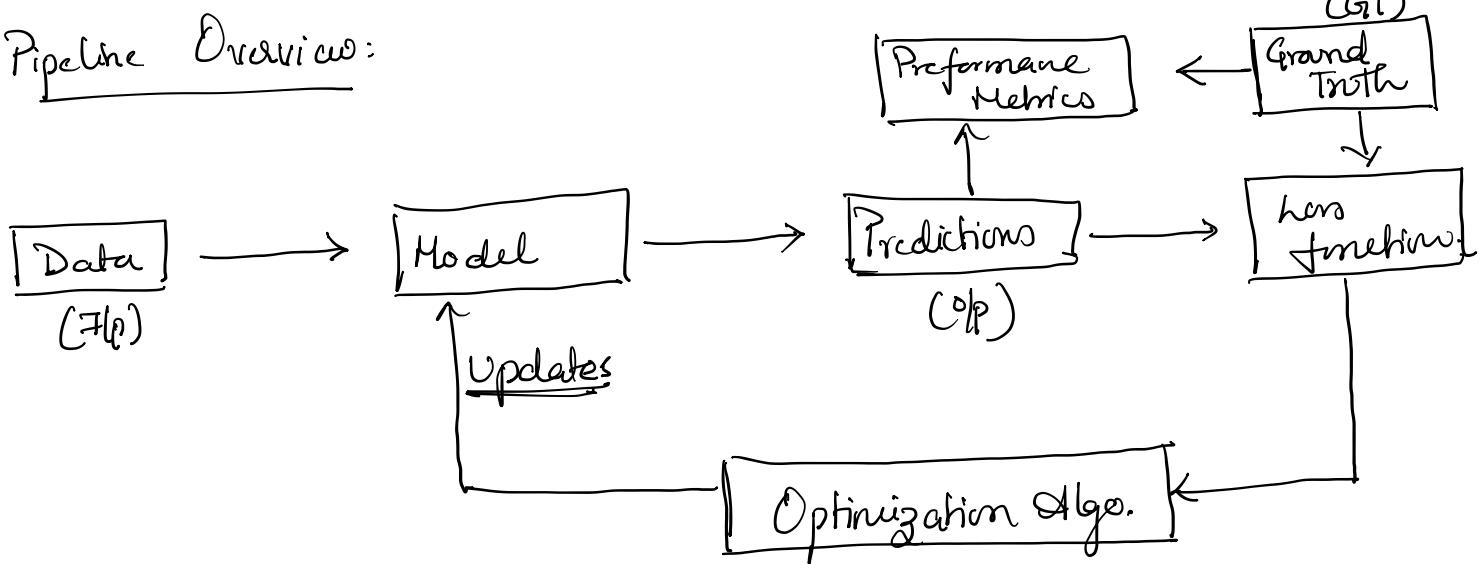
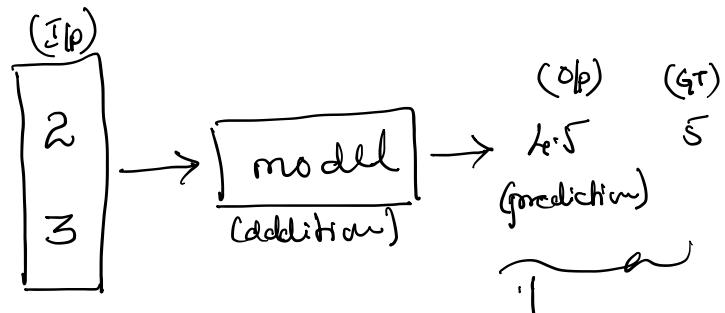


Pipeline Overview:



Eq: Task(T): addition. Aim: is to develop a model for task (T).

I/p: 2, 3
O/p: 5
(expected) \Rightarrow GT
O/p: 4.5
(computed/predicted)



In simple terms:

How far is your current prediction from expected o/p

Definition:
 Error function
 Cost function.
 Objective function.
 Loss function.

Error: \rightarrow ① What is the error. \rightarrow depends.

② How much is the error.

\propto Cost of correction \Rightarrow Cost fn.

Loss: How much are you lagging compared to your expected values.

Objective fn: Governs the learning as to how it is happening and what has to be optimized.

Generic Term ↗ { What is your goal / objective that is to be optimized (minimized / maximized) }

Types of loss fn:

GT

S

(a)

OP

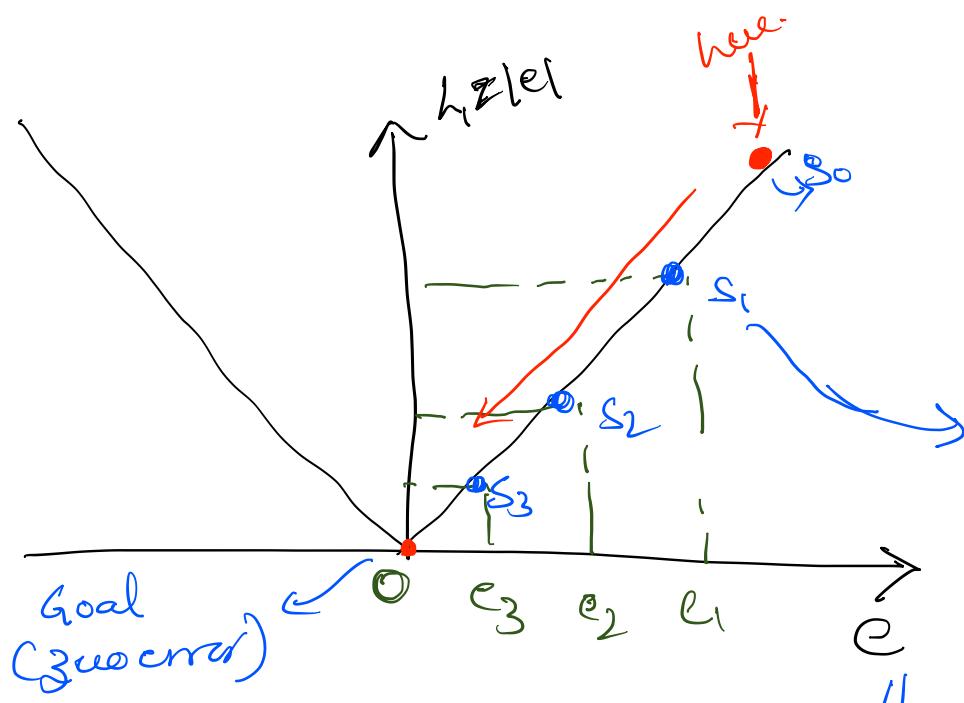
L¹

(b)

$$L_1 = \frac{|a-b|}{2} = 0.5$$

$$L_2 = \frac{(a-b)^2}{2} = 0.25$$

differed
how it
inputs
learning



Say $e = a - b$

slopes @ 0, 1, 2, 3

$$s_0 = s_1 = s_2 = s_3$$

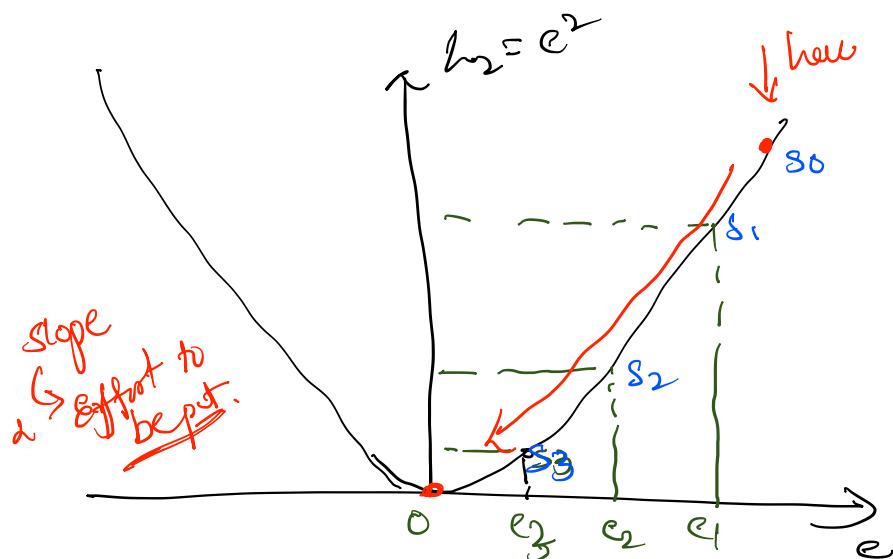
update rule:

$$\omega_i = \omega_0 - \eta \Delta \omega_0$$

descent. (mag + direction) gradient

(magnitude only) slope

gradient descent



slope
↳ effort to
be put.

$$s_3 < s_2 < s_1 < s_0 \Rightarrow \text{as } s_i \downarrow ; \Delta \omega \downarrow ,$$

Optimization Algo's

Different Algorithms
 ↓
 slight variations of
 the update rule.

- ① Gradient Descent.
- ② Stochastic " "
- ③ ADAM (uses velocity & momentum)
- ④ RMS Prop
- ⑤ AdaGrad.

Data:

{ anything that model takes as i/p }
 & { processes to give o/p }.

Dataset .
 (set of datapoints / samples)

divided
into
3 parts

Train set

Validation set

Test set

splits	
ST	7
:	:
1	1.5
:	:
1	1.5
10	10

Used for finalizing / finetuning
 hyperparameters &
 eval the model during
 training.

unseen by
 the model
 ↓
 but helps in
 choosing the best
 model.

What the
 model learns.