(1) Performance Metrics

(2) Naive Bayes — Performance Metrics

Loss — Cost

Cost $\rightarrow \Sigma(\$)$

Loss $\rightarrow$ sample wise

Metric $\rightarrow X$

$\Downarrow$

Understand the quality

Regularisation $\rightarrow L1, L2,$ Elastic

$L + \lambda_1(\ \ ) + \lambda_2(\ \ )$

-ve, +ve

		Actual Values	
		+ve	-ve
Predicted Values	+ve	TP	FP
	-ve	FN	TN

Man, Woman

FP $\rightarrow$ Type I Error

$\Rightarrow$ Pred. falsely pred.

$\Rightarrow$ Actual as -ve, model as +ve

FN $\rightarrow$ wrongly predicted — Type II Error

(1) Actual was +ve, Model predicted as -ve.

(1) Accuracy $= \dfrac{TP + TN}{TP + TN + FP + FN}$

Positive — 10,000

Negative

TP ...   100

① Precision

how many of the correct classes actually are +ve

$$\text{Precision} = \frac{TP}{TP + FP}$$

Precision is Import in some

Eg Recommendation,

FP - can cause churns

② Recall

How much actual positive cases were we able to predict correctly

$$\text{Recall} = \frac{TP}{TP + FN}$$

Recall is import in some cases,

FN → very Import.

$$\text{f1-Score} \Rightarrow \frac{2}{\frac{1}{\text{Recall}} + \frac{1}{\text{Precision}}}$$

$$\frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}}$$

Accuracy

↓

easy to interpret.

f1-Score is an measure perform. as harmonic mean

~~Reca~~ Precision of majority & Recall of the majority class.

f1 - score minority

	+ve	-ve
+ve	TP	FP
-ve	FN	TN

	-ve	+ve
-ve	TN	FN
+ve	FP	TP

100

f1 - +ve class

f1 - -ve class

f1

Cat → 10,000 → 4000

dog → 100 → 6000

TP + TN

10000 - 1

Act.

	Cat	Dog
Pred Cat	9500	500
Dog	500	0

$$\frac{9500 + 0}{10500}$$

Accuracy = 92 ... %

Naive Bayes

↳ Probabilistic algo

$$P(Y/x) = \frac{P(x/Y)\,P(Y)}{P(X)}$$

$$\boxed{P(A/B) = \frac{P(B/A)\,P(A)}{P(B)}}$$

x_1, x_2

⇒ Probability of A happening given B has occured.

⇒ B is evidence for A being hypothesis.

① naive
② All predictors will have an equal effect on the outcome (Ind. variable)

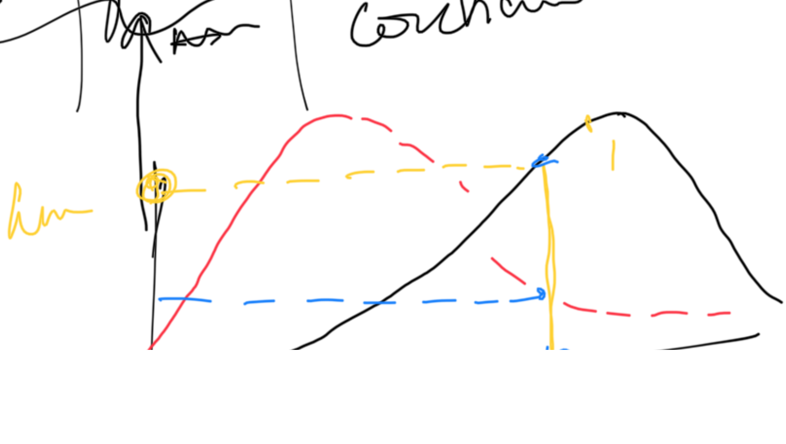
⇒ Gaussian Naive Bayes

Y/Rain/Not	humidity	cloudy
Yes	~~Yes~~	~~...~~
Not	~~No~~	

continous.

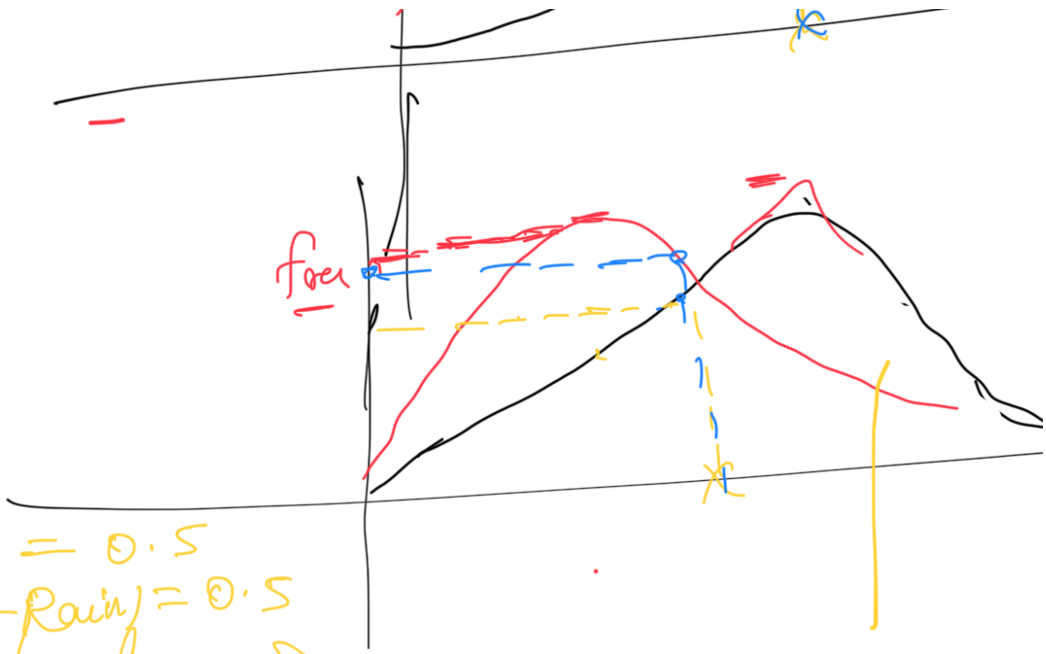
Rain

Yes No humidity



Cloudy

Yes No



P(Rain) = 0.5

P(Not-Rain) = 0.5

P(Rain happening /

→ (P(Rain) × P(humidity / Rain)

0.

× P(cloudy / Rain))

→ (0.5 × 0.06 × 0.0004

Log

When you a really small

Log( ) to prevent underflow

Log(0.5) + Log(0.06) + Log(0.0004)

0.000000

-0.70 + -6.6 + -50

→ -160

?

P(No. Rain) = 0.5 → −60

greater.

→ No-Rain

P-value Sum:-

warranty of 5 years

C & T

→ Engineer believes the engine will malfun

'<' 5 years.

40 cars → 4.8 yrs — 0.50

① State & alternate hypo

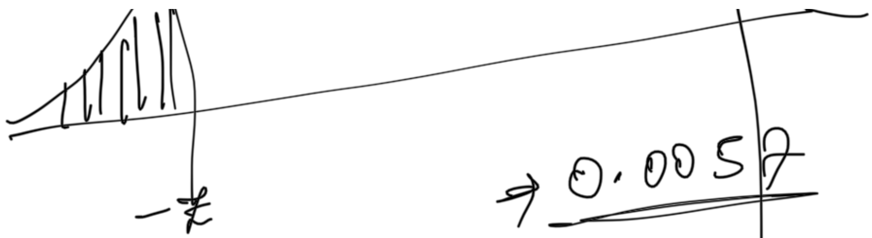
① $\mu \geq 5$; H_0

& $\mu < 5$; H_1

② C.I = 98% LOS = 2%

$$Z_c = \frac{4.8 - 5}{0.50/\sqrt{40}}$$

$$= \frac{-0.2}{0.079} = -2.53$$



→ 0.0057

[0.0057 < 0.02

[p-val < α]

Reject H_0

+ ROC-AUC
+ Ridge / Lasso / Elastic Handlers
+ Decision Trees → Random Forest