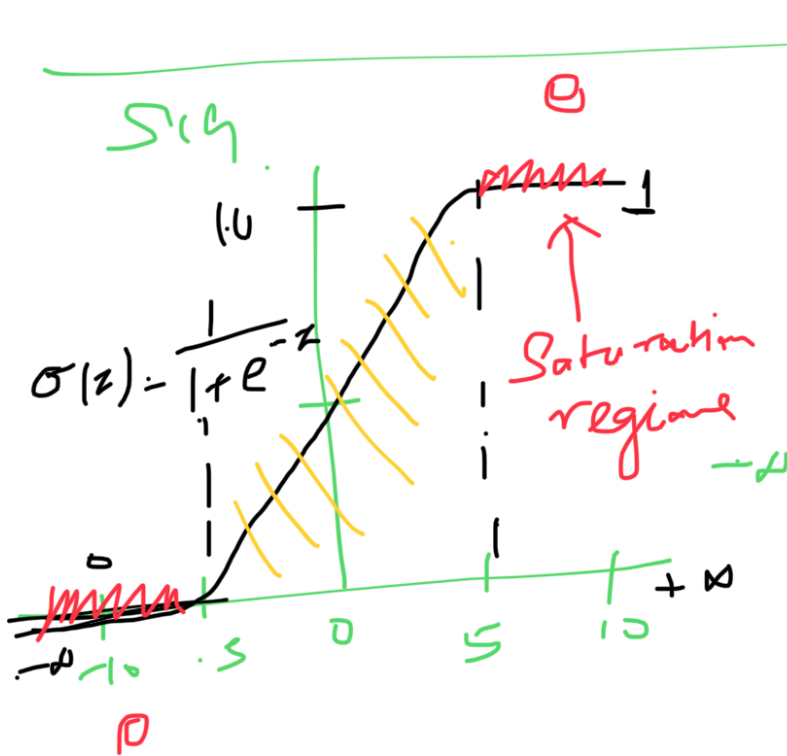
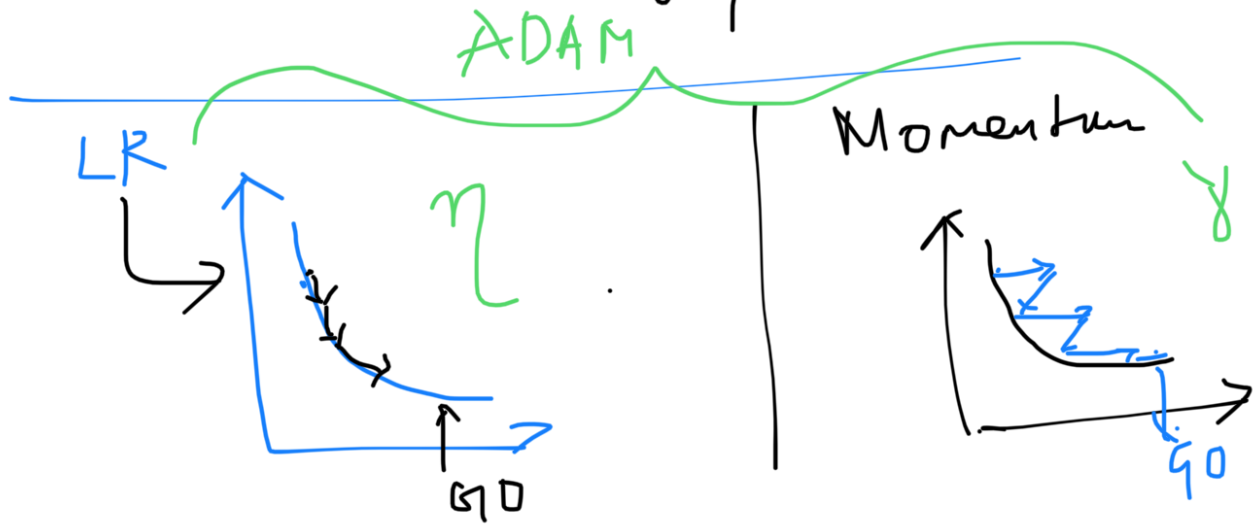
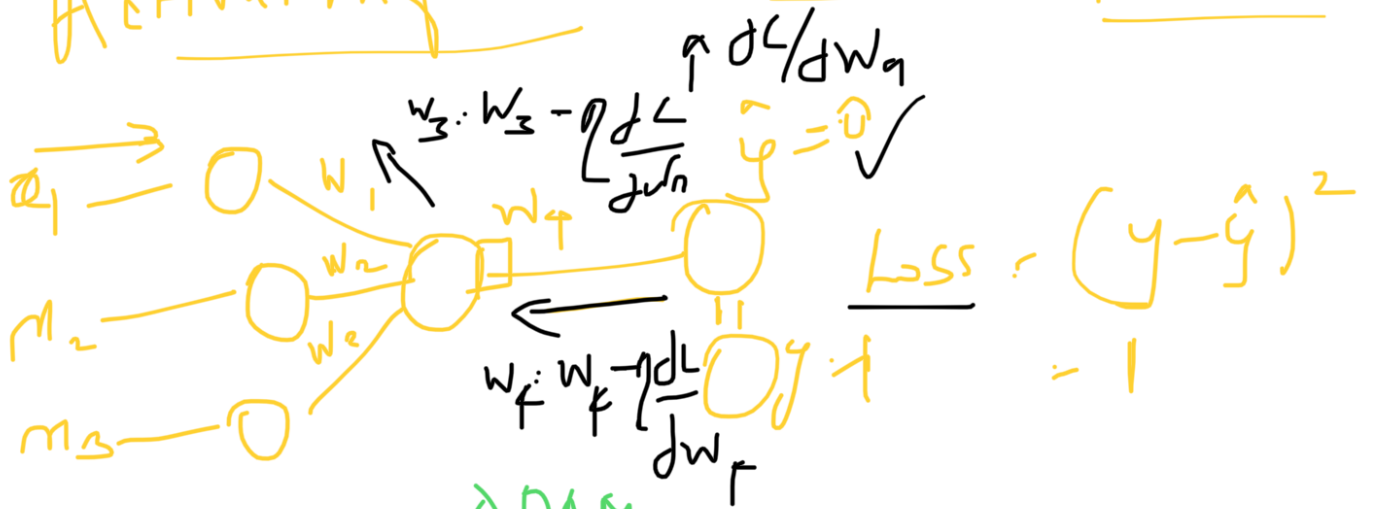
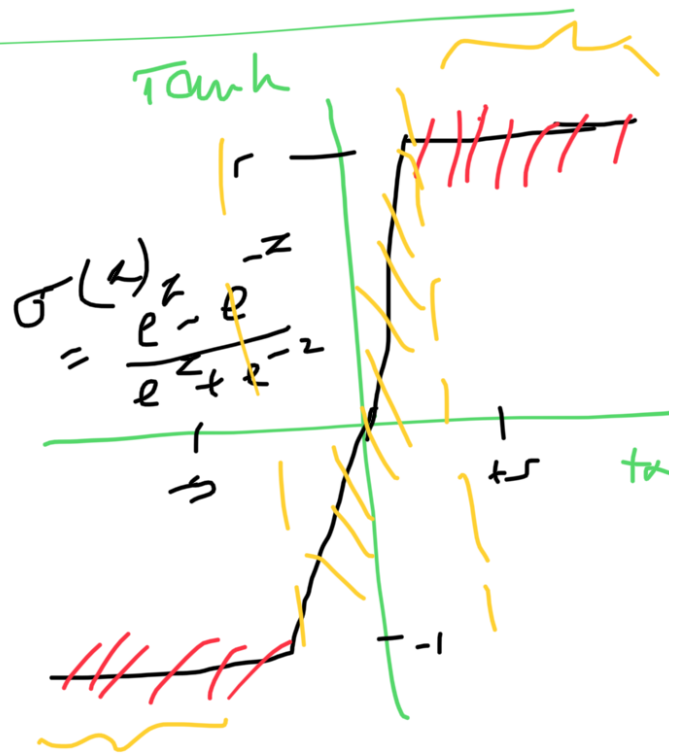


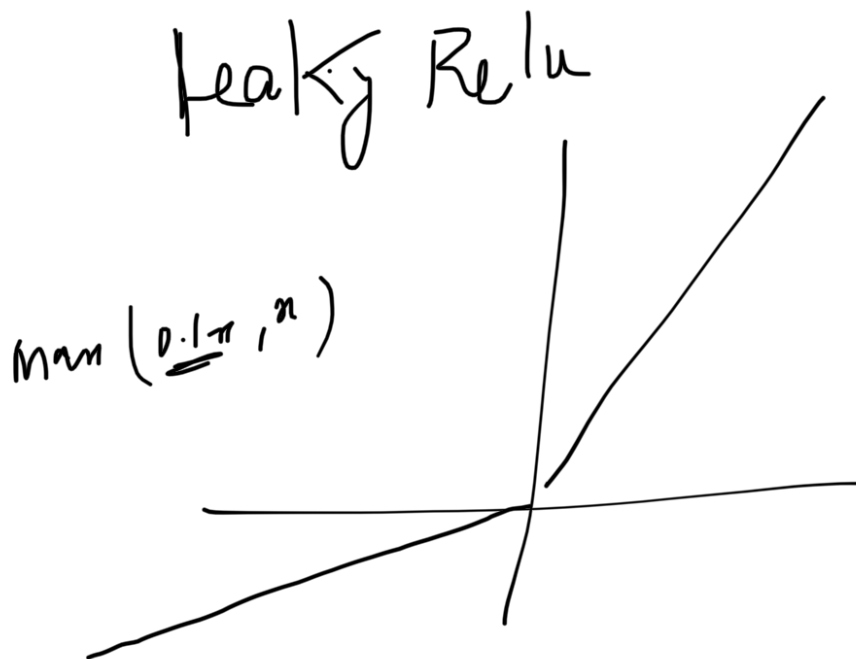
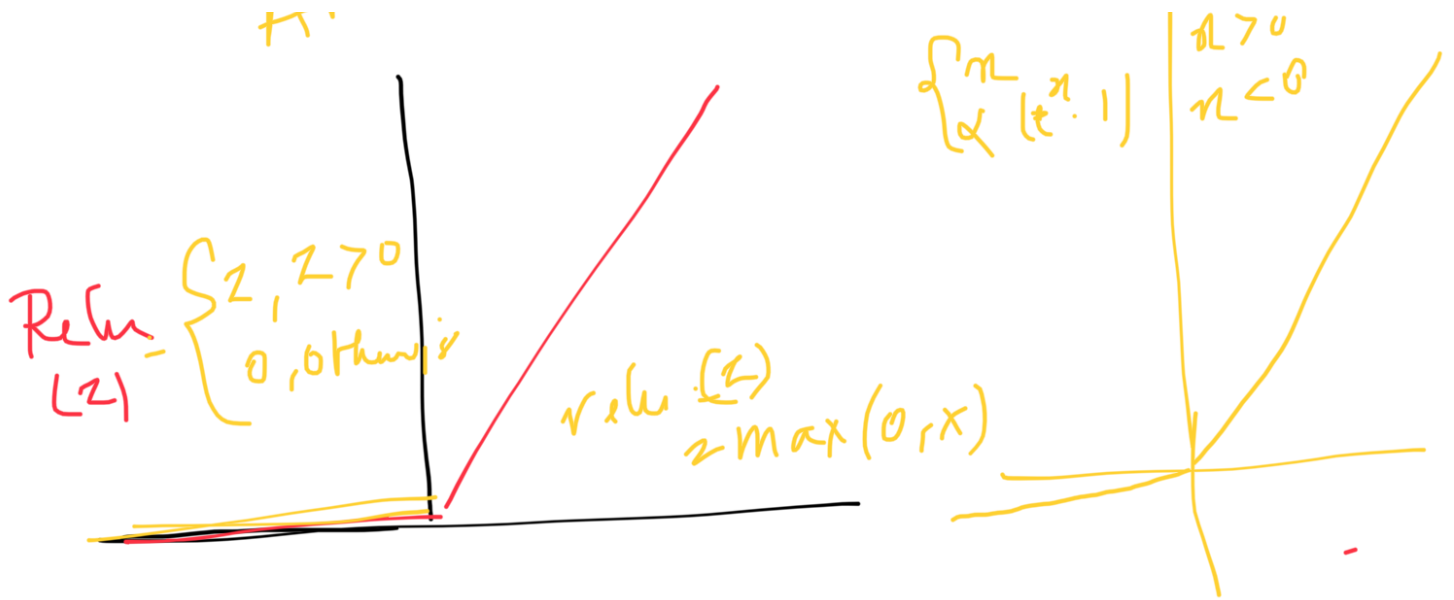
Activation functions - NN=3 \rightarrow Gradient



Non-Sat RF



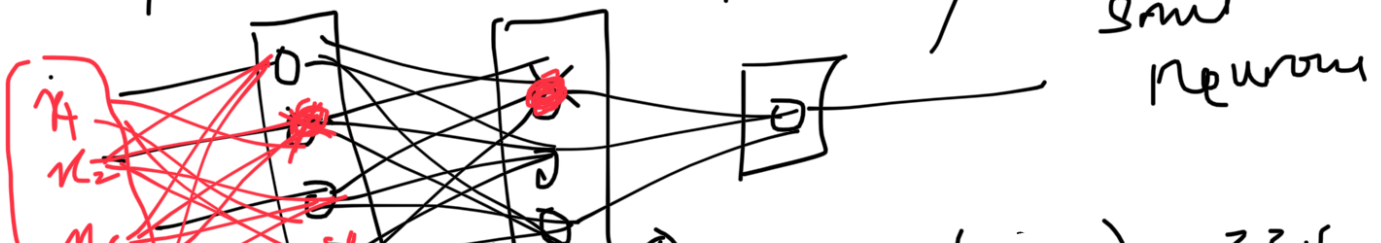
ELU



overfitting \rightarrow The model performs well in Tr \uparrow
 Poor test \downarrow

L_1 & L_2 Norm \rightarrow ①

Dropout Regularization \rightarrow Too much info.
 intentionally disable some neurons





dropout (0.335) - 55.1.

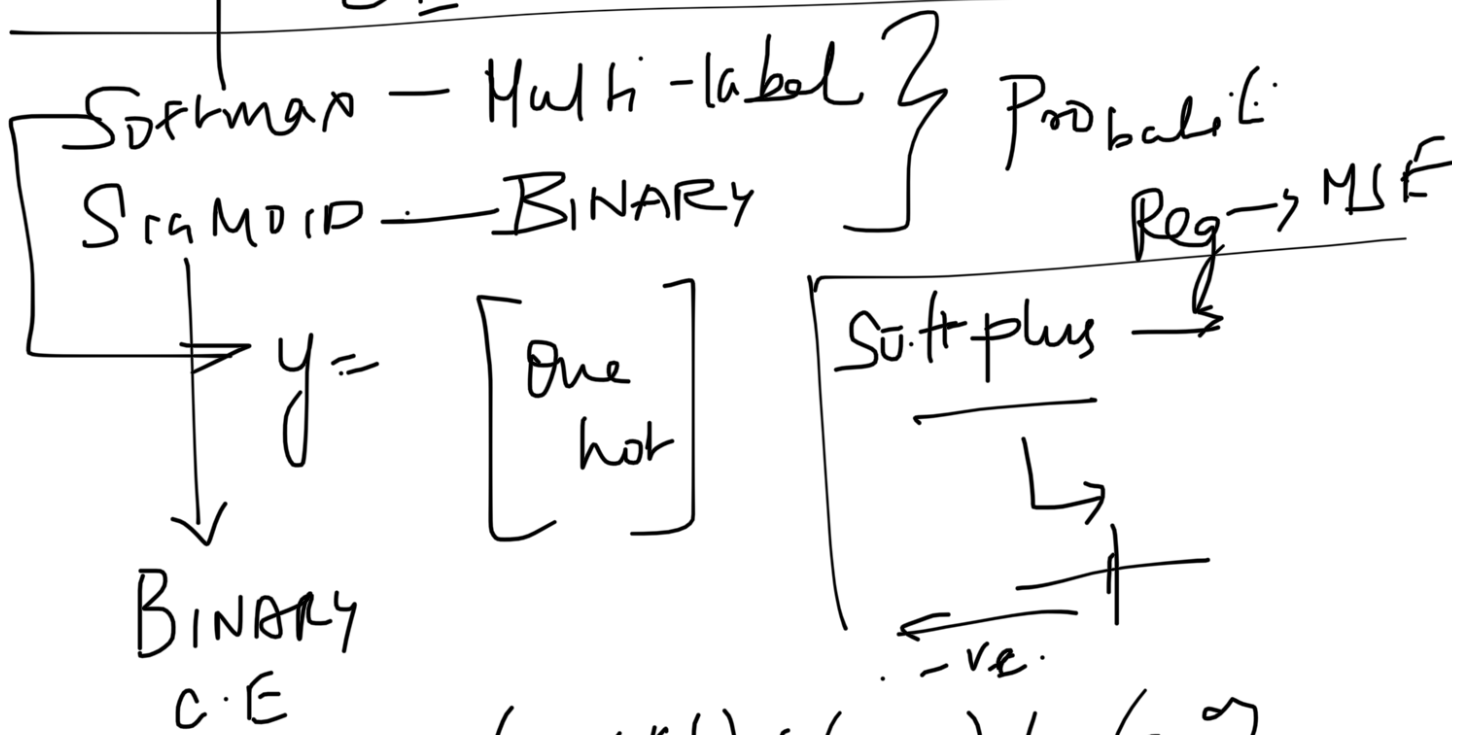
Early Stopping Epochs = 50
Monitor: val - acc

Patience: 3			train_acc = 0.85	
35:	val - acc = 0.83		"	0.84
36:	" : 0.83		"	0.90
37:	" : 0.83		"	

Saturates

Imp.

Categorical
C.E



$$\Rightarrow y (\log(\hat{y}_1)) + (1 - y) \log(\hat{y}_2)$$