MACHINE INTELLIGENCE

Performance metrics

K.S.Srinivas
Department of Computer Science and Engineering

MACHINE INTELLIGENCE

Performance metrics

Srinivas K S.
Associate Professor, Department of Computer Science

1. Accuracy
2. Precision
3. Recall
4. Specificity
5. Receiver Operating Characteristics (ROC)
6. Area Under Curve (AUC)

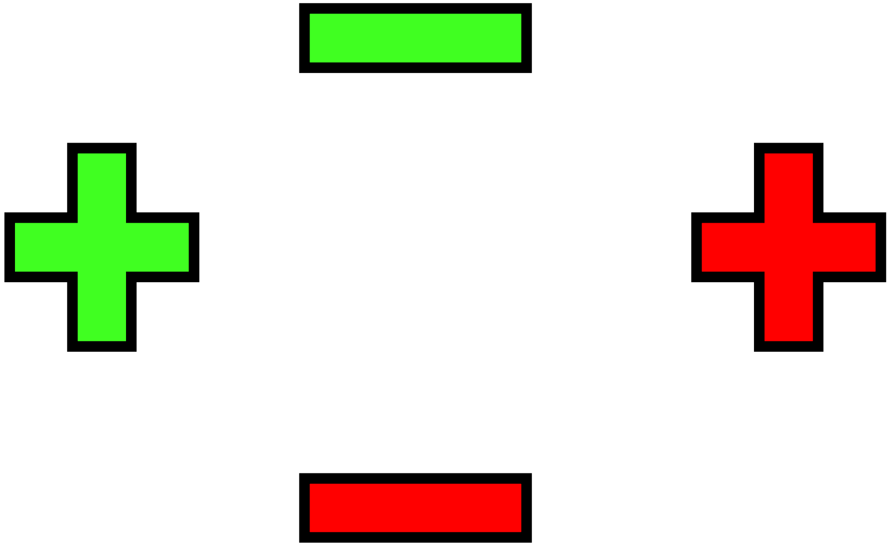
After the initial feature Engineering and implementing a model, we get some output in terms of a class or probability.

MACHINE INTELLIGENCE

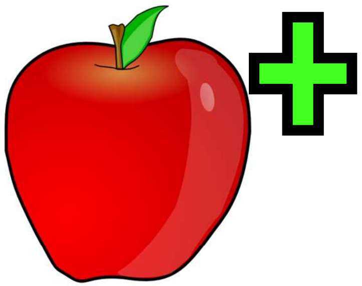
Confusion Matrix

1. True Positive
2. True Negatives
3. False Positive
4. False Negatives

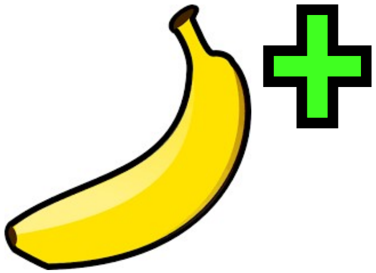
Before moving on further, we will see what is a confusion matrix.
It is a simple way to lay out ,how many predicted categories or classes were correctly predicted and how many were not.



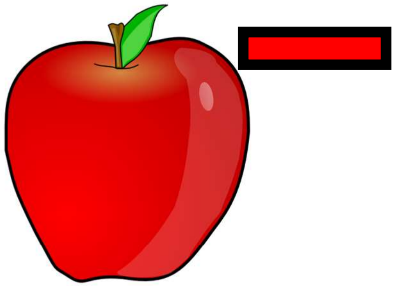
Confusion Matrix



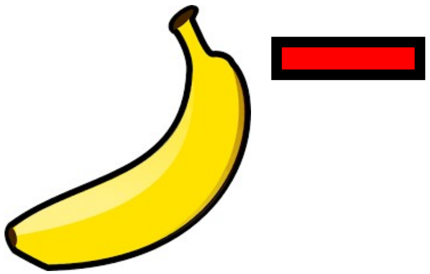
class A correctly predicted as class A



class B correctly predicted as class B



class A incorrectly predicted as class B

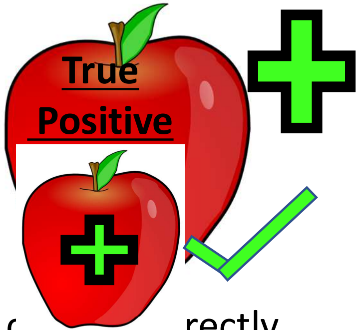


class B incorrectly predicted as class A

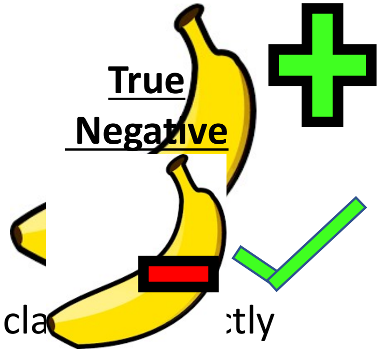
suppose our classification model has two class class A(**apple**) and class B (**all other fruits**)

essentially the confusion matrix is keeping track of

Confusion Matrix



class A correctly
predicted as class A
a positive class
A predicted as class A



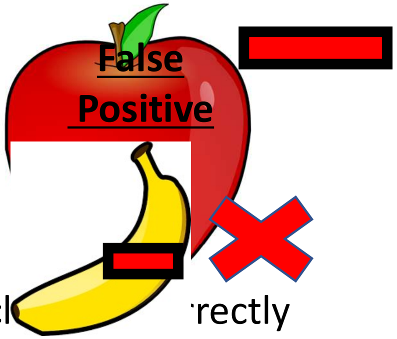
class B correctly
predicted as class B
a negative class
B predicted as class B

suppose our classification model has two classes
then our true positive and true negative are as follows
class A (**apple**) and class B (**all other fruits**)

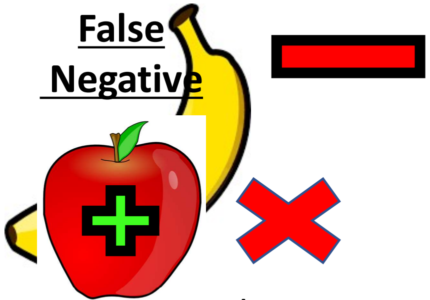
essentially the confusion matrix is keeping track of
our false positive and false negative are as follows

GOAL

As many predictions as possible
More true then false



class A incorrectly
predicted as class B
a negative class
B predicted as positive class A



class B incorrectly
predicted as class A
a positive class
A predicted as negative class B

Confusion Matrix

N: 200	**Predicted:** CLASS A	**Predicted:** CLASS B
Actual: CLASS A	**TP** 60	**FP** 40
Actual: CLASS B	**FN** 70	**TN** 30

NOTE:
FP and FN are type 1 and type 2 error respectively (same as what you learn't in IDS course in 3rd sem)

[illegible]
[illegible]
+ve and -ve class
we draw a matrix grid
with predictions from our model as below
along x axis we represent predicted values and
along y axis we represent actual values
and total wrong predictions are 40+70=110

N	class A	class B
200	100	100

- 60 of the data object were correctly predicted positive class A as in the data set
- 30 of the data object were correctly predicted as negative class B as in the data set

- most of the time it doesn't matter as the confusion matrix will tell us how many predictions were properly made from each class
- but if your classification is binary or of type one vs all you can assign the target class as +ve class

Accuracy

Accuracy is given by,

$$\textbf{Accuracy=} \frac{TP + TN}{(TP + TN + FP + FN)}$$

Accuracy in classification problems is the number of correct predictions made by the model over all kinds predictions made.

Accuracy is generally a good measure when the target variable classes are nearly balanced.

Precision and Recall

Precision: How many +ve cases did we catch?

Recall: How many did we miss? (sensitivity)

Precision is given by,

Precision= $$\frac{TP}{(TP+FP)}$$

Recall is given by,

Recall= $$\frac{TP}{(TP+FN)}$$

Specificity and F1 score

Specificity : How many -ve cases did we catch?

Recall: The harmonic mean of precision and recall.

Specificity is given by,

Specificity= $\dfrac{TN}{(TN + FP)}$

F1 score is given by,

F1 score= $\dfrac{2 * (recall * precision)}{(recall + precision)}$

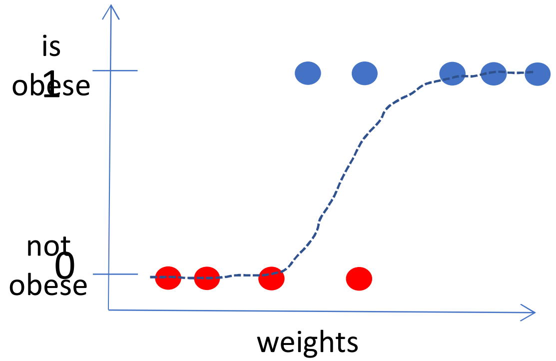
Metrics calculations

N: 200	Predicted: CLASS A	Predicted: CLASS B	CLASSIFICA-TION OVERALL	
Actual: CLASS A	TP 60	FP 40	100	precision
Actual: CLASS B	FN 70	TN 30	100	
TRUTH OVERALL	130	70	200	
	recall	specificity		accuracy

consider the following confusion matrix
Let us calculate all the metrics we discussed till now

$$\frac{\quad + \quad}{\qquad} = 0.45$$

ROC -Receiver Operating Characteristics



we fit a logistic regression curve to the data where y axis is converted to probability that the sample is obese

if we need to convert probability to classification ,one method is setting up a threshold say at 0.5 and say all sample with probability more than 0.5 is obese and vice versa

setting threshold to 0.5 we prepare a confusion matrix similarly we create for other threshold

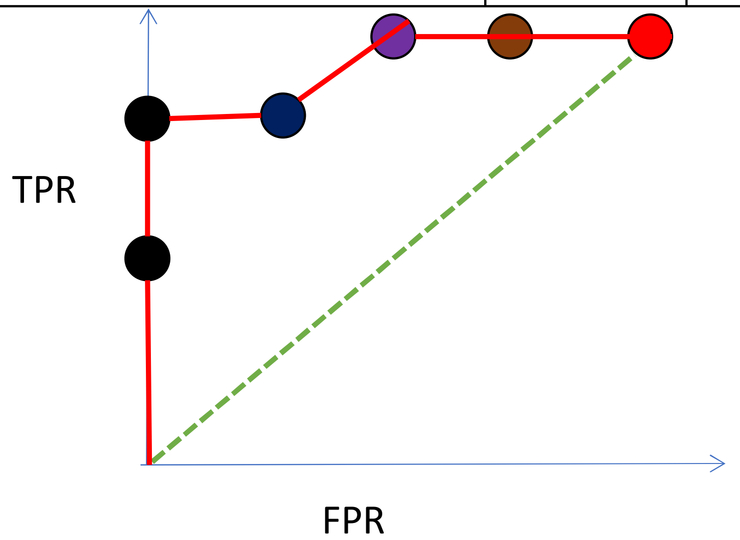
thus after we get the number of confusion matrix for n chosen threshold

now the task is to choose the with best threshold

- the y axis has two categories- obese and not obese
- the blue dots represent sample who are obese
- the red dots represent sample that are not obese
- along x axis we have weights

ROC -Receiver Operating Characteristics

Threshold	TPR	FPR
0 (all sample classified obese)	1	1
0.3	1	0.75
0.4	1	0.5
0.6	0.75	0.25
0.7	0.75	0
0.9	0.5	0



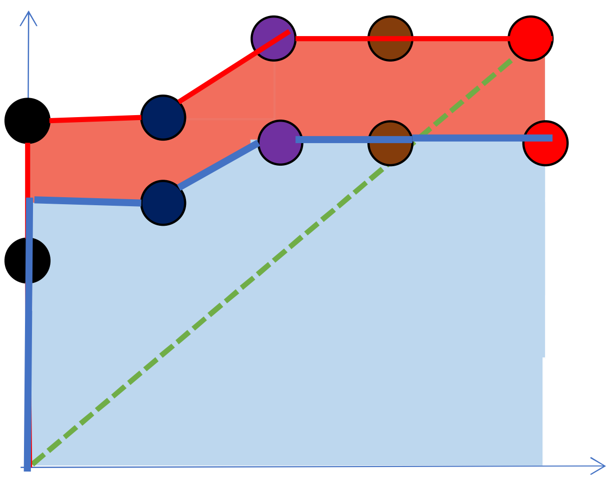
proportion of samples that were incorrectly classified as obese

this means the new threshold is better then the first one

- y axis represents true positive rate (TPR) that is sensitivity
- x axis represents False positive rate (FPR) that is specificity

AUC -Area Under The Curve

between classes.



THANK YOU

K.S.Srinivas
srinivasks@pes.edu
+91 80 2672 1983 Extn 701