Log loss is hard to interpret.
If I say log-loss = 1.0
we cannot interpret it easily

30.6 R^2 or Coefficient of determination

Lets look at some methods to determine how good a regression technique is.

We know that $\underbrace{x_i, y_i}_{\text{dataset}}, \underset{\substack{\uparrow \\ \text{model output}}}{\hat{y}_i}$

e_i = error for point i = $y_i - \hat{y}_i$
= difference between actual & model output

Sum of squares (SS)

$$SS_{total} = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \text{Total sum of squares}$$

$\bar{y}$ = Average value of all y_i
$= \frac{1}{n} \sum_{i=1}^{n} y_i$

When we are performing regression, what is simplest model we can build → return mean (y_i)

Imagine a case where we have to predict a height based on features like weight, skin colour, hair colour, etc.

Given any new person x_q, in my whole training data, if I know that avg height of a human being is 152 cm.

So, I will predict the avg height of $x_q \rightarrow 152$ cm

Thats the simplest model we can build.

$\therefore$ Simple_Mean_Model

$$x_q \rightarrow \underline{\text{Mean}(y_i)} \text{ as } y_q.$$

$$\Downarrow$$

$$\bar{y}.$$

$$\therefore SS_{Total} = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

Sum of squared error using a simple Mean Model.

$$SS_{Residues} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e_i^2$$

$$Residue = e_i = \underbrace{y_i}_{\text{Actual}} - \underbrace{\hat{y}_i}_{\text{Predicted}}$$

$$R^2 \equiv \left(1 - \frac{SS_{Residue}}{SS_{Total}}\right)$$

Case 1:

$SS_{res} = 0$ (i.e. $e_i = 0$) $\Rightarrow R^2 = 1$
(Best value)

It means all of residues or errors are 0.

Case 2:

$SS_{res} < SS_{Total}$; $R^2 = 0$ to 1

Case 3:

$SS_{res} = SS_{Total}$, $R^2 = 0$

It means the model that generated residues is same as simple mean model.

~~It mean~~

Case 4:

$SS_{Residue} > SS_{Total}$; $R^2 = -ve$.

Model is worse than a simple Mean Model.

So, when someone says.

$R^2 = 0.9 \Rightarrow$ the model is near to perfection

$R^2 = 0.1 \Rightarrow$ the model is near to simple Mean model

$R^2 < 0 \Rightarrow$ the model is worse than mean model.

30.7 Median Absolute Deviation of errors

Lets say if one e_i is very large then $SS_{Residue}$ would get corrupted.

$\therefore$ R^2 is not very Robust to outliers

For every datapoint x_i, I have y_i & $\hat{y}_i$ and a corresponding e_i

In Regression, the measure of how good a model is ~~abt~~ all about e_i

$\therefore$ If $|e_i's| \rightarrow 0 \Rightarrow$ Great

$|e_i's| \rightarrow$ Large $\Rightarrow$ Not so Good

If e_i = Random Variable
then
Median (e_i) = Central Value of errors
M.A.D (e_i) = Median ($|e_i - \text{Median}(e_i)|$).

If I know that :
Median = small } then Error is
and M.A.D = small. } small.

For calculating the efficiency of module, we can use.

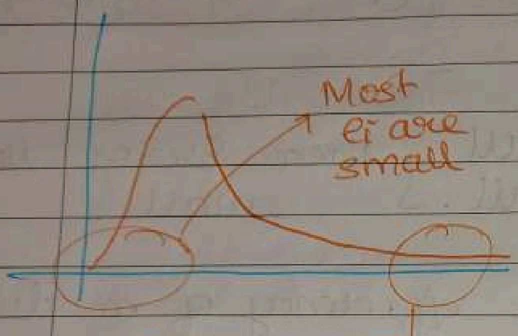
Mean OR Median of eis.
Std-Dev OR M.A.D

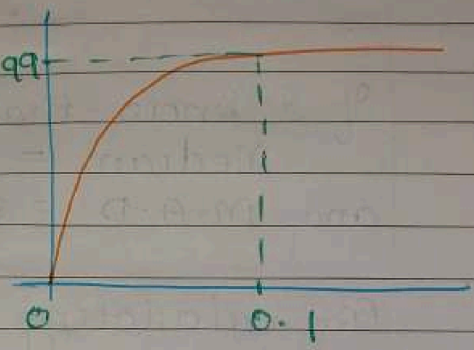
↓

Robust to Outliers

30.8 Distribution of Errors

If we plot the PDF and CDF of errors e_i





Very Few e_i are Large

↓

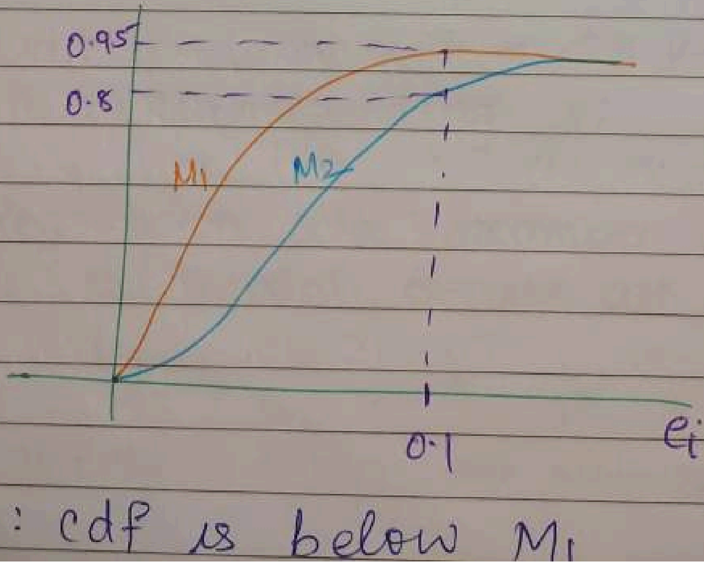
Good sign.
We should have e_i as close to 0

99% of errors are < 0.1
only 1% of error are > 0.1

↓

Good sign

Lets say I have CDF of two Model



M_2 : cdf is below M_1