

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)

TOPICS

1. Structure of an artificial neuron
 - a. Linear combination
 - b. Nonlinear activation



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 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)

1. Structure of an artificial neuron

- a. Linear combination
- b. Nonlinear activation

Also used in logistic regression

So let us look at what an artificial neuron looks like

(Refer Slide Time: 01:11)

STRUCTURE OF AN ARTIFICIAL NEURON

x_1

x_2

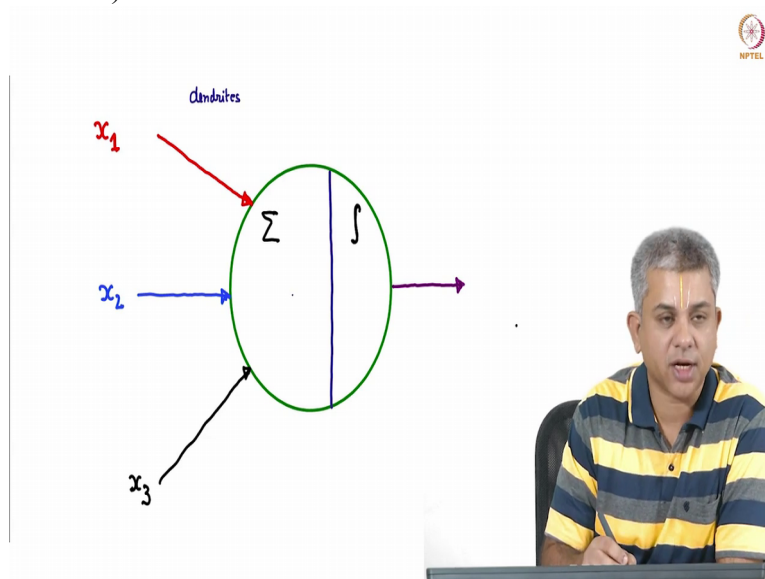
Σ

f

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

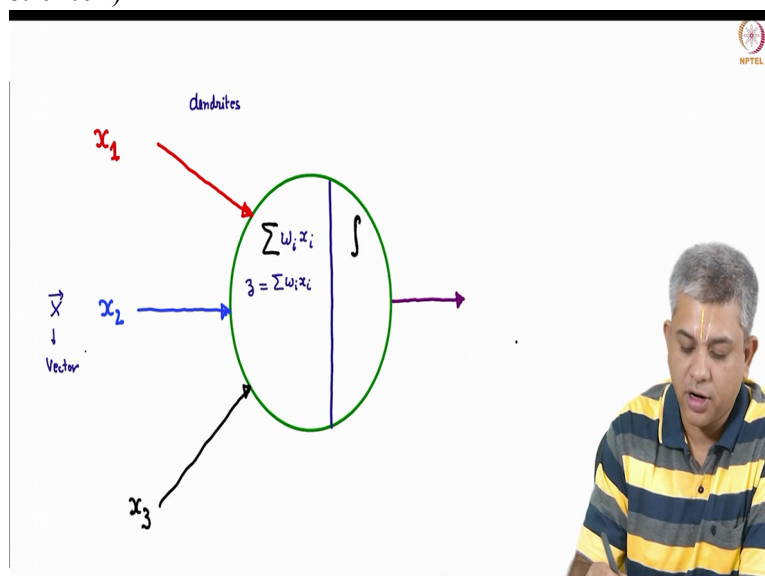
(Refer Slide Time: 01:25)



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 $\sum w_i x_i$ and this component we call z .

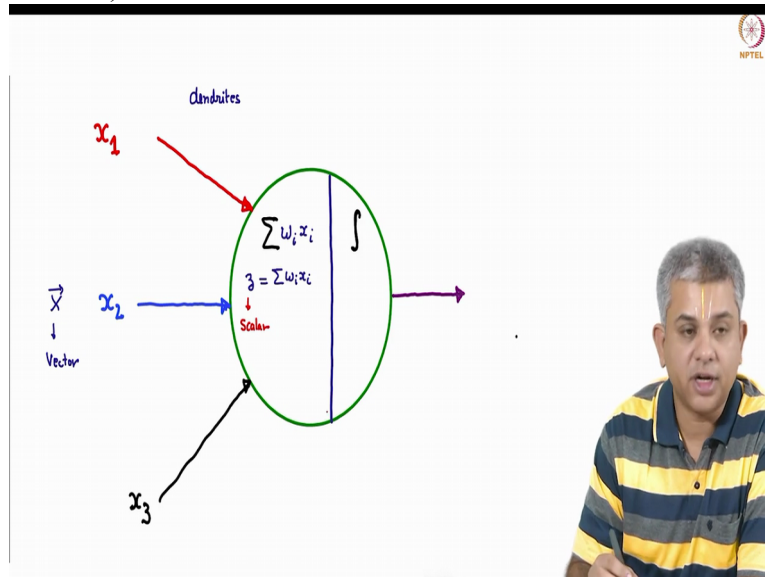
Remember the X that comes in is a vector

(Refer Slide Time: 02:02)



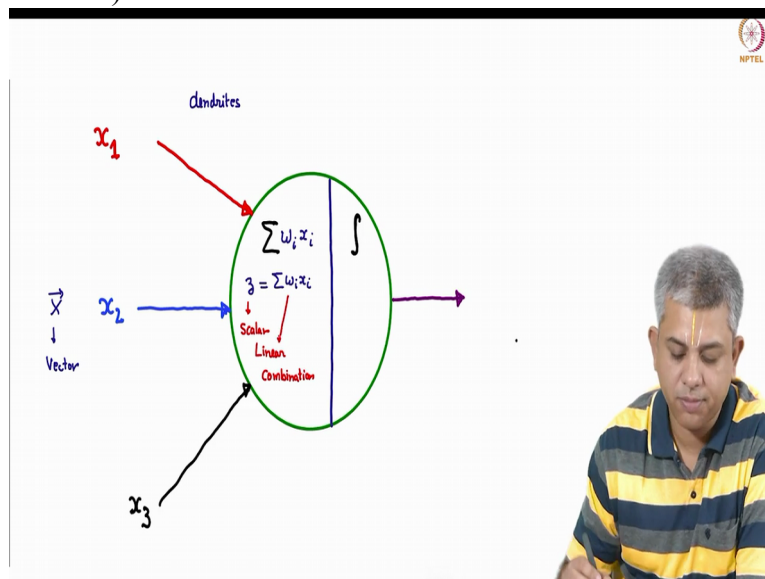
and what comes out, out of this linear combination z is a scalar. Till this point

(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)

The diagram illustrates a neuron model. On the left, three inputs are shown: x_1 (red arrow), x_2 (blue arrow), and x_3 (black arrow). x_2 is also labeled as a "Vector". These inputs are multiplied by weights w_1 , w_2 , and w_3 respectively. The weighted inputs are summed to produce the net input $z = \sum w_i x_i$. This summation is labeled as a "Linear Combination". The net input z then passes through a nonlinearity function σ , which is labeled as "Nonlinearity". The output of the neuron is shown as a purple arrow pointing to the right. The word "dendrites" is written at the top of the diagram. An NPTEL logo is visible in the top right corner.

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 $\frac{1}{1 + \exp(-z)}$

(Refer Slide Time: 03:02)

This diagram is identical to the one above, but it includes the formula for the sigmoid function: $\sigma(z) = \frac{1}{1 + \exp(-z)}$ written in red above the nonlinearity symbol σ . The rest of the diagram, including the inputs, weights, summation, and output, remains the same. An NPTEL logo is visible in the top right corner.

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

(Refer Slide Time: 03:17)

The diagram illustrates a neuron model. On the left, a vector \vec{x} is shown with components x_1 , x_2 , and x_3 . x_1 is labeled as "Clonduites". These inputs are combined into a sum $z = \sum w_i x_i$, which is labeled as "Scalar-Linear Combination". This sum z is then passed through a nonlinearity function $\sigma(z) = \frac{1}{1 + \exp(-z)}$. The diagram is annotated with "Clonduites", "Scalar-Linear Combination", and "Nonlinearity".

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

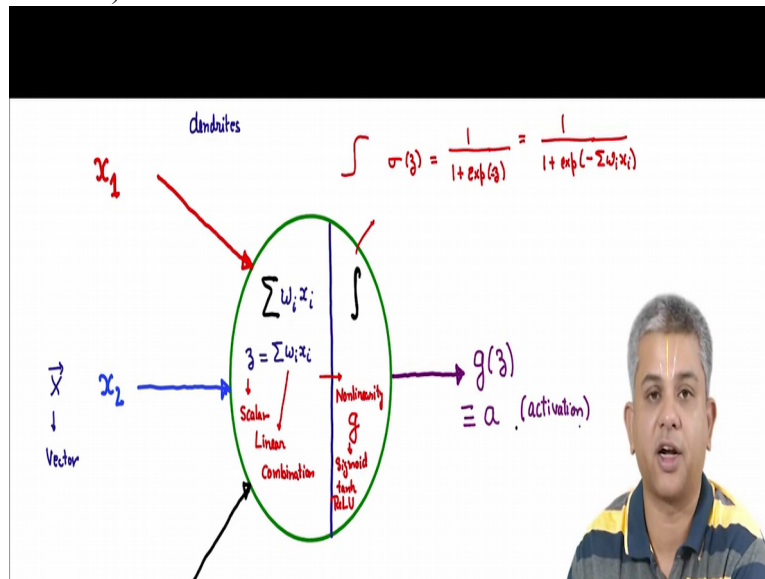
(Refer Slide Time: 03:35)

The diagram is identical to the previous one, but the nonlinearity function g is annotated with "Sigmoid", "tanh", and "ReLU".

ReLU. So any of these outputs, any of these nonlinearities could be used.

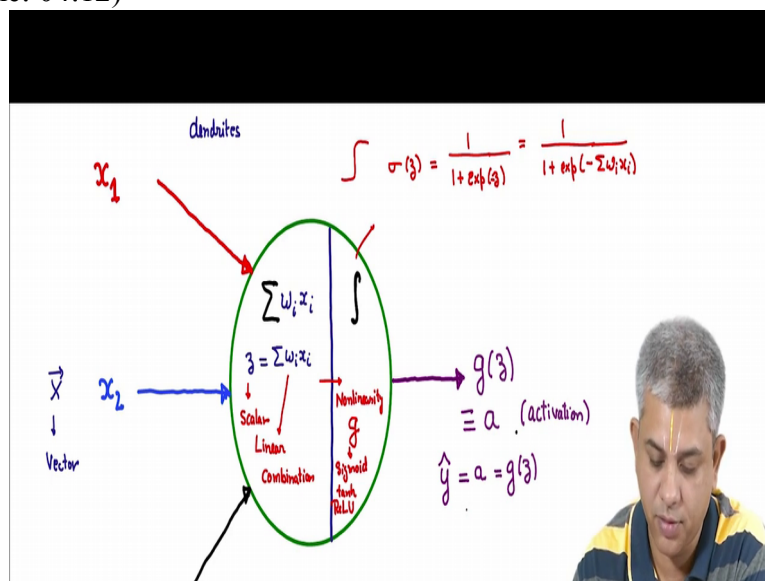
Now finally after all this, what comes out is g of z . This is denoted by a , also labeled the activation.

(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)

