Logistic Regression 2

- Code
- Accuracy
- Log - odds
- one vs Rest
- Impact of outlier

Accuracy Metric

y	$\hat{y}$	y_{pred}	
1	0.8	1	✓
0	0.1	0	✓
1	0.4	0	X

$\frac{2}{3} = 66.6\%$

100 test records

⇓

95 are correct

$\frac{95}{100} \simeq 95\%$

Q ⇒ 1500 samples → 750, 750

750 samples for class 1 & class 2

Pred :

450 – class 1
350 – class 2

} ✓ ←

$$\frac{450 + 350}{1500} = 0.53$$

⇒ Log-odds

odds : 1 : 4

$p(\text{winning}) = \frac{1}{5}$

$p = P(y=1 \mid x)$

$1-p = P(y=0 \mid x)$

odds $= \frac{p}{1-p}$

$$p = \frac{1}{1 + \exp(-(w^T x + w_0))}$$

$$p = \frac{\exp(-(w^T x + w_0))}{\exp(-(w^T x + w_0)) + 1} = \frac{e^{-z}}{1+e^{-z}}$$

$$(1-p) = 1 - \frac{\exp(-(w^T x + w_0))}{\exp(-(w^T x + w_0)) + 1}$$

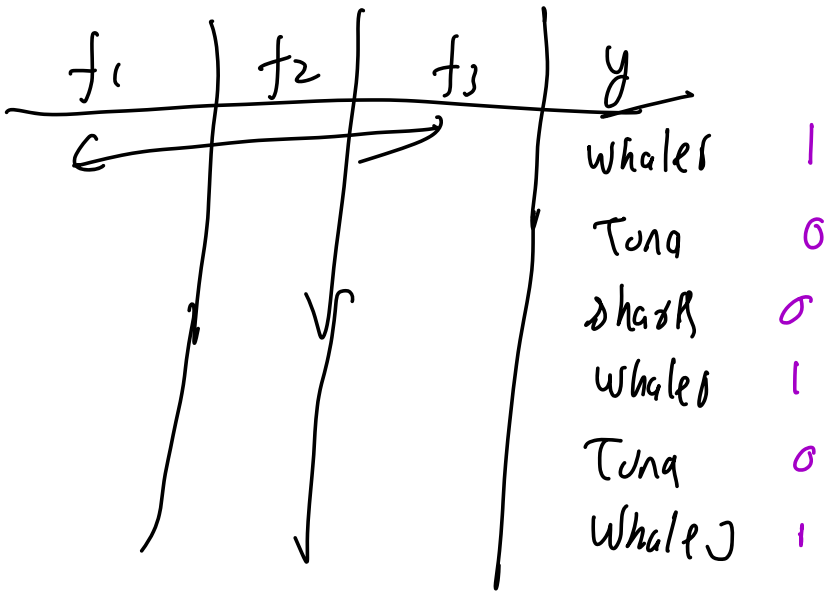
$$= \frac{1 + e^{-z} - e^{-z}}{1 + e^{-z}}$$

$$= \frac{1}{1+e^{-z}}$$

$$\left(\frac{p}{1-p}\right) = \frac{\frac{e^{-z}}{1+e^{-z}}}{\frac{1}{1+e^{-z}}} = e^{-z}$$

$$\log\left(\frac{p}{1-p}\right) = -z$$

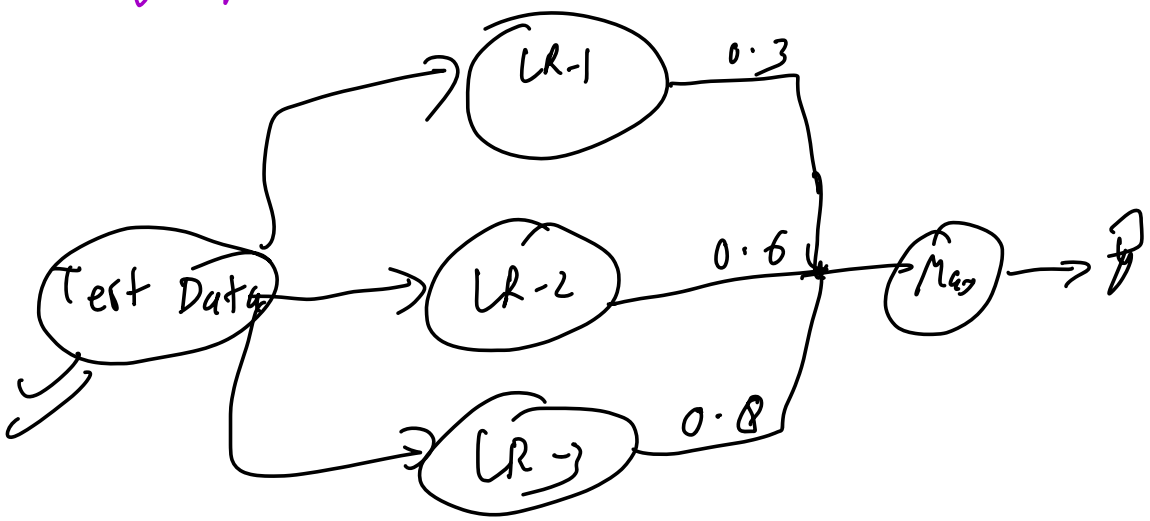
One vs Rest ⇒ > 2 o/p



Log Reg 1 ⇒ Whales vs [Tuna, Shark]

Log Reg 2 → Tuna vs [W, S]

Log Reg 3 → Shark vs [W, T]



Impact of outliers