{Classification}

When we have Balanced Dataset :→ Go for Accuracy

else :→ Precision, Recall or F1 score !!!

For Binary Classification

(1). Confusion Matrix

FP rate :- $\left\{ \frac{FP}{FP + TN} \right\}$ → $\frac{FP}{(\text{total Neg.})}$

	1	0
1	TP	FP (Type I error)
0	FN (Type II error)	TN

FN rate :→ $\left(\frac{FN}{FN + TP} \right)$

Now since we have a Balanced dataset,

Accuracy :- $\left\{ \frac{TP + TN}{TP + TN + FP + FN} \right\}$

Consider that we have an unbalanced dataset as following

$\left\{ \begin{array}{l} A : 90\% \\ B : 10\% \end{array} \right\}$ + Now let's assume the model classify everything into class A, then by default it'll have a 90% accuracy.

(2.) Precision & Recall — (Sensitivity) (Specificity)

Precision :→ $\frac{TP}{TP + FP}$

Recall :- $\frac{TP}{TP + FN}$

	1	0	
1	TP	FP	Recall
0	FN	TN	
	Precision		

The choice of Precision or Recall depends on your problem statement.

Eg :→ Spam Detection :- Need Precision ↑↑
FP is very Bad for this case

Corona Detection :→ Need Recall ↑↑
False Negative could be lethal for this.

$$F_{Beta} = (1 + \beta^2)\frac{(P \times R)}{\beta^2 \times (P + R)}$$

$\beta = 1$; $F_1 = \left\{ \frac{2PR}{P + R} \right\}$ Harmonic Mean ⇒ $\left\{ \frac{1}{\frac{\frac{1}{P} + \frac{1}{R}}{2}} \right\}$

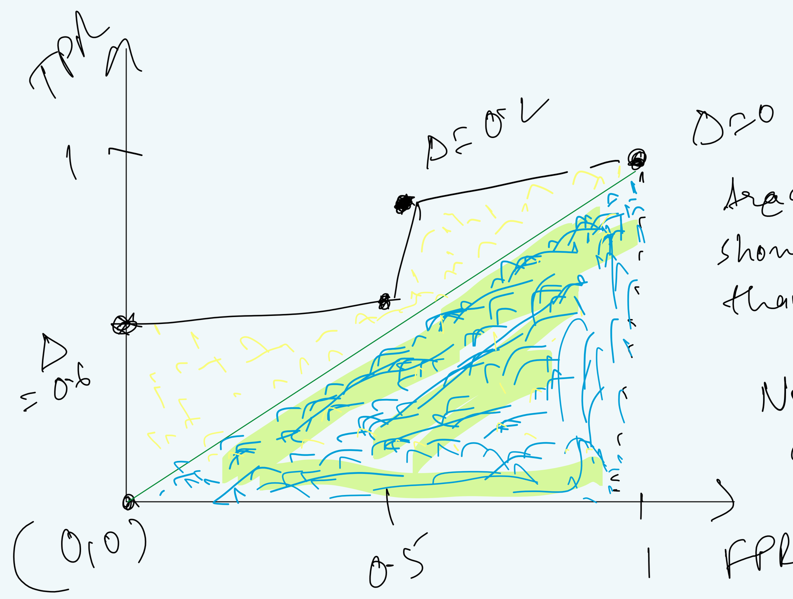
(When Both false positive & false negatives are important)

if false Negative impact is high, "increase β" β↑ (2, 3) | if false positive impact is high "decrease β" β↓ 0.5

③ AUC & ROC

y	$\hat{y}_{prob}$	$\Delta(\hat{y})$ $\Delta = 0$	$\Delta(\hat{y})$ $\Delta = 0.2$	$\Delta(\hat{y})$ $\Delta = 0.3$
1	.8	1	1	
0	.96	1	1	
1	0.4	1	1	
1	0.3	1	1	
0	0.2	1	0	
1	0.1	1	1	

$$\left\{ \begin{array}{l} TPR = \frac{TP}{TP + FN} \\ FPR = \frac{FP}{FP + TN} \end{array} \right\}$$



Area under the curve should always be greater than the base line model

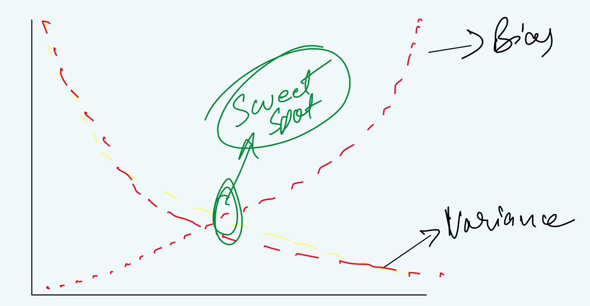
Now depending on the domain knowledge, we can tune our tradeoff b/w TPR & FPR

The Bias-Variance Tradeoff

(Bias) ↓ (Variance) ↑

(Bias) ↑ (Variance) ↓

But we need Bias ↓ , Variance ↓



Bias :- Error from mistake assumptions of the Algorithm

High Bias can make the model miss the relevant relationship b/w the feature and the target resulting in Underfitting.

Variance :→ Variance is an error that occurs from the sensitivity of the model.
High sensitivity give rise to a condition where model learns everything called overfitting.

Train Set error :- 1% } → High Variance
Test set error :- 11%

Train Set Error's → 15% } High Bias
Test set error → 16%

Train set error 18% } Apocalypse
Test set error 30%

High Bias —yes→ Bigger Model, Train longer.

↓ N

High Variance —yes→ More data or Regularization

↓ N

you've got a Great Model.

Regularization :-

L_1 ⇒ $\frac{\lambda}{2m}||W||$ ↳ $||W|| = |w_1| + |w_2| + |w_3| + \dots$ — Laplace

L_2 ⇒ $\frac{\lambda}{2m}||W||^2$ ↳ $||W||^2 = (|w_1|^2 + |w_2|^2 + |w_3|^2 \dots)^{1/2}$ — Gaussian

While you do gradient descent,

$$Loss = y \log \hat{y} + (1 - y) \log(1 - \hat{y})$$

$$(Loss)_{L_1} := Loss + \frac{\lambda}{2m} * ||W||$$ Norm

$$(Loss)_{L_2} := loss + \frac{\lambda}{2m}||W||^2$$