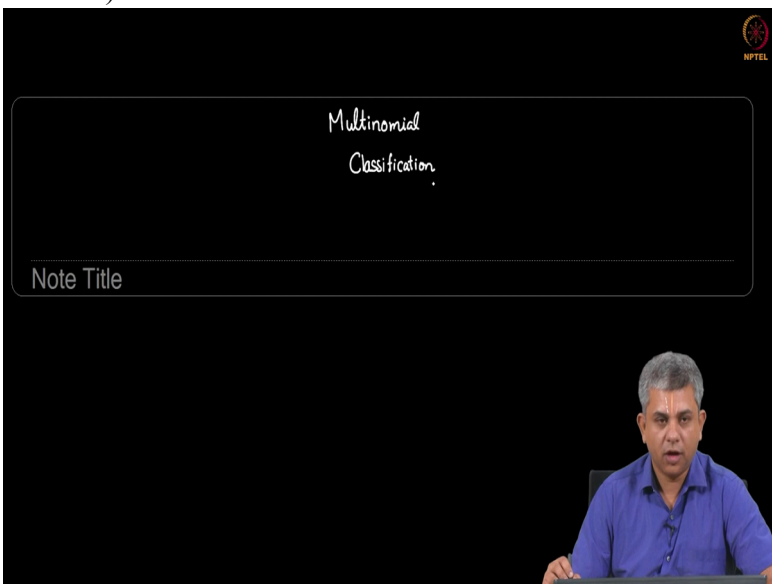
Machine Learning for Engineering and Science Applications
Professor Doctor Balaji Srinivasan
Department of Mechanical Engineering
Indian Institute of Technology Madras
Multinomial Classification Introduction

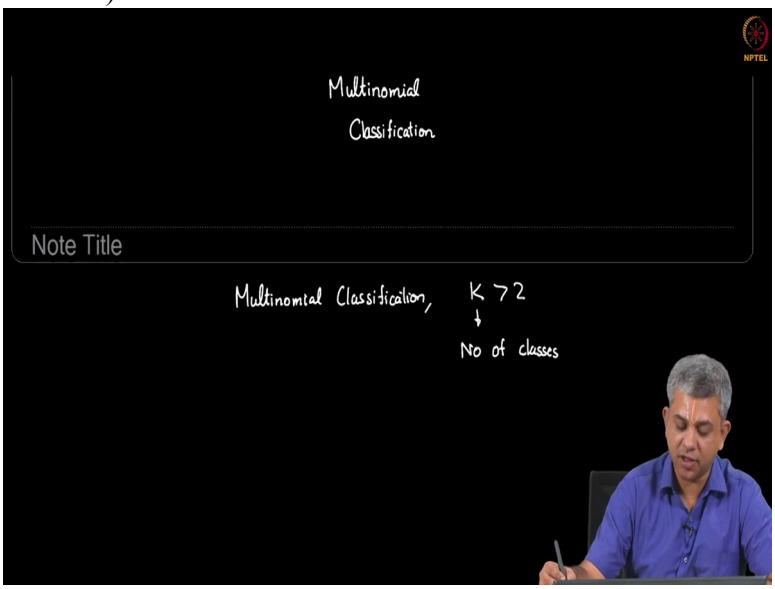
(Refer Slide Time: 00:14)



In this video we will be extending our binary classification which we did using logistic regression into the general problem of multinomial classification. Multinomial classification is simply the case when you have more than 1 classes more than 2 classes.

For example if you have a case where you want to classify digits, Ok. We will see

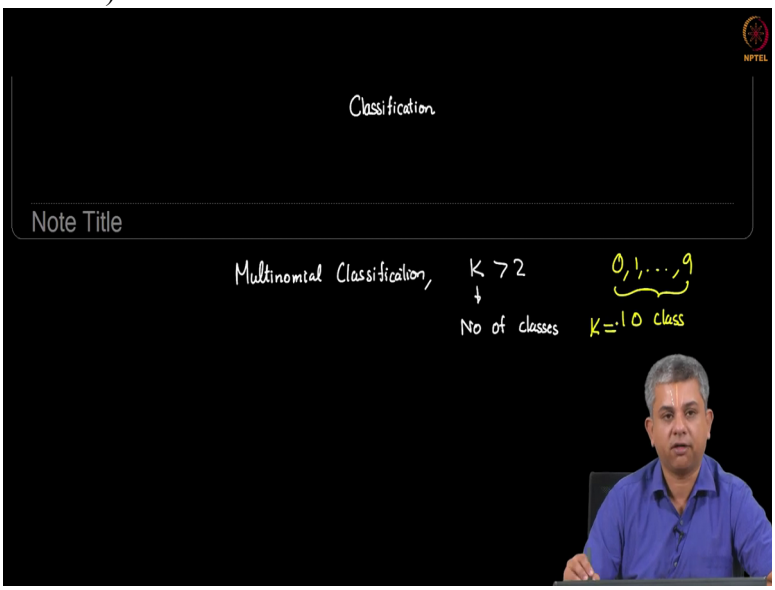
(Refer Slide Time: 00:53)



examples of this later. So let us say you have all your digits from 0 through 9. So this is a 10 class classification problem.

Somebody has handwritten digits and we have to find out which digit it is. That would be a 10 class problem. So k equal to 10

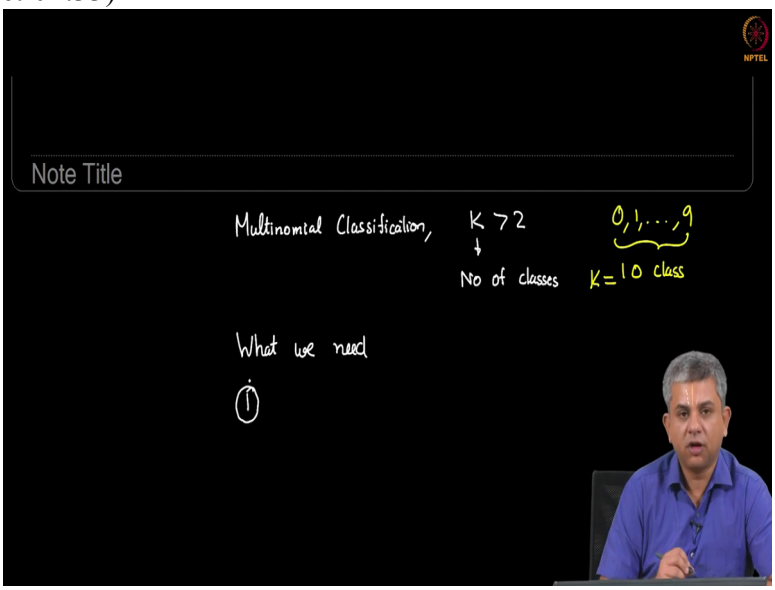
(Refer Slide Time: 01:12)



in such a case, Ok.

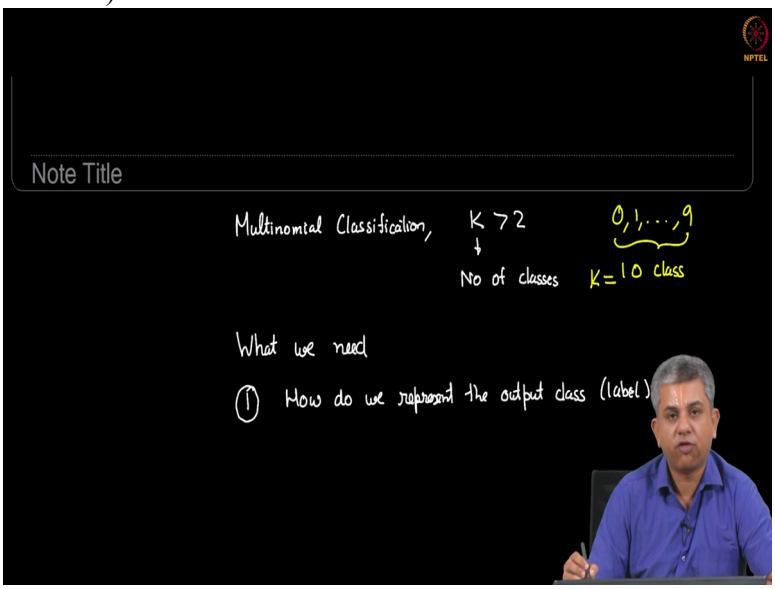
So when we try to solve such problems there are primarily a few things that we have to do over and above, over and above what we did for the binary classification problem, Ok. So what we need is the following.

(Refer Slide Time: 01:33)



First we need to know how do we represent the output class. For example in

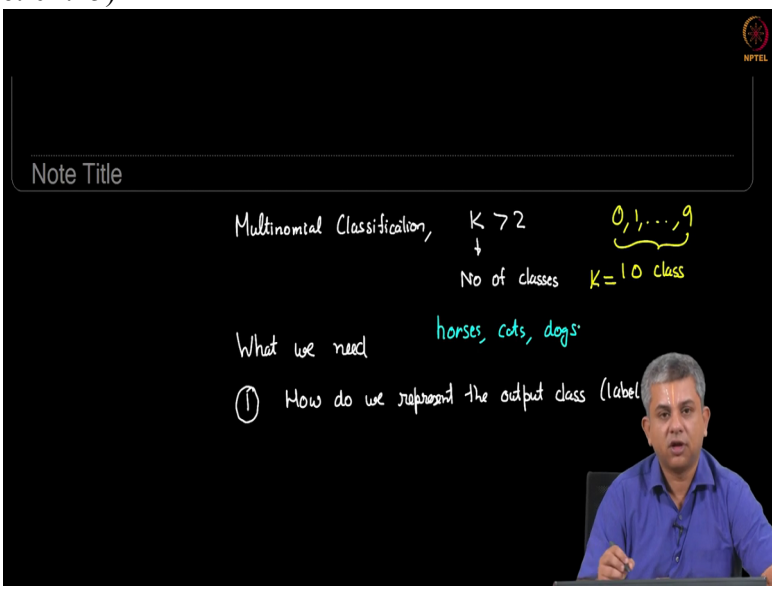
(Refer Slide Time: 01:52)



binary classification we simply decided if it belonged to one class we would label it as 0, if it belonged to another class we would label it as 1.

Let us say we are dealing with a case where our 3 classes are something, we are trying to label images. And the 3 classes are horses, cats and dogs.

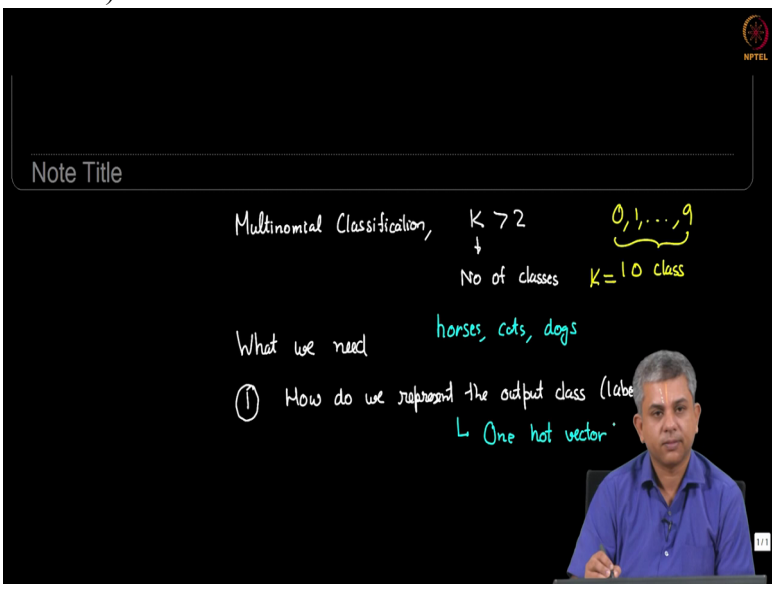
(Refer Slide Time: 02:15)



You have several choices. Obviously you cannot label it simply as words, as we have discussed in several videos before. You need to give a numerical name.

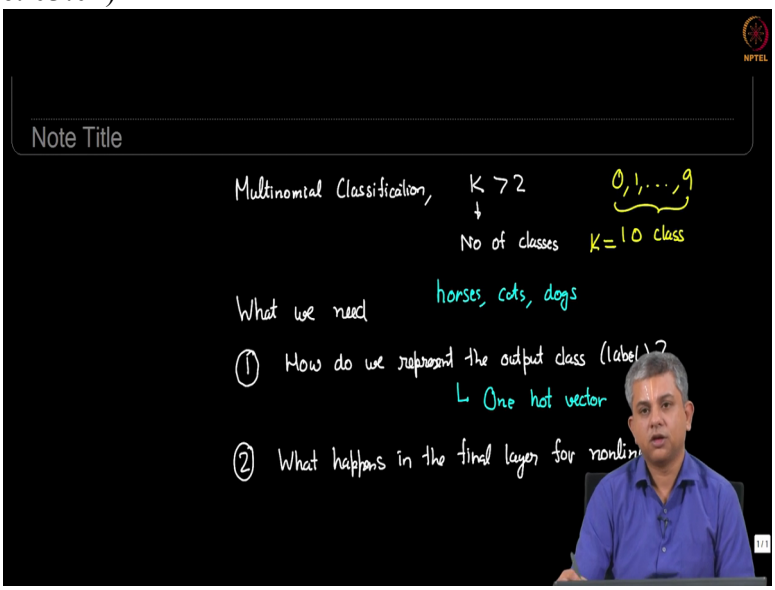
So one choice is to simply call this class 0, class 1, class 2. That is one possibility. We look at one other solution which is called the One hot vector.

(Refer Slide Time: 02:38)



So we will be looking at this. The second problem is what happens in the final layer?

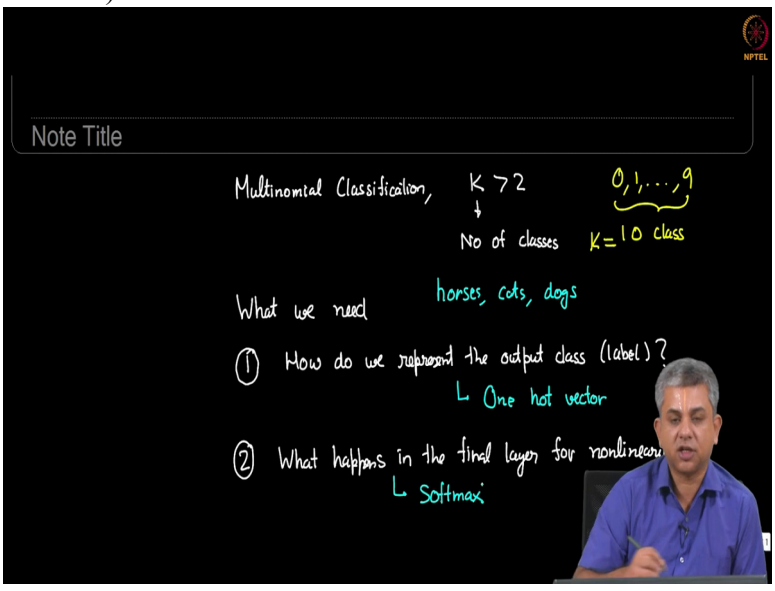
(Refer Slide Time: 03:02)



So recall that in your binary classification task we used the sigmoid because it neatly gave us a number between 0 and 1. And if it was close to 0, we knew it belonged to class 0 and if it was close to 1 we knew that it belonged to class 1.

Now what do we do in a multinomial classification case? So we will be looking at a function called the Softmax function

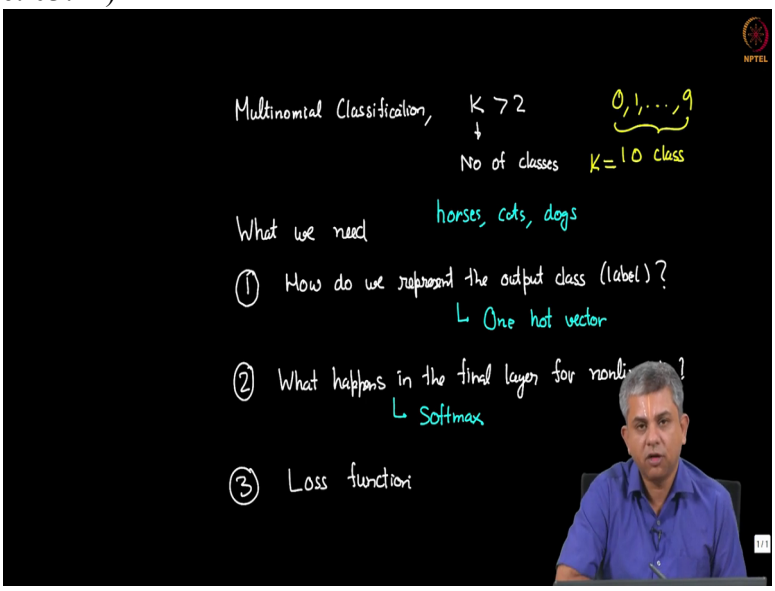
(Refer Slide Time: 03:27)



which actually corresponds very well with the One Hot Vector. So we will use that for multinomial classification.

Finally we need to answer how we are going to calculate the loss or cost function

(Refer Slide Time: 03:44)



for this case. We looked at the binary class entropy as the cost function for binary classification.

It turns out that we will use something very, very similar even for the multinomial classification. So in the videos that follow, we look at one hot vector, we look at Softmax and we look at how we are going to look at the loss function for multinomial classification.