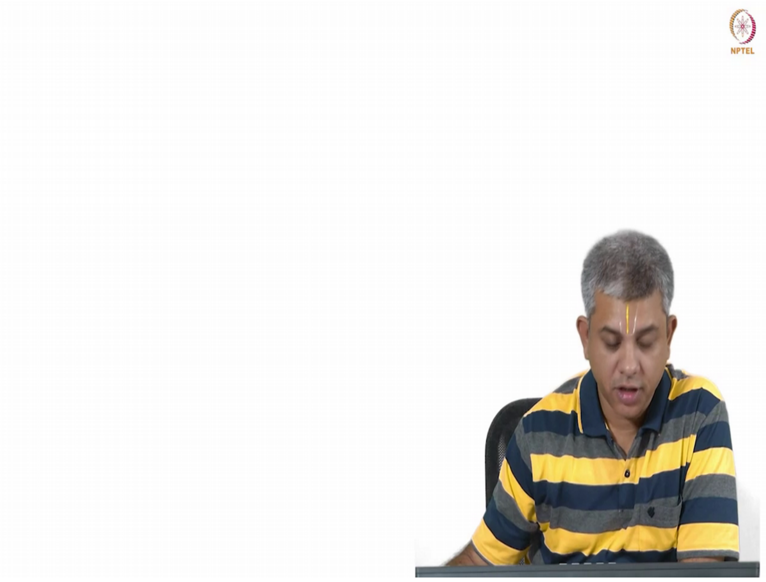
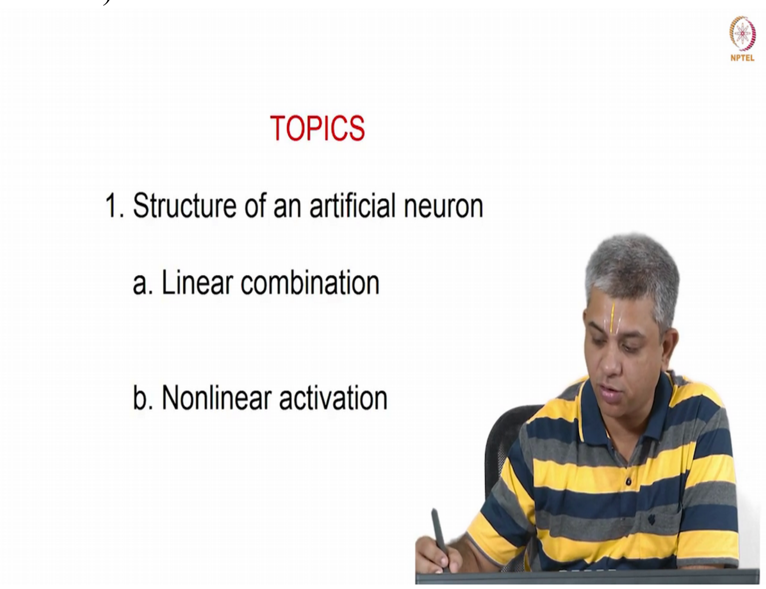
Machine Learning for Engineering and Science Applications
Professor Doctor Balaji Srinivasan
Department of Mechanical Engineering
Indian Institute of Technology Madras
Structure of an Artificial Neuron

(Refer Slide Time: 00:13)



In this video we will be looking at the structure of an artificial neuron.

(Refer Slide Time: 00:17)

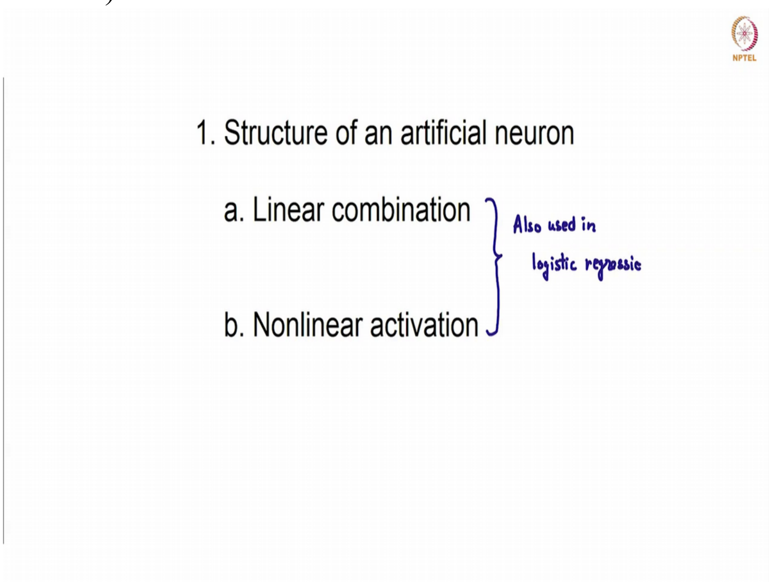


So the artificial neuron is a simple abstraction of a biological neuron.

It is supposed to abstract out all the details that actually exist in the biological neuron and just give you whatever is usable. Please do remember that this is a simplification and this is not how actually our brain neurons work.

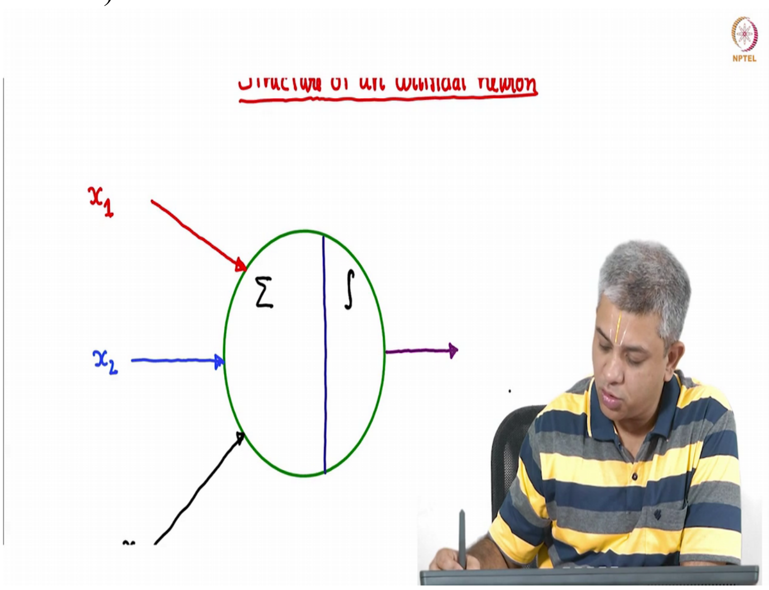
In this video the topics that we will be covering are what goes in into an artificial neuron and the two operations. Really speaking both of these operations were things that you had already seen within logistic regression, Ok. We also use these two in logistic regression. We are just going to formally combine these operations into a single thing called an artificial neuron.

(Refer Slide Time: 01:06)



So let us look at what an artificial neuron looks like

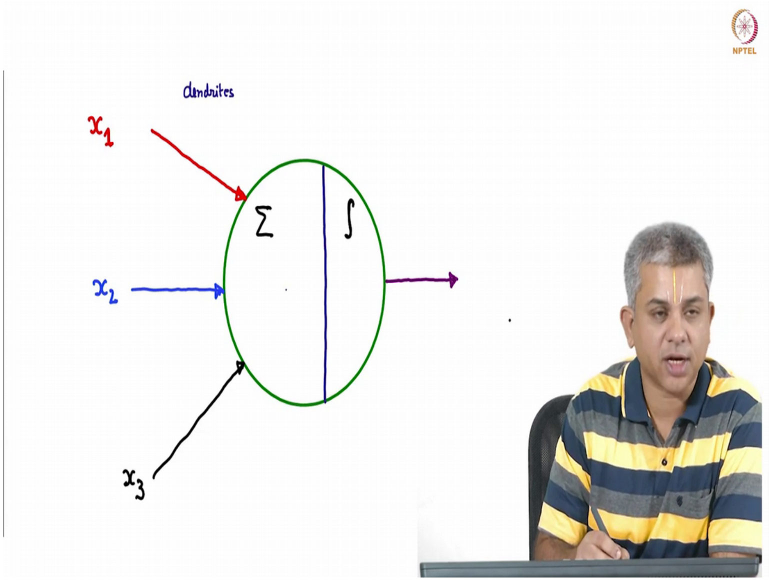
(Refer Slide Time: 01:11)



in a neural network. So suppose you have some inputs coming in from one end. We can call these, biologically these are supposed to be equivalently of dendrites.

So you have

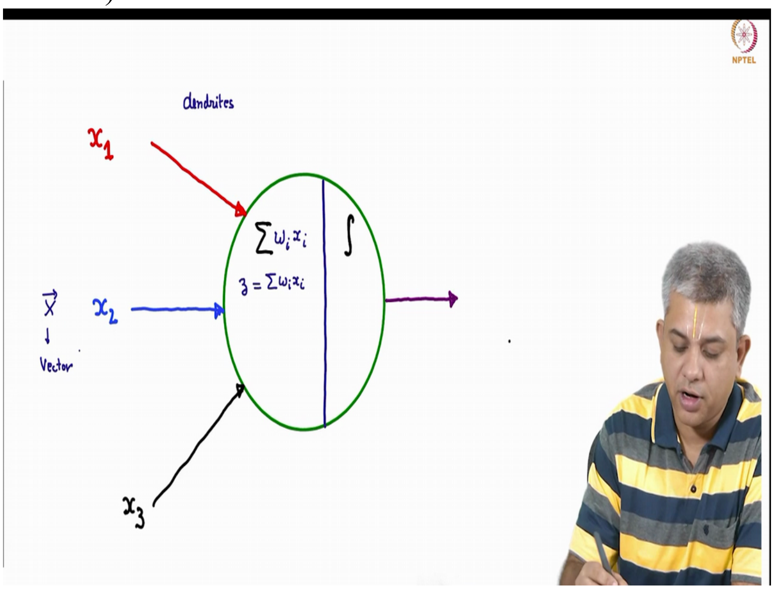
(Refer Slide Time: 01:25)



some 3 variables. In general some vector X vector which has these features, x 1, x 2, x 3. All these three come in and as usual as we did with both linear regression as well as classification we simply take a linear combination. So we do sigma w i x i and this component we call z.

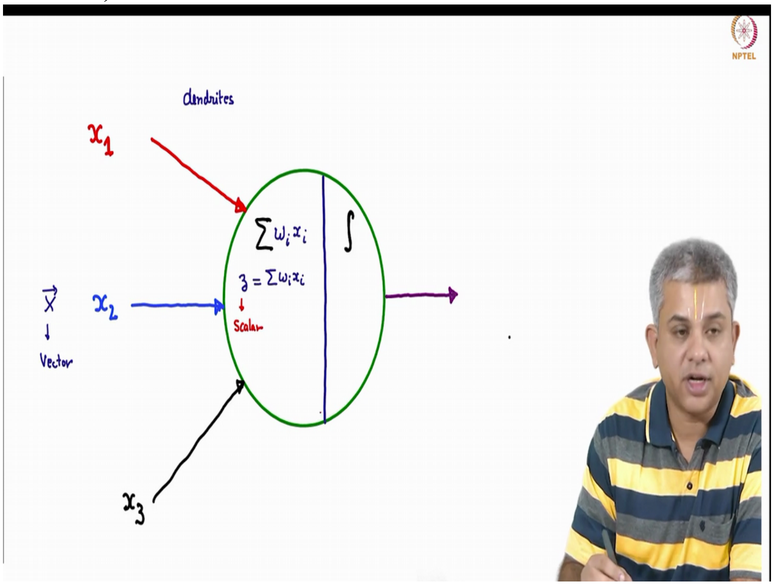
Remember the X that comes in is a vector

(Refer Slide Time: 02:02)



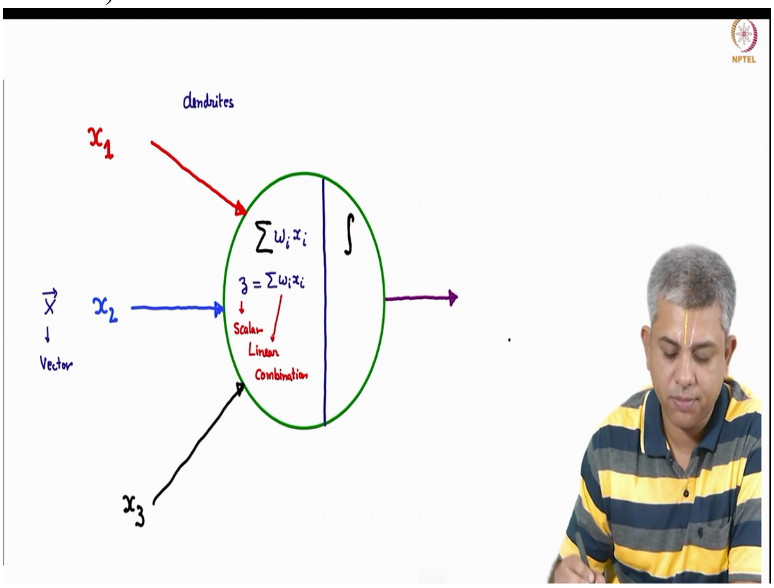
and what comes out, out of this linear combination z is a scalar. Till this point

(Refer Slide Time: 02:11)



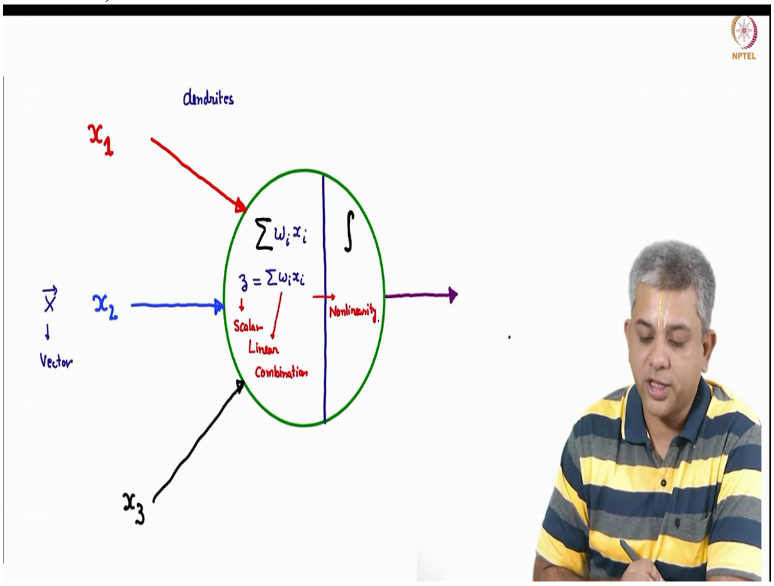
what we have here is essentially a linear combination.

(Refer Slide Time: 02:22)



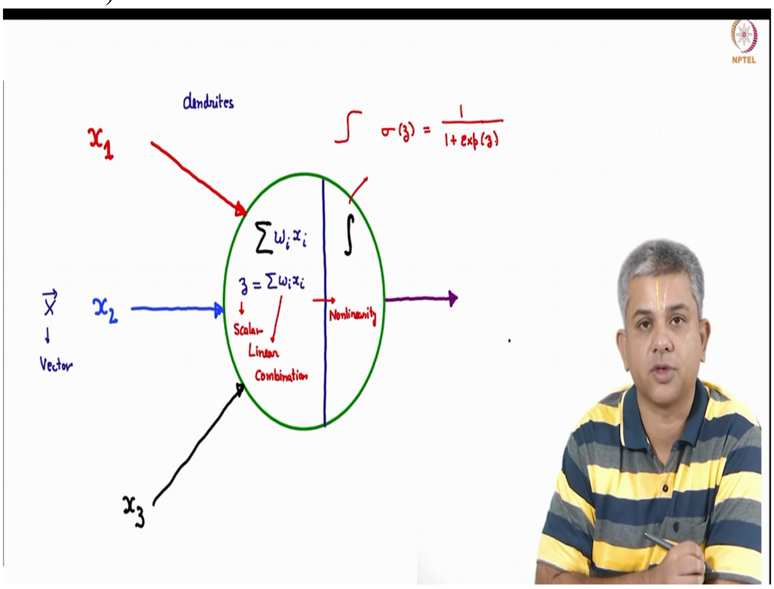
Now this z goes into the next part which is the nonlinearity. And what makes neural networks work really is this nonlinearity.

(Refer Slide Time: 02:39)



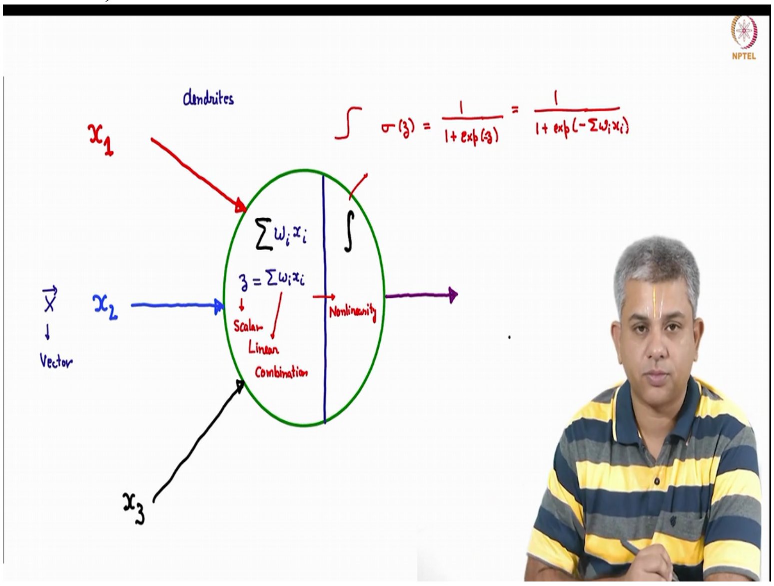
One simple nonlinear function which we have already seen, for example is the sigmoid function. Graphically we denoted by the shape of the sigmoid curve. So sigmoid of z would be 1 plus 1 by 1 plus exponential

(Refer Slide Time: 03:02)



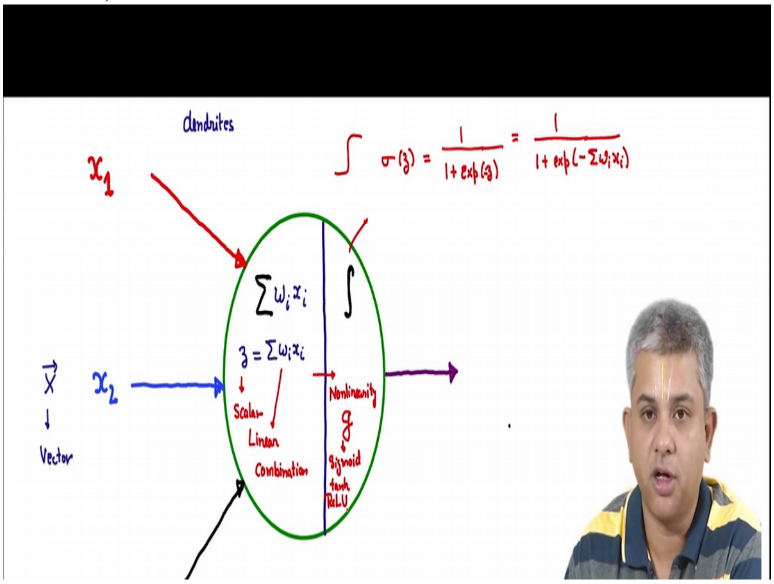
z, sorry minus z which is the same as...

(Refer Slide Time: 03:17)



So we call this nonlinearity in general g. This might be the sigmoid. It could be other functions which we will see later, for example tan h. We also have another thing called the rectified linear unit

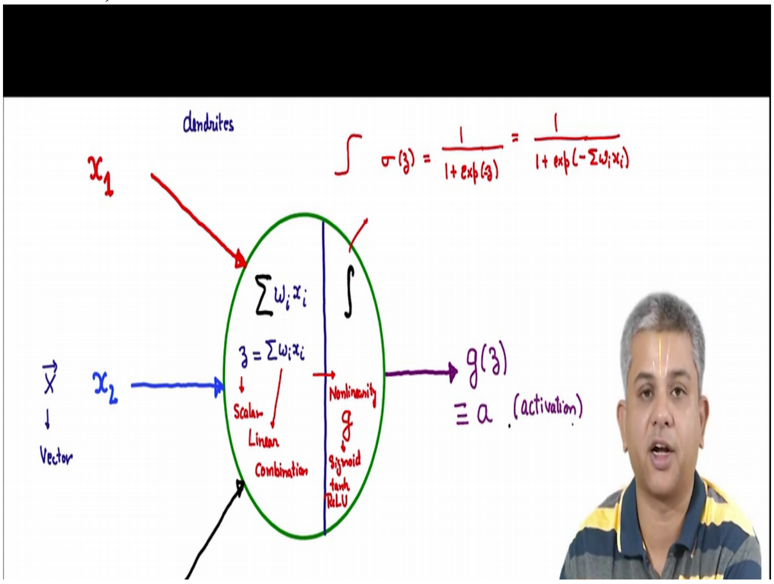
(Refer Slide Time: 03:35)



Re L U. So any of these outputs, any of these nonlinearities could be used.

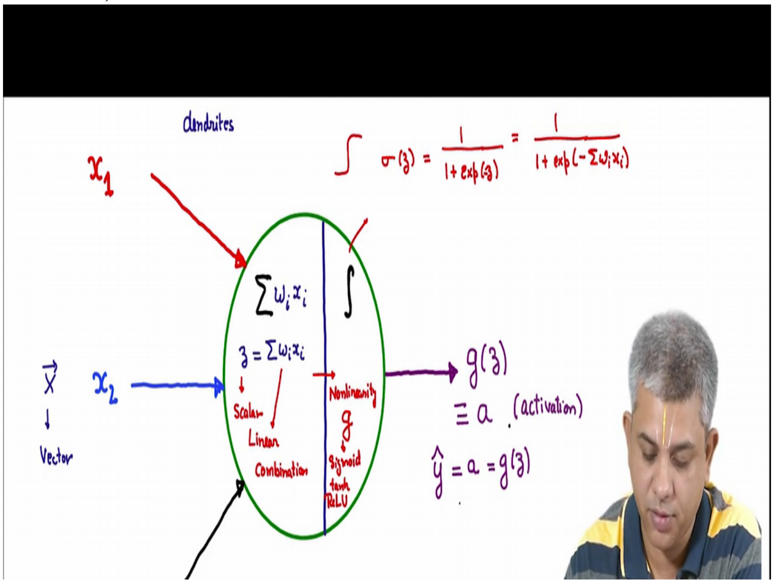
Now finally after all this, what comes out is g of z. This is denoted by a, also labeled the activation.

(Refer Slide Time: 03:59)



If this is the only thing in your network then all you have is your prediction is simply the activation of this neuron which is g of z.

(Refer Slide Time: 04:12)



So remember the two portions of an artificial neuron are linear combination and an additional nonlinearity.

(Refer Slide Time: 04:31)