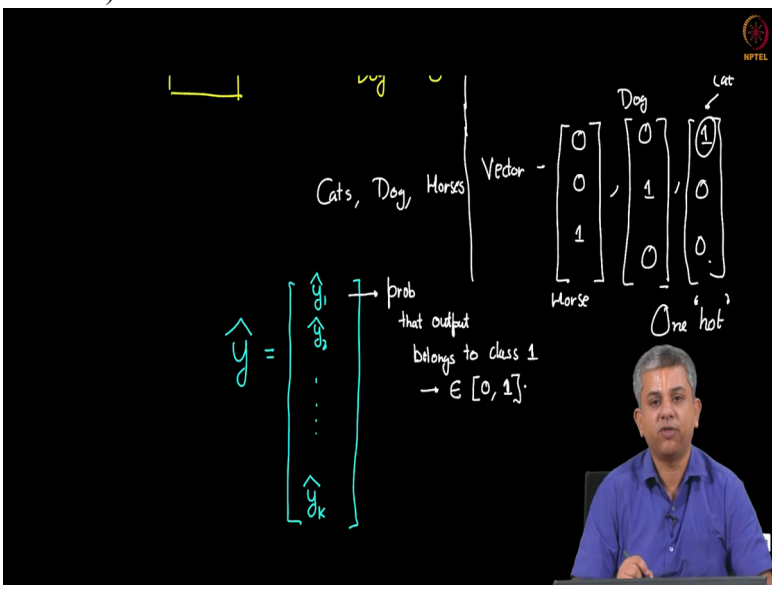
all the others are 0. That is why it is called One Hot Vector. All the others effectively are cold.

(Refer Slide Time: 04:01)



However we know through logistic regression that we are not going to get precisely 0 or 1. We will get some number. This has to belong to some number

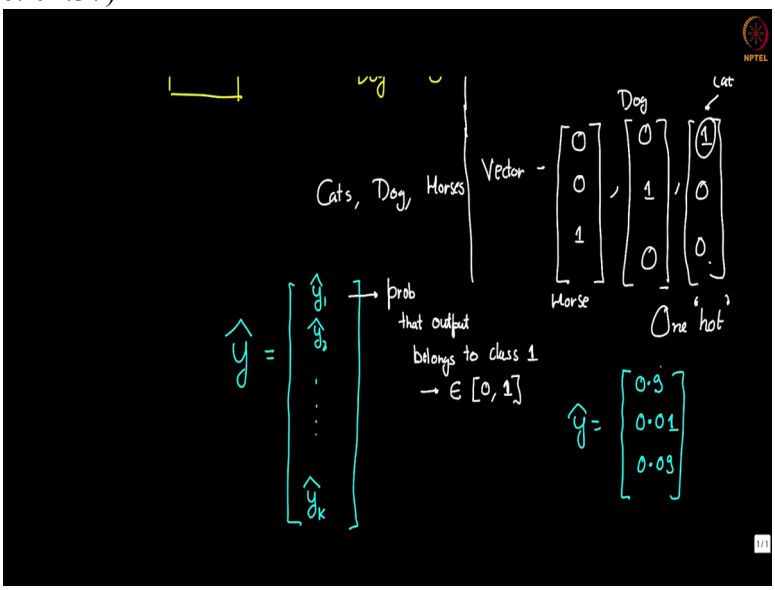
(Refer Slide Time: 04:15)



between 0 and 1. What our actual prediction will typically look like something like this, point 9, point 0 1, point 0 9, something of that sort.

This would represent to us, if I have a number of this sort, that this has

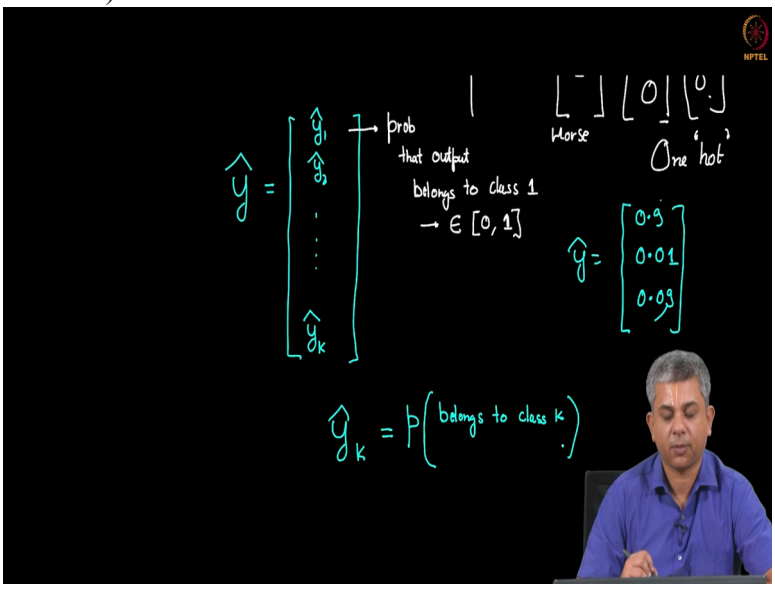
(Refer Slide Time: 04:37)



a probability of point 9 then it is a cat, it has a probability of point 0 1 that it is a dog. And it is a probability of point 0 9 that it is a horse, Ok. So each element here, the element y hat k represents the probability, Ok.

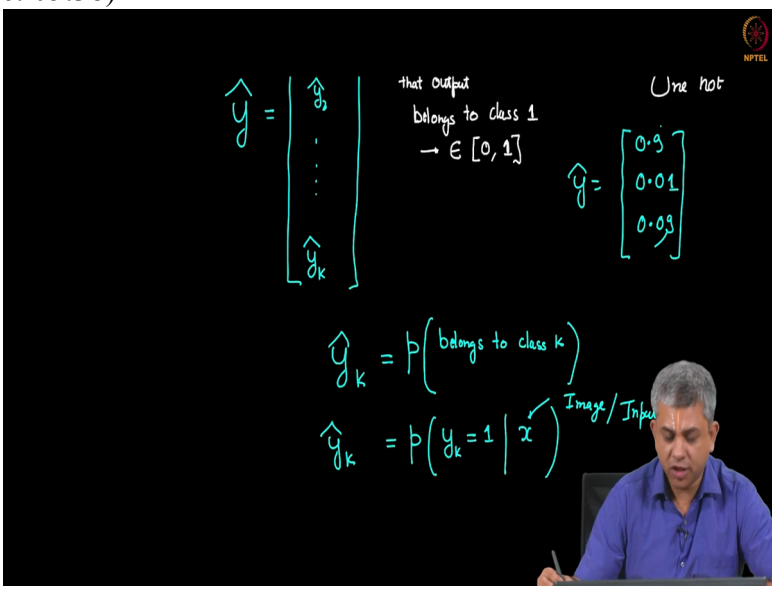
I will first write it in words and then I will write it in more mathematical term, belongs to class k.

(Refer Slide Time: 05:12)



More precisely this is probability that y k equal to 1 given whatever our input image, Ok. x here denotes the image in the example and in general it would denote whatever input we have here.

(Refer Slide Time: 05:38)



So this is the One Hot Vector and as you can see it can generalize to a, from a binary to a k class problem. We will be using this for multinomial classification.