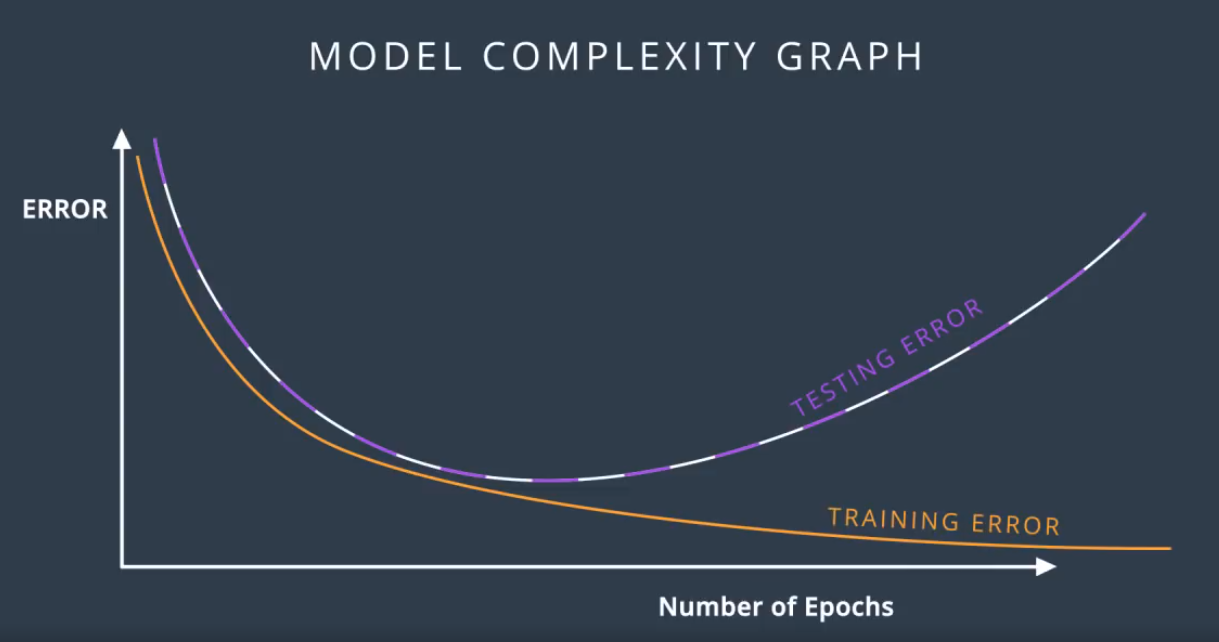
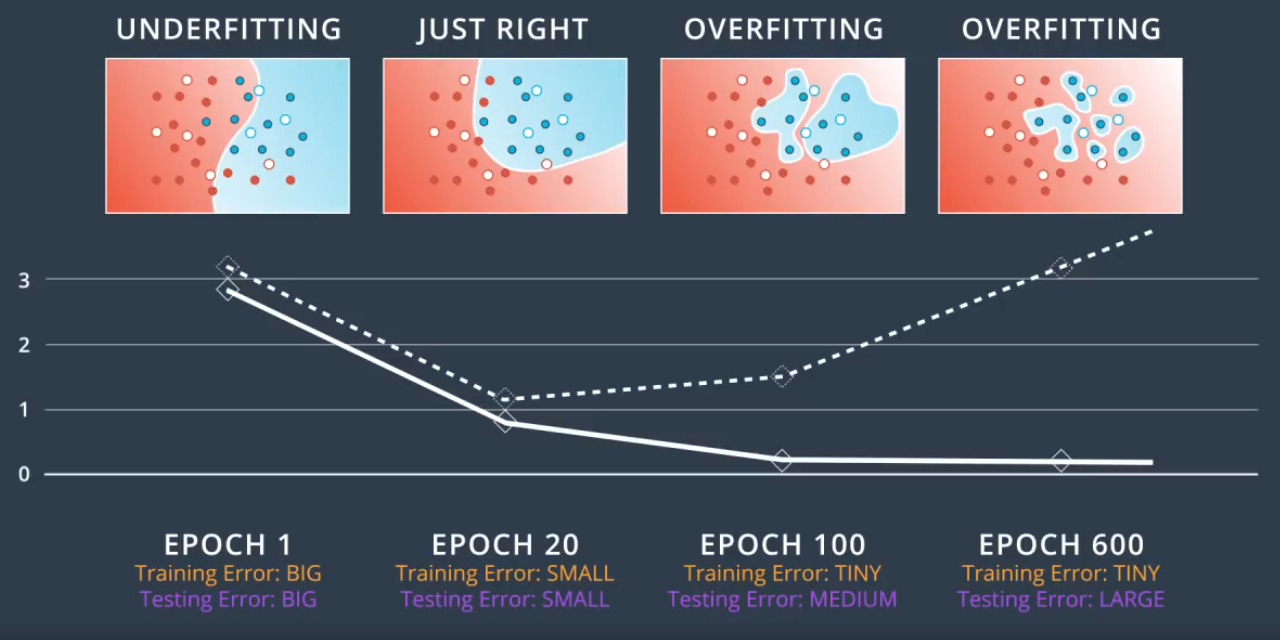
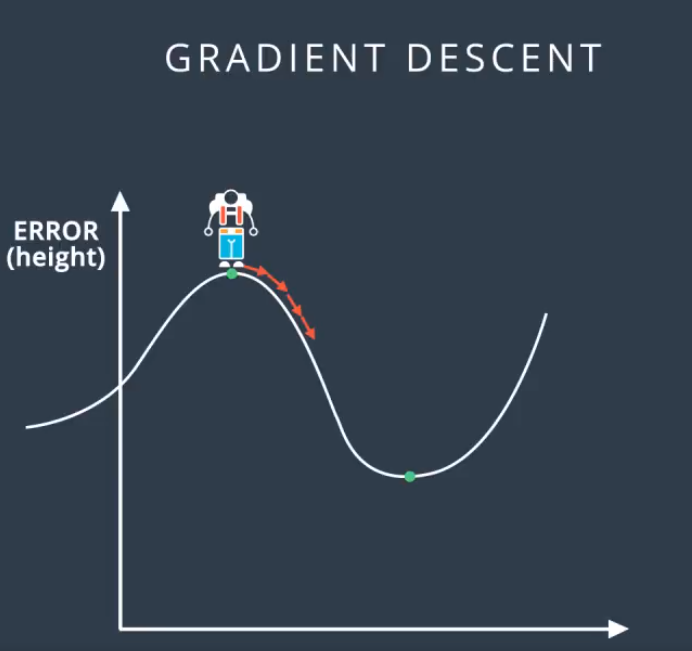
@Linear Algebra:

@Training:



**Dropout:** Buffed in all areas of body rather than one hand.

Weights get **buffed**. Randomly we turn off the notes so other nodes take up the process.

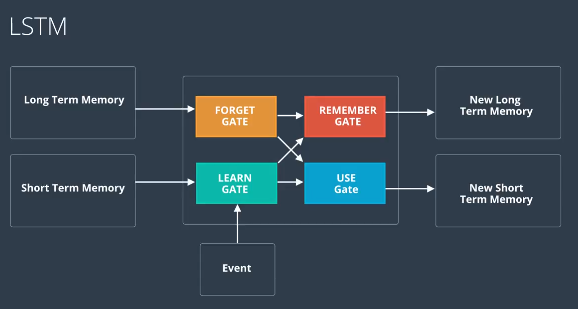
Dropout = at each epoch 0.2 each node gets off and on average will make sure all the nodes are

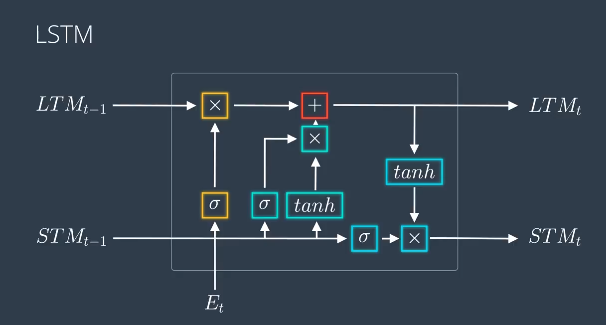
Vanishing gradient:

Problem Tiny changes on weights. Never able to find minimum.

Solution: ReLU,

**@LSTM**



**Why like this only it works:  
Try to come up with a better architecture**