Today's Agenda

1) Vanishing Gradient / Exploding
2) Improving NN
3) Early Stopping
4) Dropouts

Vanishing / Exploding Gradients

1) Weights / Bias → Trainable Params

While Training → update → loss minimum
local minima

Weight Update

$$L = \hat{y} - y$$

$$w_n = w_0 - \eta * \frac{\partial L}{\partial w}$$

(η → L.R)

$\frac{\partial L}{\partial w}$, [high, low]

LR = 0.01

$= .1 \times .1 \times .001 \times .01$

$= .00000001$

$W_0 = 1$

New Weights

$$W_n = W_0 - .01 \times 0.00000001$$

$$= 1 - .0000000001$$

$$W_n = .0000000099$$

$W_n \approx W_0$ * Problem of Vanishing gradient

* $\boxed{W_n = W_0}$ No weight updates

↓

Cannot Reach local Minima

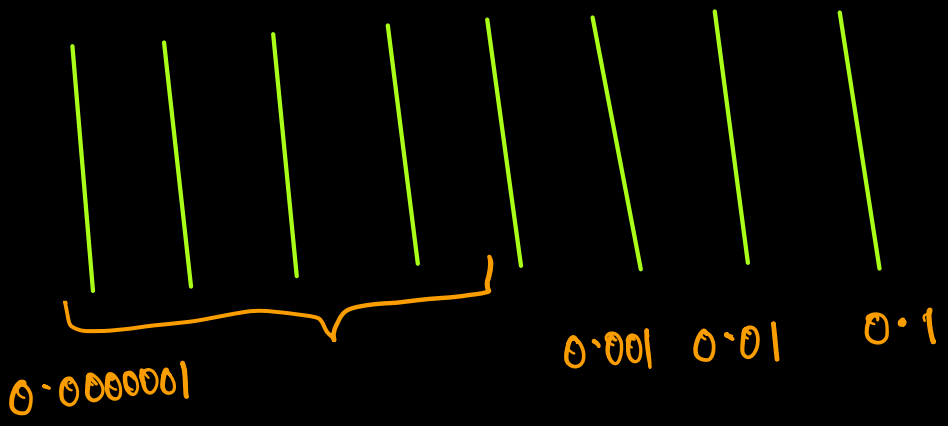
Loss values will be stuck at a certain point without any changes.

EPOCH 1 :- 6.5

101 :- 5.7

201 :- 5.7

301 :- 5.7



Improving Neural Networks

1. Finetune NN hyperparameters
   1) Epochs
   2) Hidden layers
   3) No of Neurons
   4) LR, optimizer, AF
   5) Batch Size
2. Few Major Problems :-
   1. Vanishing / Exploding
   2. Low Data
   3. Slow Training (* optimizers) (LR)
   4. Overfitting

Epochs

1) Sufficient (Not low - Not high)

Early Stopping

* Loss Tracking *

https://keras.io/api/callbacks/early_stopping/

```
keras.callbacks.EarlyStopping(
    monitor="val_loss",
    min_delta=0,
    patience=0,
    verbose=0,
    mode="auto",
    baseline=None,
    restore_best_weights=False,
    start_from_epoch=0,
)
```

Hidden Layers

3 - 8 Range

[5]

No of Neurons

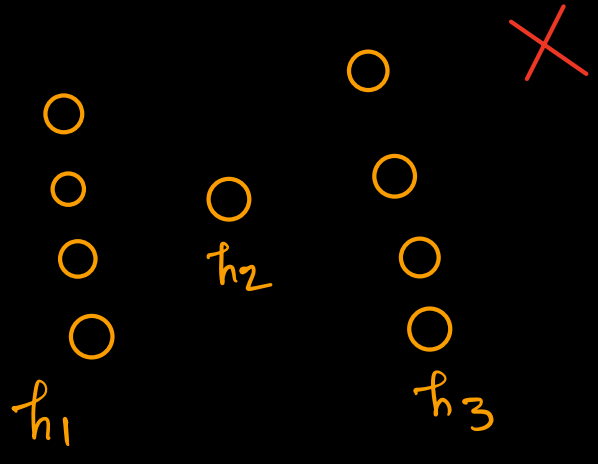
Ex:- 1 HL → 512 Neurons

* Bad Practice

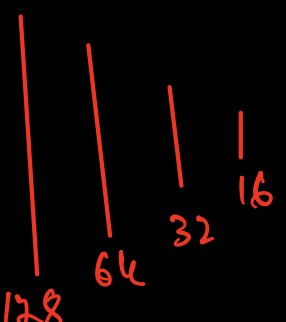
3 HL → 128 Neurons

* Good Practice

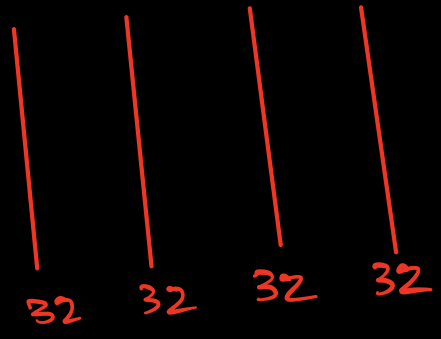
{ Low → Mid → High }



1) Pyramid Rule



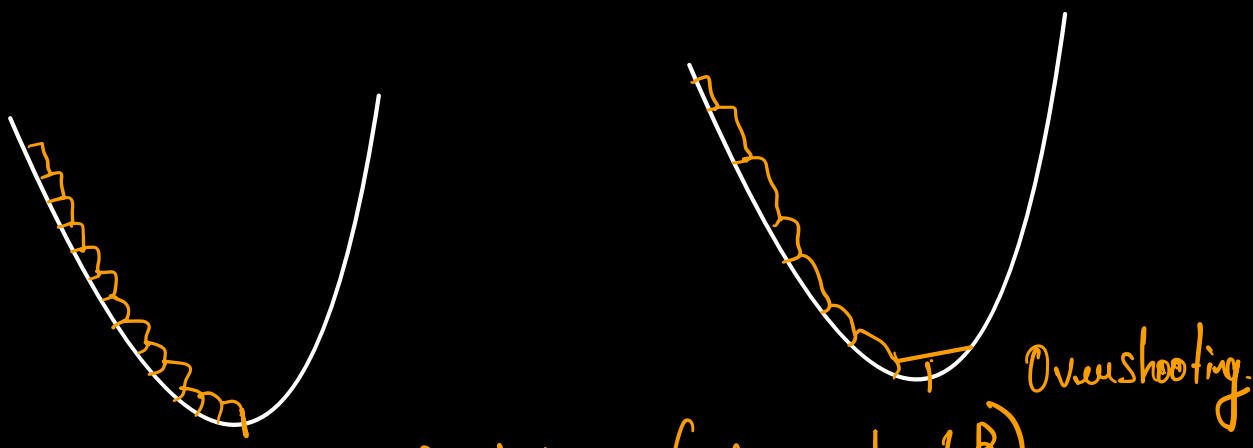
2) Box Rule



Learning Rate

0.01 → 0.00001

$$W_{new} = W_{old} - \eta \frac{\partial L}{\partial w}$$



LR Schedular (Warmup LR)

if epochs = 100

then LR = 0.001

elif epochs = 500

LR = 0.0001

else

LR = 0.000001