

Lecture-4 (Performance Measures)

Friday, 11 August 2023 4:30 PM

Actual	Predicted	
	+	-
+	TP	FN
-	FP	TN

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

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

- $0 < P < 1, 0 < R < 1$
- Precision is undefined : $TP + FP = 0$
→ Classifier predicts everything as -
- Recall is undefined : $TP + FN = 0$
→ Samples are from - class only
- $P = 1 \Rightarrow \frac{TP}{TP + FP} = 1 \Rightarrow TP = TP + FP \Rightarrow FP = 0$
- $R = 1 \Rightarrow FN = 0$
- ~~$\frac{P+R}{2}$~~ $F1\text{-score} = \frac{2PR}{P+R} = \frac{TP}{TP + \frac{1}{2}(FP+FN)}$

Multi-Class Classification - Performance Measures

		Urgent	Actual Normal	Spam
	Urgent	8	10	1
P R D	Normal	5	60	50

Confusion matrix

1 - vs - rest

	Actual		
Spam	3	30	200
	↓	↓	↓

		Actual	
		Urgent	Not Urgent
Predicted	Urgent	8	11
	Not Urgent	8	340

$$P_{\text{Urgent}} = \frac{8}{8 + 11} = \frac{8}{19}$$

$$R_{\text{Urgent}} = \frac{8}{8 + 8} = \frac{8}{16}$$

		Actual	
		Spam	Not Spam
Predicted	Spam	200	33
	Not Spam	51	83

$$P_{\text{Spam}} = \frac{200}{233} \approx 0.91$$

$$R_{\text{Spam}} = \frac{200}{251} \approx 0.88$$

		Actual	
		Normal	Not Normal
Predicted	Normal	60	55
	Not Normal	40	212

$$P_{\text{Normal}} = \frac{60}{115} \approx 0.52$$

$$R_{\text{Normal}} = \frac{60}{100} = 0.6$$

macro score

N_c
↖ D.

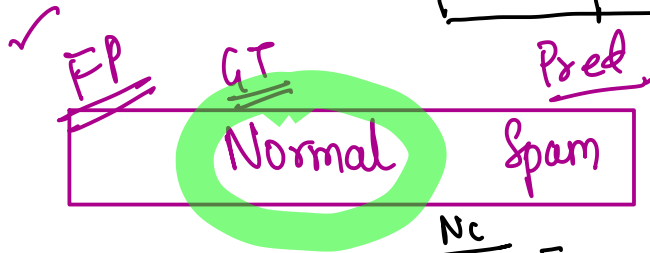
$$P_{\text{macro}} = \frac{P_{\text{Urgent}} + P_{\text{Normal}} + P_{\text{Spam}}}{3} = \frac{\sum_{i=1}^N r_i}{N}$$

pooled

		Actual ³	
		+	-
Pred	+	268	99
	-	99	35

$$P_{\text{micro}} = \frac{268}{268 + 99}$$

FP = FN



Spam

$$F1_{\text{micro}} = \frac{2 P_{\text{micro}} R_{\text{micro}}}{P_{\text{micro}} + R_{\text{micro}}}$$

scikit-learn

$$F1_{\text{macro}} = \frac{\sum_{i=1}^{N_c} F_i}{N_c}$$

ALT

$$\frac{2 P_{\text{macro}} R_{\text{macro}}}{P_{\text{macro}} + R_{\text{macro}}}$$

$$F1\text{-score} = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$

$$F1\text{-score}_{\text{micro}} = \frac{\sum_{i=1}^{N_c} TP_i}{\text{sample-centric}}$$

$$\sum_{i=1}^N TP_i + \frac{1}{2} (\sum FP_i + \sum FN_i)$$

	TP	FP	Prec	
A	1	1	0.5	$Prec_{macro} \approx 0.62$
B	1	1	0.5	$Prec_{micro} = \frac{93}{93+4}$
C	90	1	≈ 1	
D	1	1	0.5	≈ 0.95

* $n/2$ Classifier: Random Guessing

- $n/2$ $P(\text{Pred} = +) = \left(\frac{1}{2}\right)$
 $P(\text{Pred} = -) = \frac{1}{2}$

Accuracy = $\frac{\text{\# of correct predictions}}{n}$

$\xrightarrow{(n/2)}$
 $\frac{n/2}{n} = \frac{1}{2}$
 R.V

+ + + + + - - - - -
 ✓ + + + + + + + + + +
 ✓ + + + -

$E[\text{\# of correct predictions}]$
 \downarrow

✓
✓
✓
✓

binomial distribution

0 ... n [0, n]

n p

How many times do you
get k successes?

$$E[\cdot] = n \cdot p$$
$$= n \left(\frac{1}{2} \right)$$