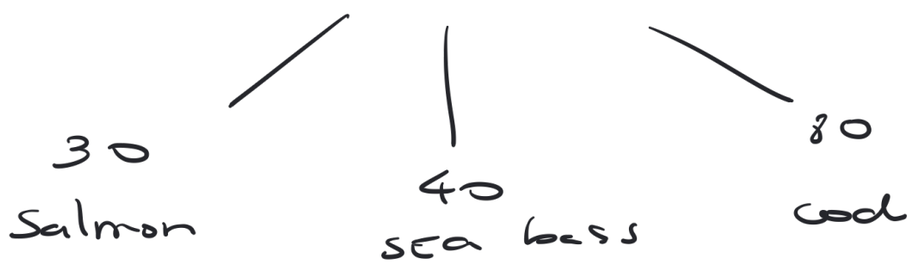
FISH DATASET: 150 samples

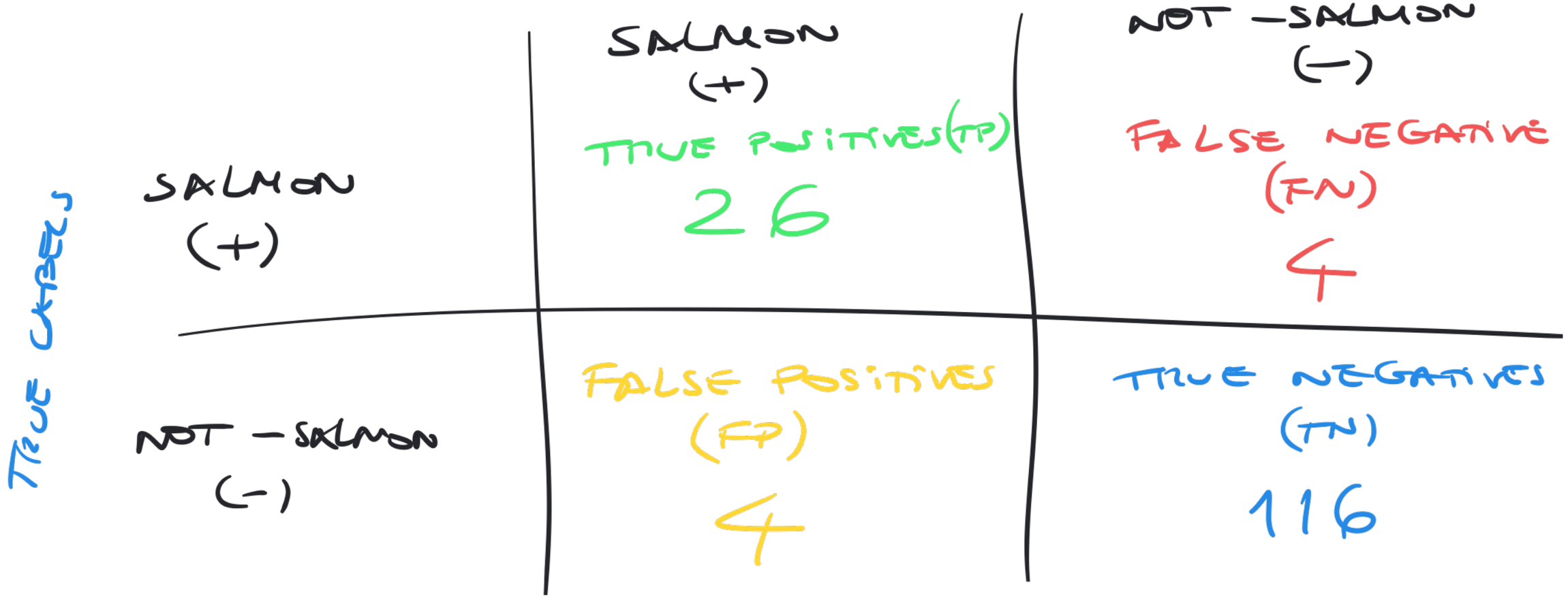


Confusion Matrix: $C \times C$ matrix

$C \equiv$ # classes

TRUE LABELS	PREDICTED LABELS: SALMON	SEA BASS	COD	TOTAL
SALMON	26	2	2	30
SEA BASS	2	38	0	40
COD	2	3	75	80
				150

For each class, we can compute a binary confusion matrix (2x2)



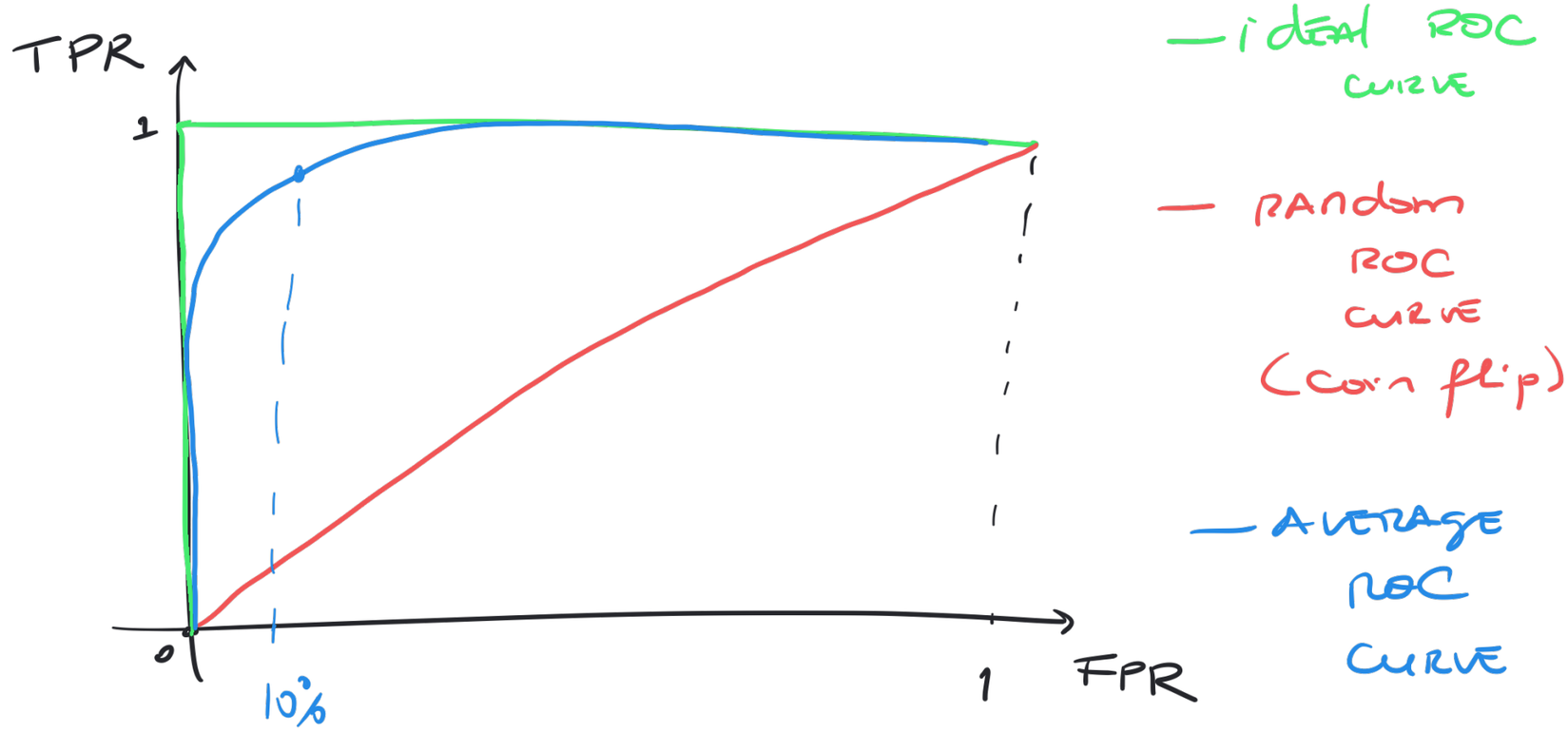
① FPR ≡ FALSE POSITIVE RATE

$$FPR = \frac{FP}{FP + TN}$$

② TPR ≡ TRUE POSITIVE RATE

$$TPR = \frac{TP}{TP + FN}$$

ROC CURVE (Binary measure)



— ideal ROC curve

— random ROC curve (coin flip)

— AVERAGE ROC CURVE

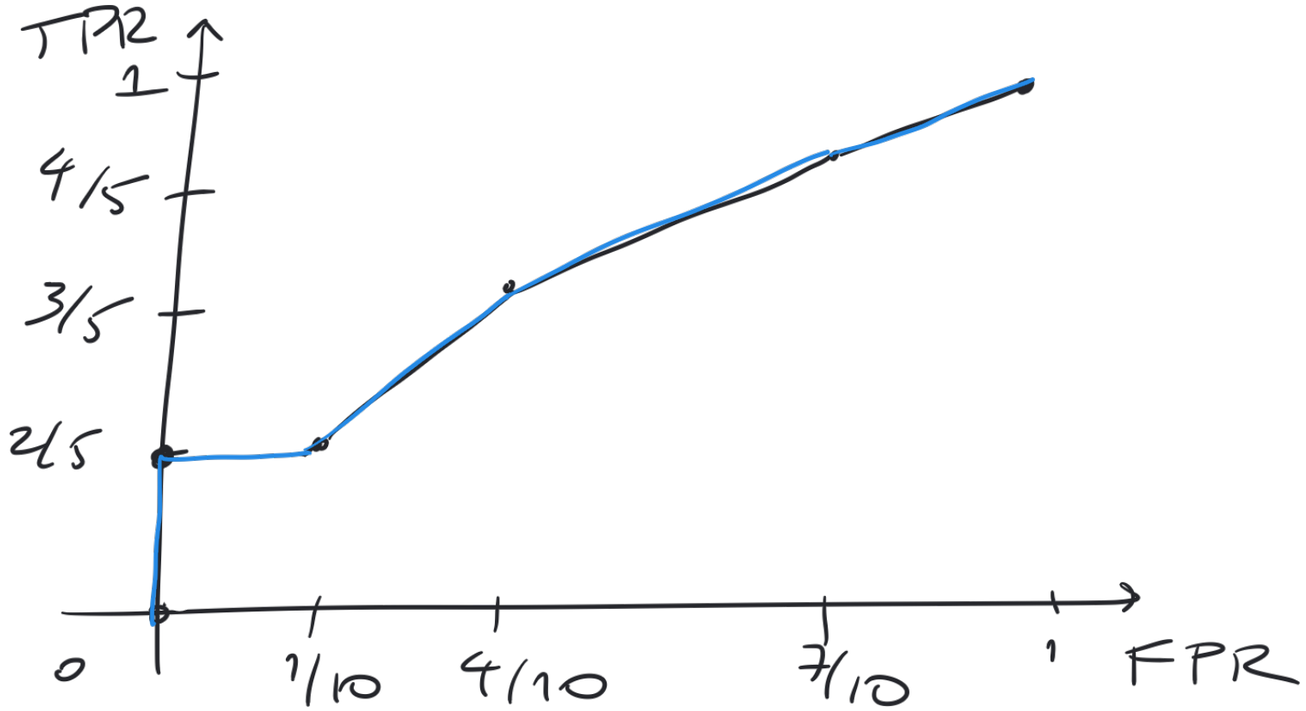
AREA UNDER THE CURVE (AUC)

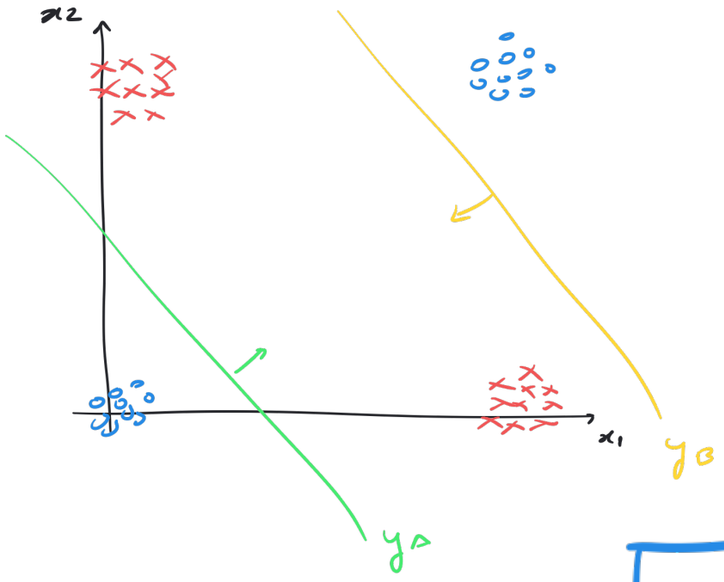
- FOR random: $AUC = \frac{1}{2}$
- FOR ideal: $AUC = 1$
- FOR typical: $\frac{1}{2} \leq AUC \leq 1$

$$TPR = \frac{TP}{TP+FN}, \quad FPR = \frac{FP}{FP+TN}$$

Confidence	Ground Truth	Prediction
0.91	T	T
0.9	T	T
0.8		
0.79	T	
0.77		
0.75		
0.5		
0.4	T	
0.39		
0.38		
0.37		
0.25		
0.10		
0.09	T	
0.01		

THRESHOLD	TPR	FPR
≥ 0.5	3/5	4/10
0.8	2/5	1/10
0.3	4/5	7/10
0.9	2/5	0/10 = 0
0	5/5 = 1	10/10 = 1

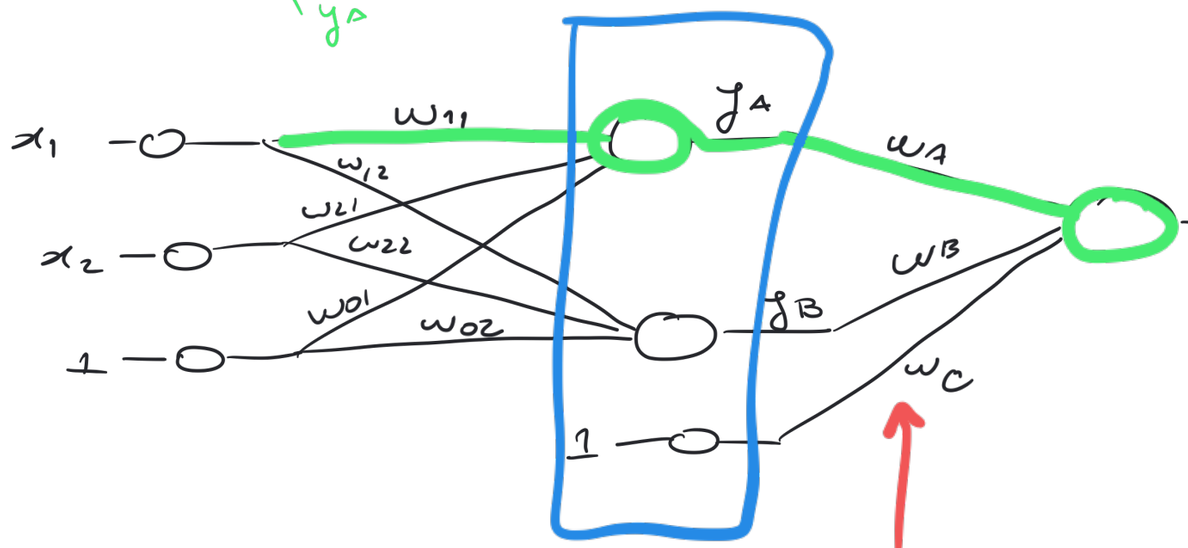


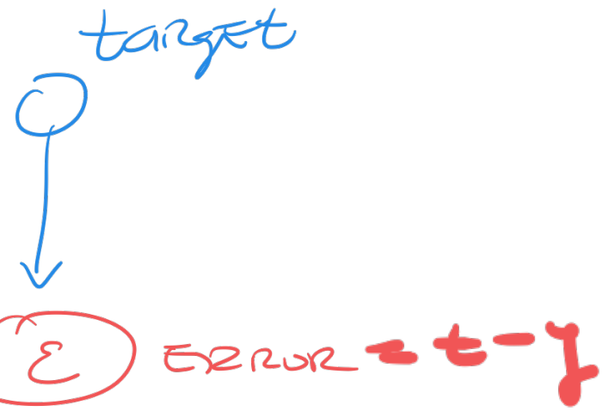


o – class 1

x – class 2

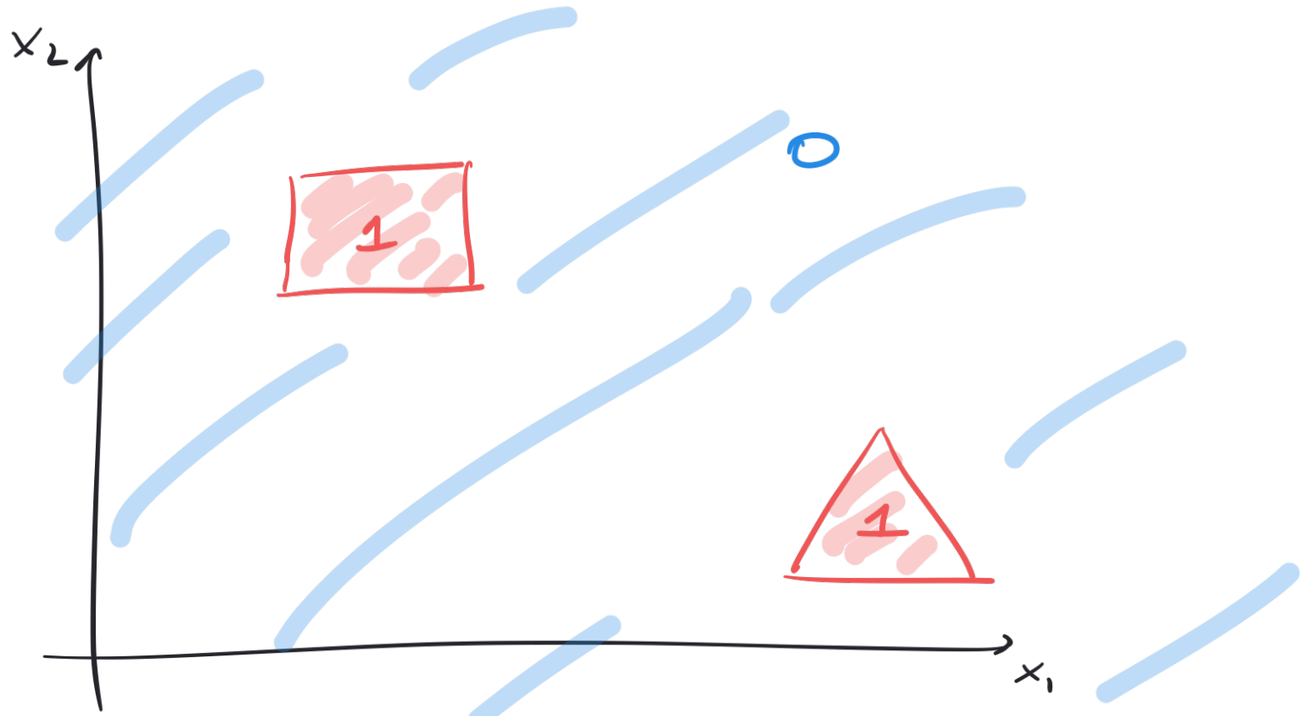
Hidden layer





Learning algorithm

OBJECTIVE function



DRAW the MLP for this classification task.

① How many units in the input layer?

② How many units in the hidden layer?

③ How many units in the output layer?