$(MSE, MAE, RMSE)$ [Cost function] → Performance Metrics

R^2 and Adjusted R^2

Salary

$J(\theta, \theta_1)$

→ Global Minima

=) Squaring

EXP   Salary
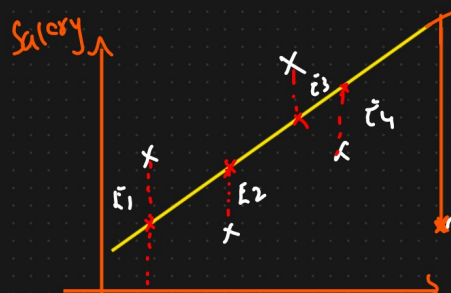
Increase error

penalizing

EXP the outlier

Impact
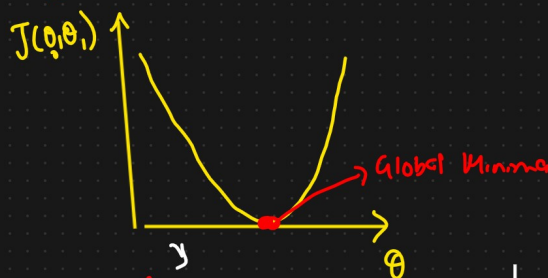
$NSE \uparrow \uparrow$

MSE

$$MSE = \frac{\sum_{i=1}^{n} (y - \hat{y})^2}{n}$$

=) Cost function ↓↓

→ Quadratic Equation

$ax + by + c$

$(a-b)^2 = a^2 - 2ab + b^2$

=) Differentiable

→ global Minima

Error less

convex function

Slope = 0

Converge

local
Minima

→ Global Minima

Non Convex functions
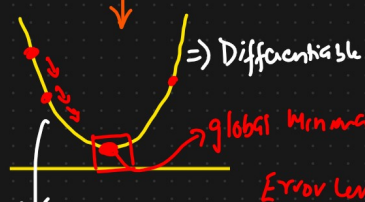
Salary (lakhs)

X   Y $(y_i - \hat{y}_i)^2$ (lakh)^2
               (MSE)

Advantage

① Differentiable at all point. ✓

② It has one local and one global Minima

③ Converge faster
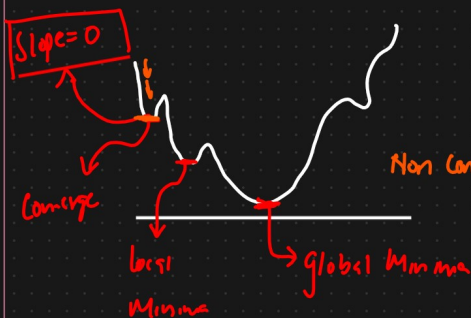
true   predicted

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$(lakh)^2$

Disadvantage

ⓧ Not Robust to outliers

ⓕ It is not in same unit.

Error c ?.5 =) $(lakh)^2$ ⇐

② Mean Absolute Error (MAE)   MAE↑↑

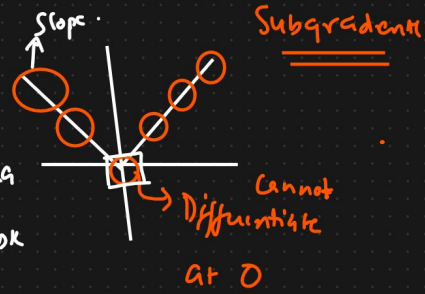$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y - \hat{y}|$$

lakhs
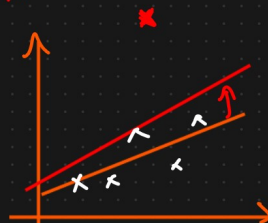
Advantage

① Robust to outliers ✓

② It will be in

    the same unit

Disadvantage

Ⓕ Convergence usually

    take more time. Optimiza
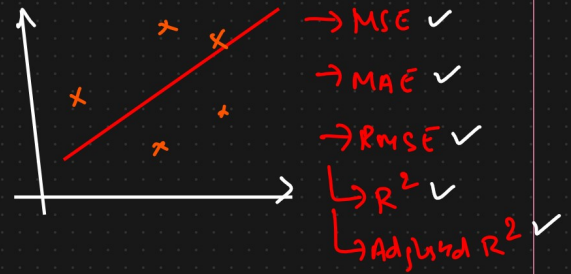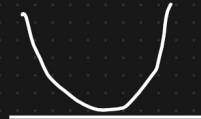
    -tion is a complex task

Ⓕ Time Consuming.

Slope ·

Subgradent

Cannot
→ Differentiate
at 0

---

③ RMSE { Root Mean Squared Error }

$$RMSE = \sqrt{MSE}$$

$$= \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

| MAE, MSE, RMSE |

Performance
↑ metric

→ MSE ✓
→ MAE ✓
→ RMSE ✓
↳ R^2 ✓
↳ Adjusted R^2 ✓

Advantage

① Same unit

② Differentiable

Disadvantage

     Not
Ⓕ Robust to outliers

MSE vs MAE vs RMSE ⎫
               ⎬
R^2 vs Adjusted R^2 ⎭