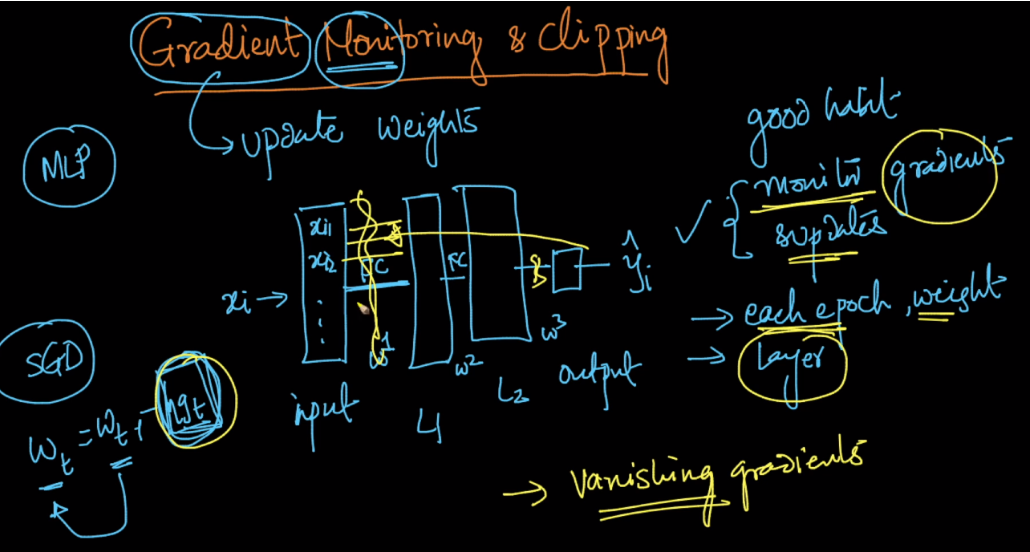
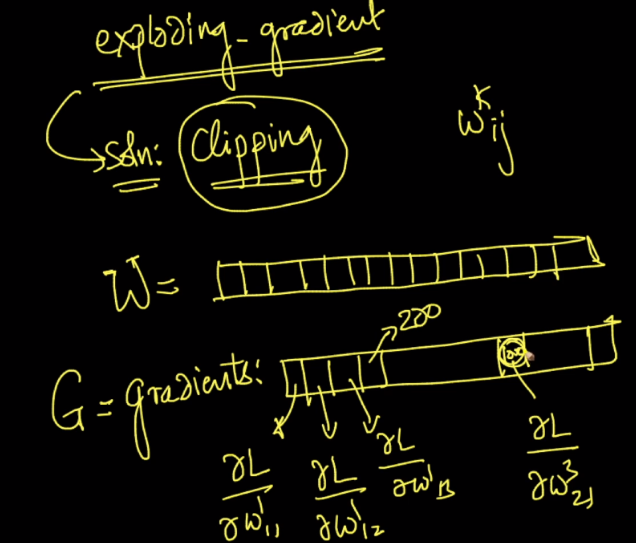
**Gradient Monitoring:**

We should monitor gradients at each epoch with weight and at each layer to know whether the process is going fine or there is a problem of gradient vanishing and gradient exploding.

If there is a problem of **gradient exploding** then we can overcome this using **Clipping**





**Clipping:**

Let’s for a we have gradients in an array, now before using gradients for updating weights we do the following thing as given in below image

Basically we are dividing each gradient with L2 norm of gradients which square root of sum of square of all gradients. And after dividing we multipy it with threshold.

