1. Mean squared Error (MSE)

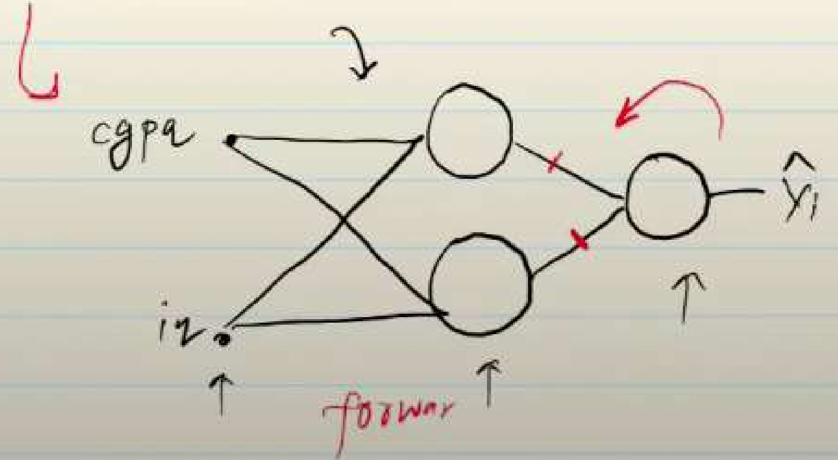
squared loss L2 loss

$\hookrightarrow (y_i - \hat{y}_i)^2$

$(\text{true} - \text{predict})^2$

$\rightarrow (6.3 - 6.1)^2 = -$

cgpa	Iq	package (y_i)	Prediction ($\hat{y}_i$)
6.3	100	6.3	$\boxed{6.1}$
7.1	91	4.1	4
8.5	83	3.5	3.7
9.2	102	7.2	7

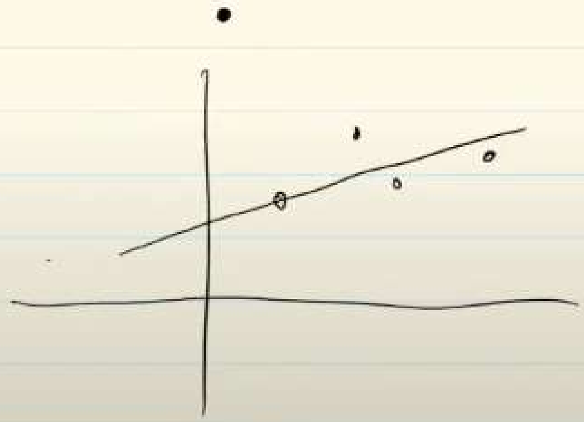


Advantages

1> Easy to intrepret
2> Diffrentiable (GD)
3> 1 local minima

Disadvantage

1> Error unit (squared) → diff
2> Robust to Outliers (Not)

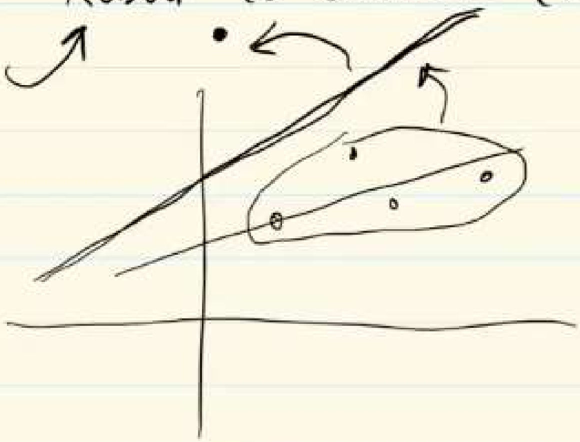


Advantages

1) Easy to intrepret
2) Diffrentiable (GD)
3) 1 local minima

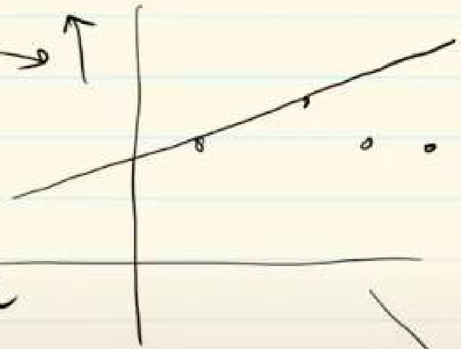
Disadvantage

1) Error unit (squared) → diff
2) Robust to Outliers (Not)



$$C = \frac{1}{n} \sum_{i=1} |y_i - \hat{y}_i|$$

punish

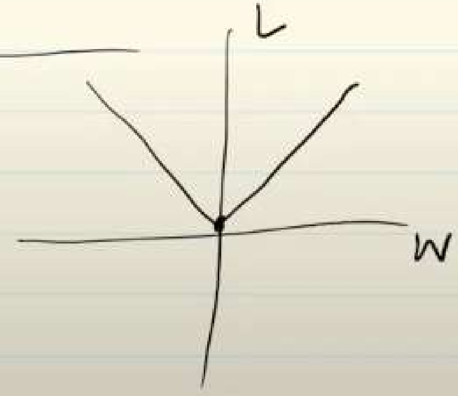


Advantages

1) Intutive and easy
2) Unit → same - y
3) Robust to outliers

Disadvange

1) Not diffrentiable



5. Categorical Cross Entropy [used in Softmax Regression]

→ [Multi-class] {classification}

1 point

$$\boxed{L = -\sum_{j=1}^{K} y_j \log(\hat{y}_j)}$$

where K is # classes in the data
↳ 3

cgpa	iq	placed?	Yes	No	maybe
8	80	Yes 1	1	0	0
6	60	No 2	0	1	0
7	70	Maybe 3	0	0	1

1 point

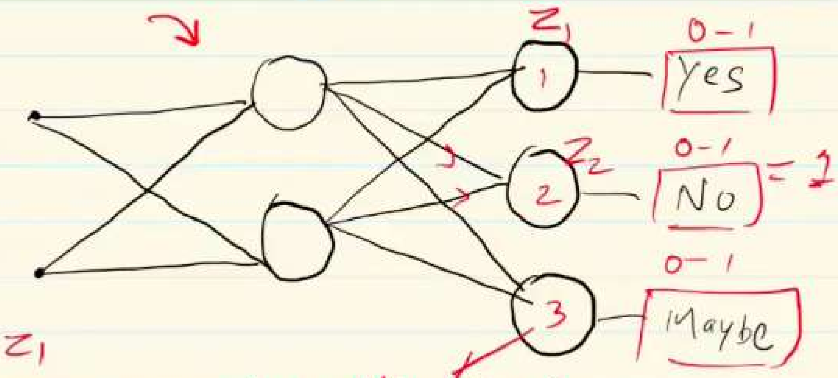
$$L = -y_1 \log(\hat{y}_1) - y_2 \log(\hat{y}_2) - y_3 \log(\hat{y}_3)$$

$$\frac{e^{z_3}}{e^{z_1} + e^{z_2} + e^{z_3}}$$

$$\frac{e^{z_2}}{e^{z_1} + e^{z_2} + e^{z_3}}$$

$$f(z) = \frac{e^{z_1}}{e^{z_1} + e^{z_2} + e^{z_3}}$$

softmax



Activation

neurons

class ← categories