Loss functions classification

28 September 2023 07:06

① Binary cross Entropy.

9 yes 1 NO   1 yes 9 NO.

uncertainity.

10 y 0 N   0 y 10 NO

Certain.

5 y 5 N   5 y 5 N

uncertainity ①

$\boxed{\text{Entropy} = 1.}$

$$\boxed{Loss = -y_{act}\log(\hat{y}) - (1-y_{act})\log(1-\hat{y})}$$

y_{act} = actual
$\hat{y}$ = predicted.

$$-\frac{1}{n}\left[\sum_{i=1}^{n} \underbrace{y\log(\hat{y})}_{\text{class 1}} + \underbrace{(1-y_{act})\log(1-\hat{y})}_{\text{class 0}}\right]$$

$y_{act} = 0$

$= 0 \times \log(\hat{y}) + (1-0)\log(1-\hat{y})$

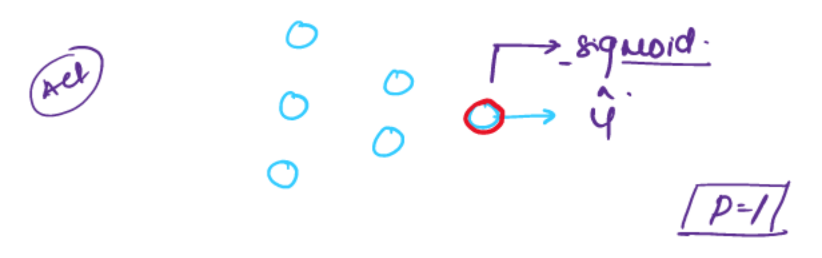
$= 0 + (1)\log(1-\hat{y})$

$= \log(1-\hat{y})$

$y_{Act} = 1$

$= 1 \times \log(\hat{y}) + 0 \times \log(1-\hat{y})$

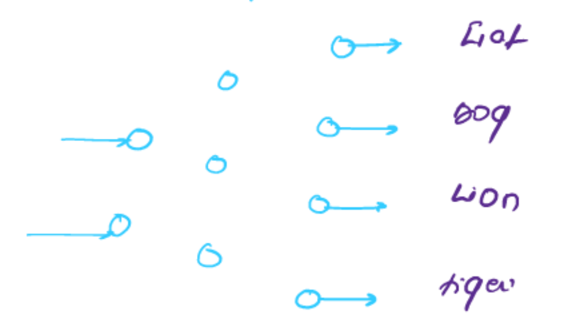
$= \log(\hat{y})$



$\boxed{P=1}$

Categorical cross Entropy.

used in multiclass classification problems. Approved, Receive, Lapipeline



$$Loss = -\sum_{i=1}^{K} y_{act} \times \log(\hat{y})$$

$$\underset{\text{actual}}{-y_1}\log\overset{\text{predicted}}{\hat{y}_1} - y_2\log\hat{y}_2 - y_3\log\hat{y}_3 - y_4\log\hat{y}_4$$

Age	Sal	class.	OHE y_1 y_2 y_3	$\hat{y}_1$ $\hat{y}_2$ $\hat{y}_3$
7	100	A	1 0 0	[0.3, 0.2, 0.5]
8	110	D	0 1 0	[x 0 x]
9	700.	tp.	0 0 1	

$-1\log(0.3) - 0 \times \log(0.2) - 0 \times \log(0.5).$

$\boxed{-\log(0.3).}$

all the probabilities summation of

Basically we are softmax into the output layer.

target variable must be converted into one hot Encoding.

it produces an array of one hot encoding for probable matches of each category.

Sparse categorical cross Entropy.

it is frustrating when you are dealing with or working with large classification problems.

And if suppose you are doing one hot encoding it Requires significant amount of memory.

Age	Sal	class.	CCE OHE	SPCCE
7	100	A	1 0 0	1 [x x x]
8	110	D	0 1 0	2 [x x x]
9	700.	tp.	0 0 1	3 [x x x]

Sparse categorical cross Entropy performs same calculations without Requiring one hot encoding of target column.

$\boxed{HP}$

vol

199996 word.

2 word

1 1 1 1

$\boxed{R \times C}$

Lion Long

1 0 1 1 0 0 0 0

CCE   1-1000

Shirt   999

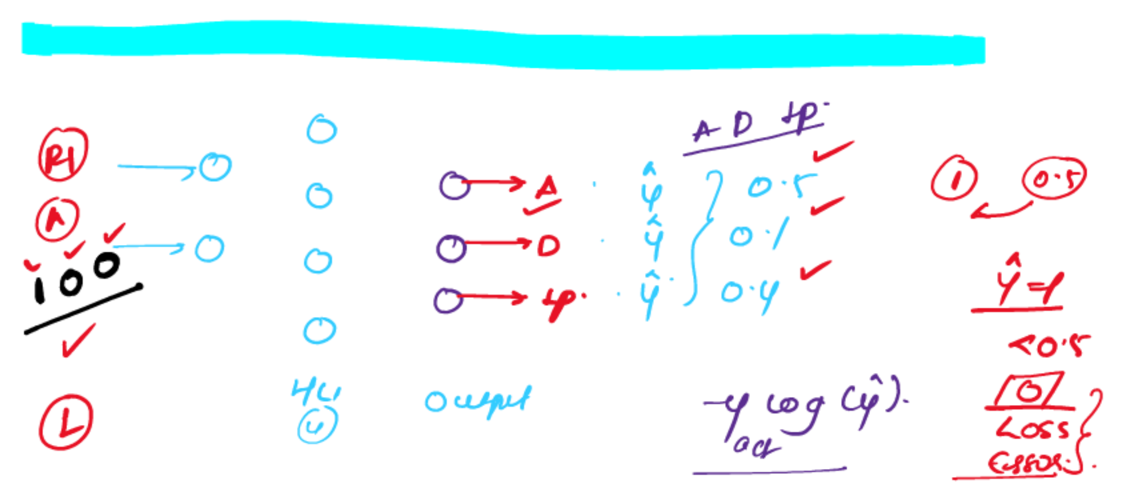
0.

10

1x10   999 x10

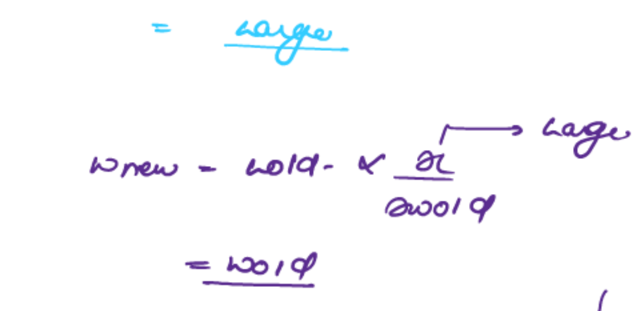
$\boxed{10}$ 9990

✓1   open   0.



$-y_{act}\log(\hat{y}).$

Exploding gradient.



$$w_{new} = w_{old} - \alpha \frac{\partial L}{\partial w_{old}}.$$

$$\frac{\partial L}{\partial w_{old}} = \frac{\partial L}{\partial O_{31}} \times \frac{\partial O_{31}}{\partial O_{21}} \times \frac{\partial O_{21}}{\partial O_{11}} \times \frac{\partial O_{11}}{\partial w_{11}}.$$

$= 1 \times 10 \times 0.2 \times 0.5.$

$=$ large

$$w_{new} = w_{old} - \alpha \underbrace{\frac{\partial L}{\partial w_{old}}}_{\text{large}}$$

$= w_{old}$

wnew   wold   significant

$\boxed{\alpha \uparrow}$

Solution.

① proper weight Initialization.

gradient get larger and large in case of Back propagation.

Model experiences high learning. Exponential growth in model parameters.