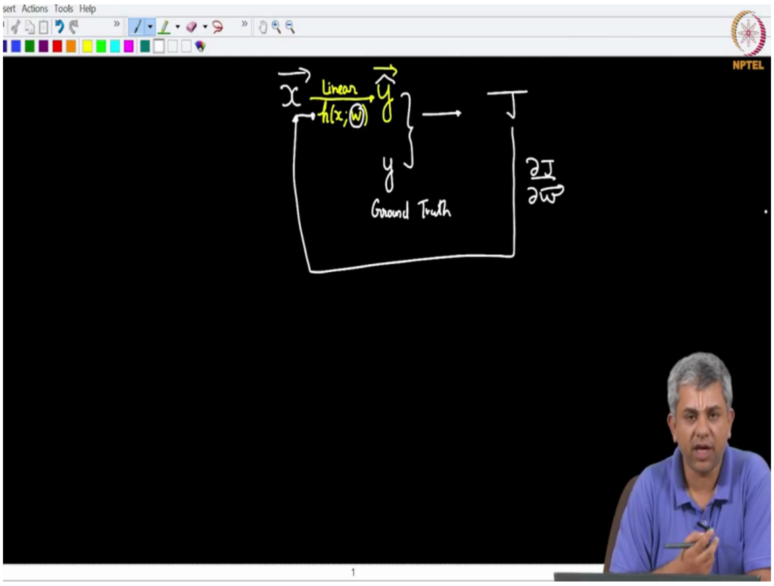
Machine Learning for Engineering and Science Applications
Professor Dr Balaji Srinivasan
Indian Institute of Technology, Madras
Department of Mechanical Engineering
Introduction to Week 5 (Deep Learning)

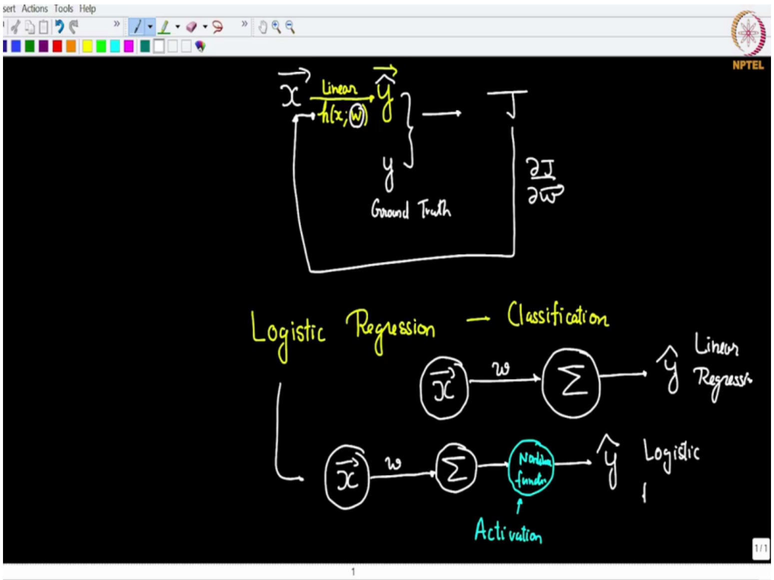
Welcome to week 5, this week we will be introducing you to deep planning. Last week we saw linear regression which was a simple model that connected input to output via linear model. This week we will be looking at more models, one is something called logistic regression which is a classification model and the next is the neural network and subsequently we will go into what are called the neural network which is usually the terminology used for deep learning.

(Refer Slide Time: 0:50)



So recall that last week what we looked at was given some input vector X, if you want to connect it to some output Y via a linear model, for some reason you think that the connection between input and output, the regression connection is actually through a linearity connection ok. In that case let us say our h of x with parameters W is assumed to be a linear model, in that case all you do is you take a hypothesis function h of x, you say that my guess is Y hat, you will have already got some ground truth Y and using these two you calculate the cost function J and you feed it back so as to improve W ok by looking at Del J Del W. So this is what we saw in the last week and we saw that the same model could be used for linear, quadratic, cubic and any type of polynomial fit.

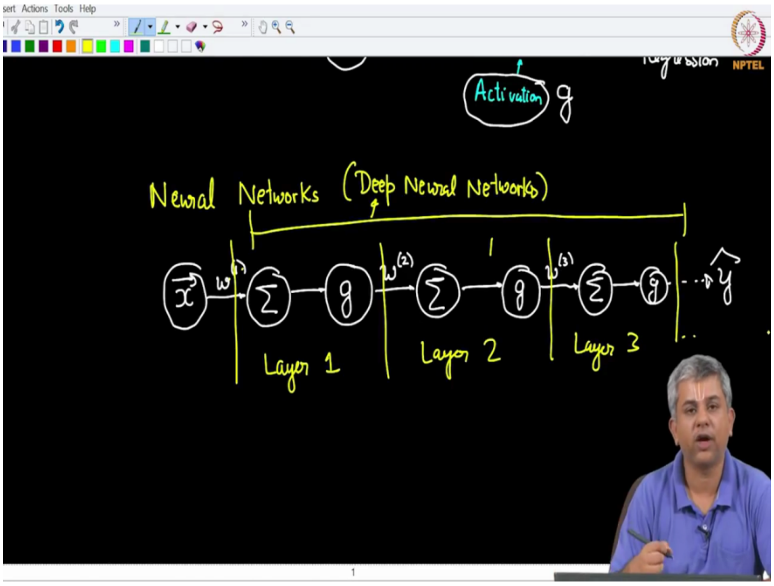
(Refer Slide Time: 3:38)



This week we will just do us very simple change to this, actually you will be very surprised that how simple this change is in case you have not seen this before and you will be able to achieve almost universal computation. So the first example of this that we will see is something called logistic regression which is a classification algorithm, this is what we see first this week, let me point out how we will do that and you will see the details in the next few videos. All we do is a very simple thing, remember in linear regression you had X, I showed you a notation, you multiply by W and run it through a submission and you get Y hat, this was linear regression.

What we will see this week is logistic regression is a very small change over this, you take X, again the same parameters W, run it through a submission and we add a one small change, we add a non-linear function. This is called a non-linear activation function and this gives our Y hat, this is called logistic regression for certain choices of activation functions. So please remember this name activation function, all an Activision function is after your linear combination you add nonlinearity over this ok, so we will typically denote the non-linear activation function by G ok so G stands for some non-linear function.

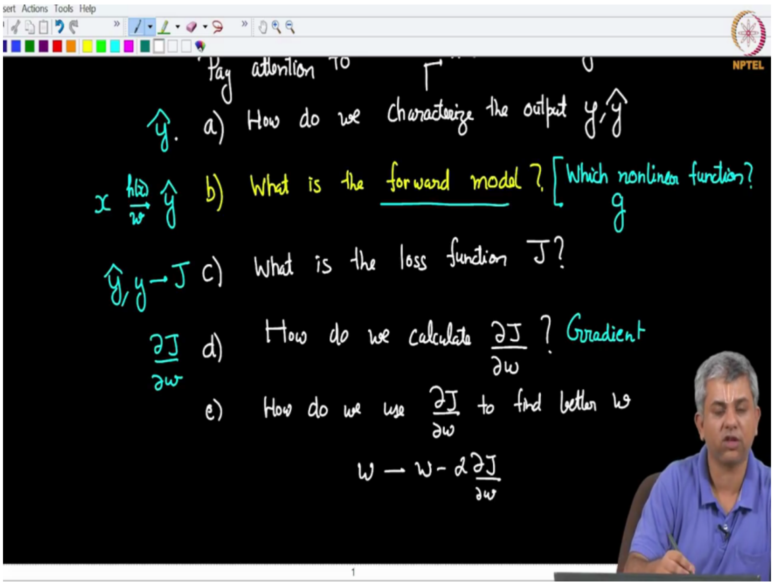
(Refer Slide Time: 4:11)



All this is put together let us achieve classification in a very simple way, you will see this in the future videos. After this we will look at neural network, we will also look at what are called DNN or Deep Neural Network hence the name deep learning, and once again let me give you a schematic of what this is and this is actually a very straightforward as well. So you take your X, run it through a linear combination with some weight W, run it through a non-linear function G ok. Then run it through another linear combination with some other weights let us call them W1, some other weights W 2, and another non-linear combination G and so on and so forth Sigma G and finally you get your output prediction Y hat.

Now all this put together is called a Deep Neural Network that is it, it is a very simple idea. Each of these combinations of Sigma and G is called a layer, this is layer 1 hands W 1, layer 2, layer 3 so on and so forth, so a deep neural network is the one which has more than one layer, so that is all that is if you have more than 1 layer is called a deep network. So this in a nutshell is all there is to learn with deep learning at least in terms of simple implementation, let me come to a few more details that we will see as we move on through this week. You will see as we go ahead that we need to pay attention in any model of this sort to the following factors; 1st one is how do we characterize the output Y or Y hat.

(Refer Slide Time: 7:55)



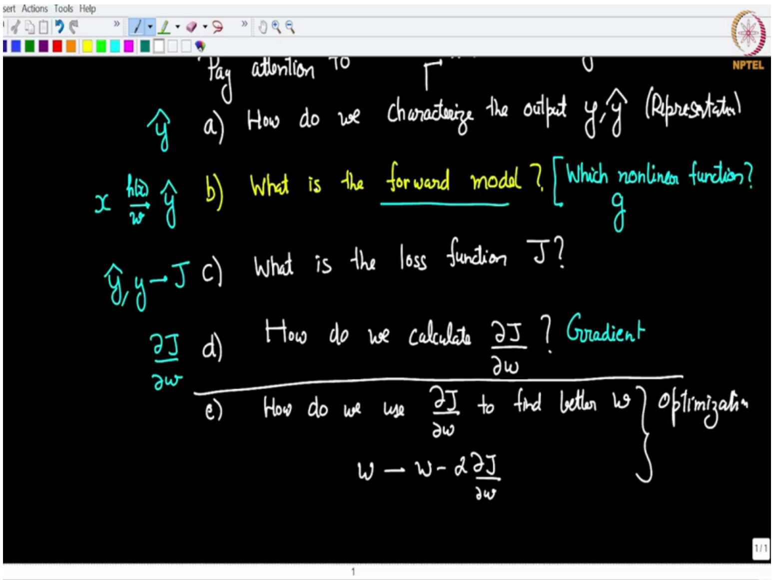
So as we saw in the last week, suppose the output itself is a number okay so the number for example last time was we had our Alpha which was the coefficient of thermal expansion or you could have a house price or something that you are trying to actually predict, a number that you are actually trying to predict as a regression problem. If the output is a number, it is easy to say what Y is. Sometimes however you are looking at slightly more qualitative thing for example, success or failure of a machine part or it could be a classification for example, this is a cat or dog or a horse okay so you have something of that sort.

Remember that what we saw in the 1st week, every single thing that a machine learning algorithm does maps one set of numbers to another set of numbers. So when I say characterize, it means how do we assign numbers, so how do we assign to something a number to something like a cat or dog. So we will come up with a very simple idea, for a regression problem usually it is obvious, for classification problem how do we do for a certain a few cases it can be slightly subtle so we will look at that so 1st thing. Next is of course what is the feed forward model okay, I have already spoiled that for you, I have told you that for logistic regression all it is is for a linear combination followed by a non-linearity.

Which nonlinearity do we use that usually also plays a part, for a neural network it is linearity-nonlinearity, linearity-nonlinearity so on and so forth that is usually what happens in a neutral network ok. And auxiliary problem is which non-linear function ok, so how do we choose G? We will look at some rules of thumb again there is no hard and fast rule, we will give you some common choices that are available within the literature this week. The 3^{rd} thing is what is the loss function?

Or just to recapitulate you have Y hat, how do I give numbers to Y hat, the 2^{nd} thing is how do I go from X to Y hat given that I had decided some numbers for X, some numbers for Y, what the forward model decides is what is this function form. Remember we distinguish between the functional form and functional parameters so how do we take this functional form from input to output that is what we have to choose next. The 3^{rd} you have to choose is, given that you have Y hat and you have some idea output Y how do I get J ok, so that is the 3^{rd} thing that you have to decide.

(Refer Slide Time: 11:47)



Fourth, what you have to find out is how do we calculate Del J by Del W, in other words this is the gradient problem ok. There is a 5^{th} problem which will not be discussing very much which is how do we use Del J Del W to find better W, as of now more or less what I am assuming is to be will simply use some form of gradient design and as Dr ganpati had told you last week, you can actually use several variants of this, typically pure gradient design is almost always never used in practice.

We use some option or the other which is a slightly modified version of gradient design but for our understanding of the algorithm what we will split our problem into is finding these 4 things; a number representation okay characterized or a representation of Y hat, so the representation problem, the forward model, the loss function and the gradient. If you know these 4 things, you have a deep learning model one way or the other you can always get a deep learning model, so this is just an optimization problem after that okay. So I will just request you to pay attention to these 4 as we go over logistic regression as well as deep neural network, you will see these 4 once again in different forms very move onto the next few weeks which will be called Convolution neural networks and recurrent neural networks, but did you get these 4 during this week, you have a very good picture of what is actually needed to setup a deep learning model. Thank you.