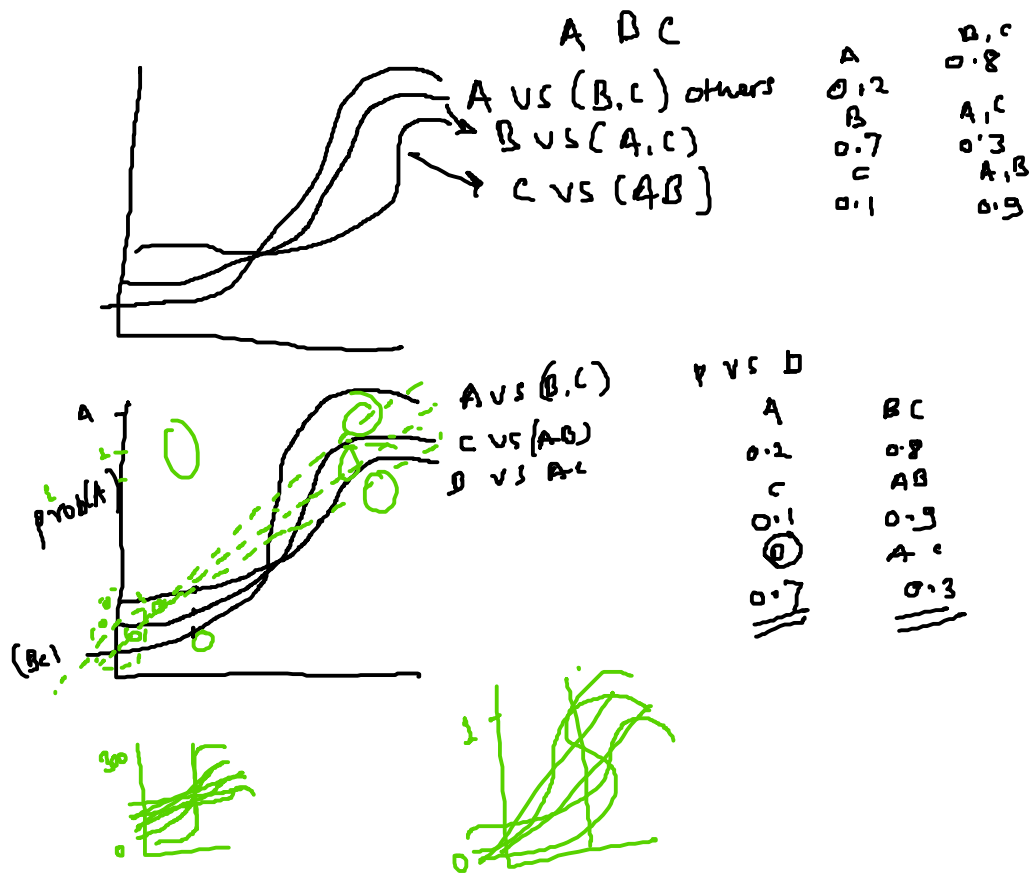
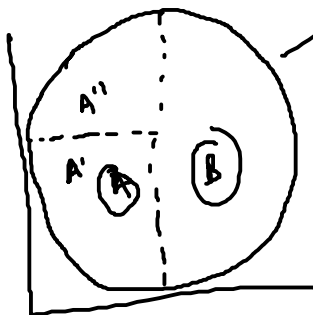


Multiclass (one vs rest) OVR

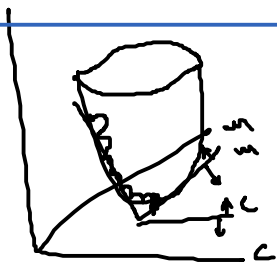
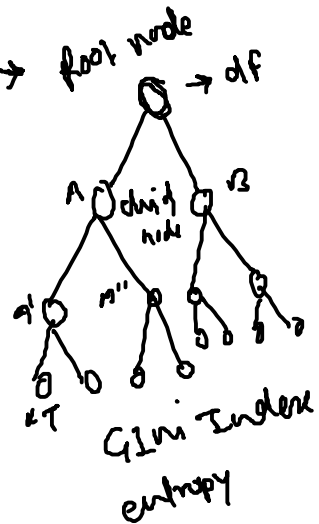


Decision Tree



test
S

test



Logistic uses gradient descent

Log Loss function (cost function)

L $y = \max c$ $\text{limit} = 1$

predict

$$\text{Log loss} = \frac{1}{N} \sum_{i=1}^n y_i (\log(y_p)) + (1 - y_i) \log(1 - y_p)$$

y_q = actual value (0, 1)

y_p = probability

	T	T
default	spam	1
not default	not-spam	0
	rusty	0
	strong	1
	not	0

For multiclass

low, medium, high
0, 1, 2

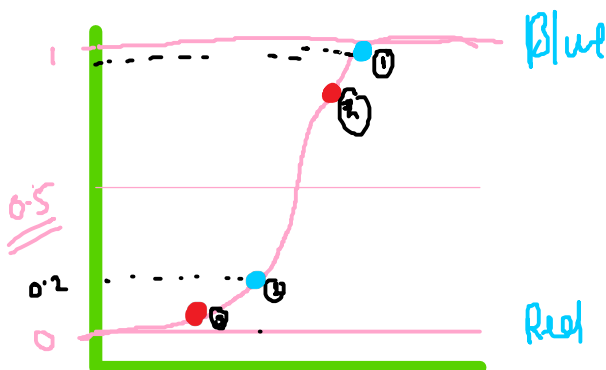
low vs (m, h)
0 vs (1, 2)

med vs (low, h)

0 vs 1

h vs low, m

0 vs 1



Case 1: $y_q = 1$, $y_p = \text{high} = 0.9 \gg \text{correctly classify}$

$$y_q \times \log(y_p) + (1 - y_q) \log(1 - y_p)$$

$$L = 1 \times (\log(0.9)) + (1-1) \times \log(1-0.9)$$

$$= -0.1 + 0$$

$$L = -0.1$$

>> correctly classified

When $y_a = 1 \gg \text{part 2} = 0$
 $y_a = 0 \gg \text{part 1} = 0$

case (2) $y_a = 1, y_p = \text{low} = 0.2$

model is saying because of prob of (2)
 belonging to class Blue is less/low.

model predict it belongs to Red class

Misclassified / incorrectly classified

$$L = 1 \times \log(0.2) + (1-1) \times \log(1-0.2)$$

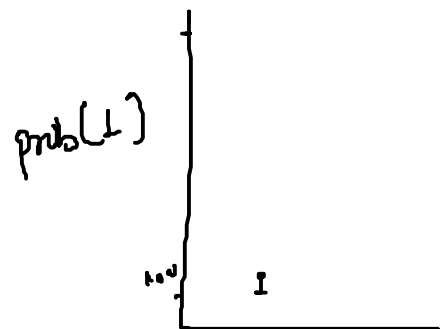
$$= 1 \times 1.6 + 0$$

$$L = 1.6$$

case (3) $y_a = 0, y_p = \text{low}$

model is smaller prob of
 (3) being blue

model is predicting it
 belongs to Red



Correctly classified

$$L = 0 \times (\log(0.1)) + (1-0) \log(1-0.1)$$

$$= 0 + 1 \times \log(0.9)$$

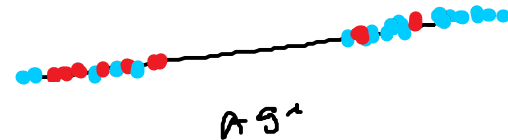
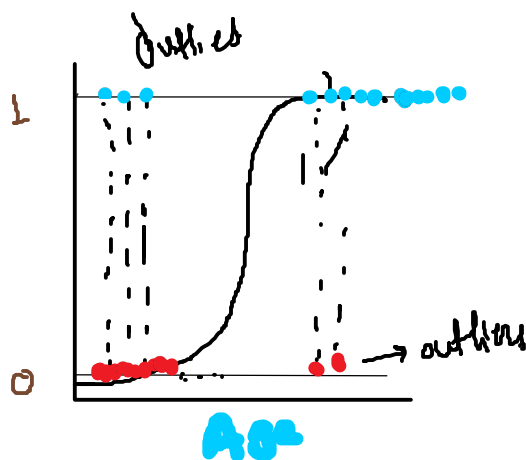
$\bar{L} = \bar{0.1}$; near to zero

Case (ii) $y_a = 0$, $y_p = \text{high} = 0.75$

model is saying the probability of (ii) belonging to class Blue is high.
model predict it belongs to class Blue.

Misclassified / incorrectly classified

$$\begin{aligned} L &= 0 \times (\log 0.75) + (1-0) \log (1-0.75) \\ &= 0 + 1 \times \log (0.25) \\ L &= -1.38 \end{aligned}$$



Metrics for classification

- 1) Confusion Matrix
- 2) Classification report
- 3) Recall, precision, F1-score, accuracy_score.
- 4) AUC, ROC.

hi