**Deep Learning**
Mitesh M. Khapra
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Lecture – 01
Chapter 5: Faster, higher, stronger

So now, so this is what the progression was right that, in 2006 people started or the study by Hinton and others led to the survival, and then people started realizing the deep neural networks and actually we use for lot of practical applications and actually beat a lot of existing systems.

But there are still some problems and we still need to make the system more robust, faster and even scale higher accuracies and so on right.

(Refer Slide Time: 00:41)
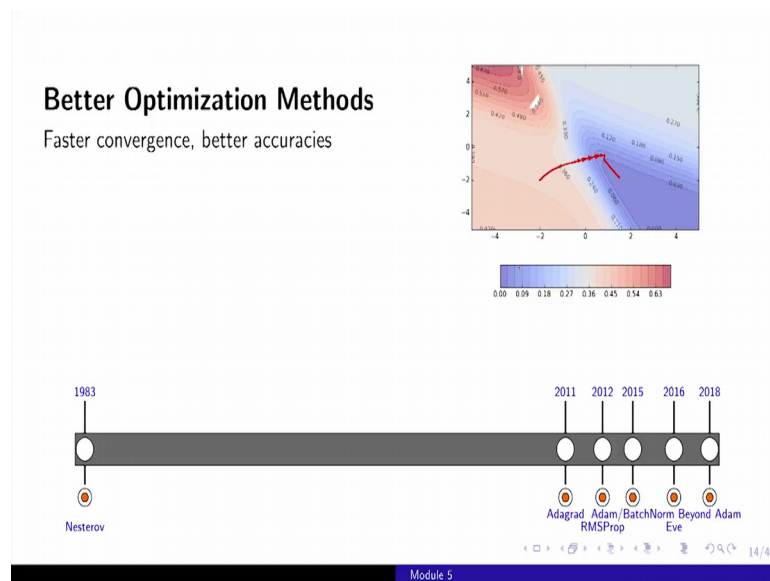
So, in parallely while there was lot of success happening from 2012 to 2016 or even 2010 to 2016. In parallel there will also a lot of research to find better optimization algorithms which could lead to better convergence, better accuracies.

And again some of the older ideas which were proposed way back in 1983. Now this is again something that we will do in the course. So, most of the things that I am talking about we are going to cover in the course. So, we are going to talk about the image the

challenge, we are going to talk about all those networks the winning networks that I had listed there Alex net ZF net Google net and so on.

We are going to talk about Nesterov gradient descent which is listed on the slide. And many other better optimization methods which were proposed starting from 2011. So, there was this parallel resource happening while people were getting a lot of success using traditional neural network, they are also interested in making them better and robust and lead for lead to faster convergence and better accuracies and so on.

So, this led to a lot of interest in coming up with better optimization algorithms, and there was a series of these proposed starting from 2011. So, Adagrad is again something that we will do in the course, RMS prop, Adam Eve and many more right. So, many new algorithms I have been proposed, and in parallel a lot of other regularization techniques or weight initialization strategies have also been proposed for example, Batch normalization or Xavier initialization and so on. So, these are all things which were aimed at making neural networks perform even better or faster and even reach better solutions or better accuracies and so on, this all that we are going to see in the course at some point or the other.