

ML 평가지표

240217_7기_파이썬문법응용반

ML에서 모델을 평가하기 위해서 어떤 것을 기준으로 평가하였나?
실제값과 예측값을 두 개를 비교하는 한다.

실제는 1,0,1
모델이 예측한 값도 1,0,1 인가? 아니면 1,1,1 인가? 0,0,0 인가
이 결과를 보고 실제 값과 비교하여 정말 잘 맞았는지를 평가하는 것

임계값은 디폴트로 생각을 안 하지만 0.5 1과 0으로 구분 하는 것
임계값이 1 되거나 0되면 어떤 결과가 일어날까? (이진분류1,0 분류할 때)

임계값은 1이라는 것은 1만 1로 예측한다는 것 (기준 엄격하다.)
0이면 거의 대부분을 다 1로 예측한다는 것 (기준이 엄격하지 않다.)

모델에서 머신이 예측한 값들 ? 어떤 식으로 나올까-?
1,0,1 을 예측하는 이진 분류라면 예측 값은? 0.75(1) , 0.28(0), 0.89(1)

혼동 행렬 Confusion Matrix

예측

실제값

		Predicted		
		Negative (0)	Positive (1)	
Actual	Negative (0)	True Negative TN	False Positive FP (Type I error)	Specificity $= \frac{TN}{TN + FP}$
	Positive (1)	False Negative FN (Type II error)	True Positive TP	Recall, Sensitivity, True positive rate (TPR) $= \frac{TP}{TP + FN}$
		Accuracy $= \frac{TP + TN}{TP + TN + FP + FN}$		Precision, Positive predictive value (PPV) $= \frac{TP}{TP + FP}$
				F1-score $= 2 \times \frac{Recall \times Precision}{Recall + Precision}$

0, 1 N, P 의미가 없음 그냥 반대로 생각할 수도 있다.

정확도 Accuracy

TP+TN 이 모델이 전체 관측치 데이터 중에서 1은 1로 0은 0 으로 모델이 올바르게 분류한 비중

클래스가 불균형한 경우는 대부분 문제가 발생한다.

만약 클래스 불균형인 상태로 그냥 train,test 어떤 한 데이터를 나눈 구간에서는 0에만 쏠려 있을 수 있다.

1을 예측하고 싶은데 0 을 예측한 것도 정확도에 포함이 되니 0만 예측했는데 정확도가 높게 나온다.

클래스가 불균형일 경우는 정확도가 잘못된 결과가 나올 수 있다.

Accuracy		Actual	
$\frac{TP + TN}{TP + FP + FN + TN}$		Positive	Negative
Predicted	Positive	TP	FP
	Negative	FN	TN

오분류율 (Error rate)

$$\frac{\text{오분류율}}{\text{(Error Rate)}} = \frac{FP + FN}{TP + FN + FP + TN}$$

정밀도 (Precision)

정밀도는 1로 예측하여서 분류한 관측치 결과 중에서 실재값도 1인 비중인 경우
정밀도 하나만 보게 되면 문제가 될 수 있다.

만약 모델이 하나의 관측치만 1로 분류했고, 실재값도 이게 1이면 모델의 정밀도가 100%

$$Precision = \frac{tp}{tp + fp}$$

$$Recall = \frac{tp}{tp + fn}$$

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn}$$

재현율 Recall

민감도라고도 부른다. 실제 1인 관측치 중에서 모델이 정확히 1로 예측한 것을 분류하는 비중
정밀도 기준으로 볼 때 아무리 예가 우수해도 실제 1인 관측치가 너무 적게 찾아내면 좋은 모델은 아니다.

$$Precision = \frac{tp}{tp + fp}$$

$$Recall = \frac{tp}{tp + fn}$$

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn}$$

특이도 (Specificity)

TN

FP + TN

F1-스코어

정밀도와 민감도 동시에 고려한 모델 0,1 사이의 값이 나온다.
정밀도와 민감도의 Trade off 관계에서 값이 변경이 된다.

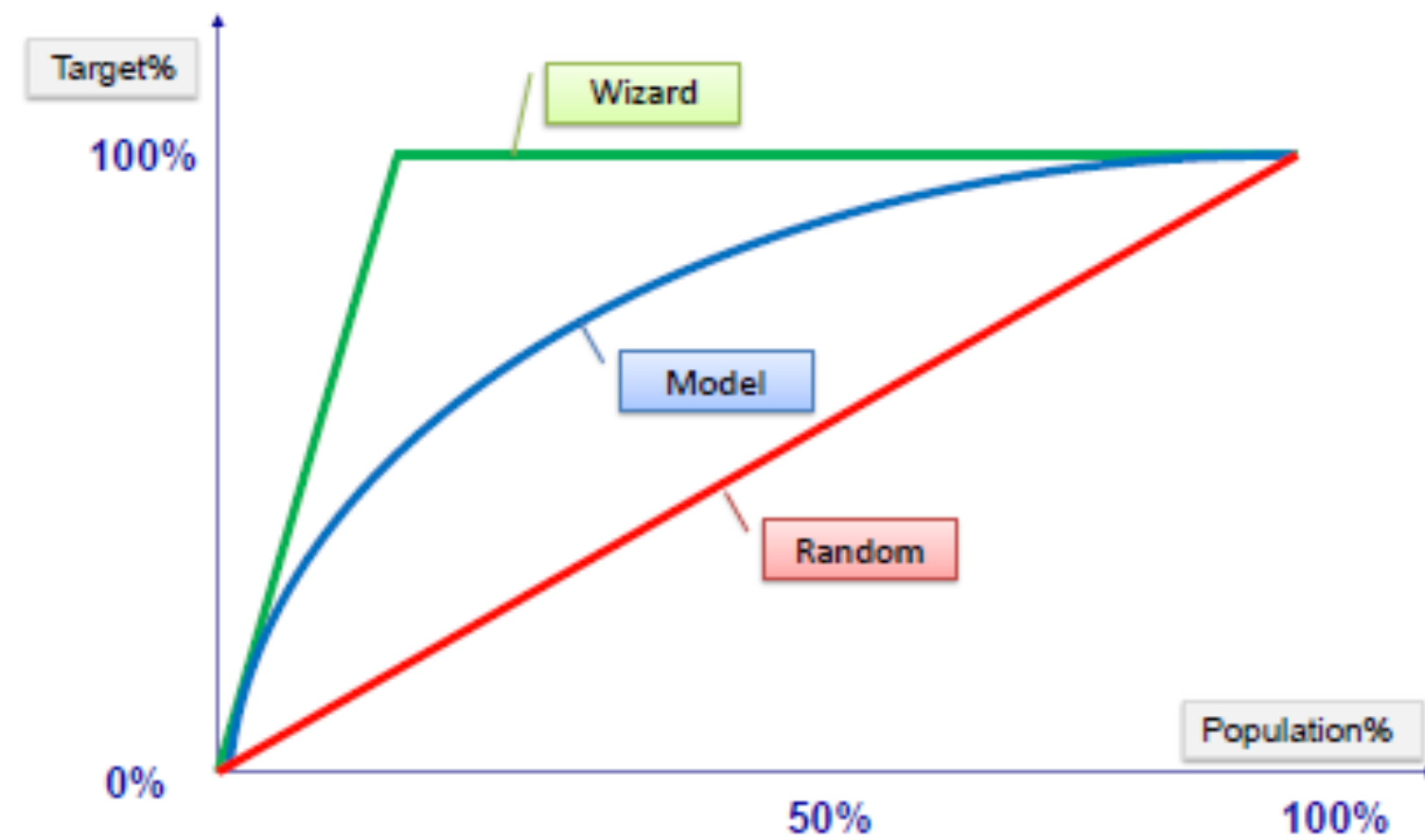
정밀도 높인다 -> 실제로 1일 것인 거의 확실한 관측치만 1로 분류
전체 실제 1중에서 1로 분류되는 비중인 민감도(재현율) 감소할 수 밖에 없다.

$$F1 = \frac{2}{\frac{1}{precision} + \frac{1}{recall}} = \frac{2 * (precision * recall)}{precision + recall}$$

$$F\beta = \frac{(1 + \beta^2) * (precision * recall)}{\beta^2 * precision + recall}$$

정밀도와 민감도 한쪽으로 치우치지 않을 때 상대적으로 높은 값을 가지고 온다.
정밀도 중요하다면 1미만, 재현율 중요하다면 1초과로

향상도 차트 Lift Charts



Example:

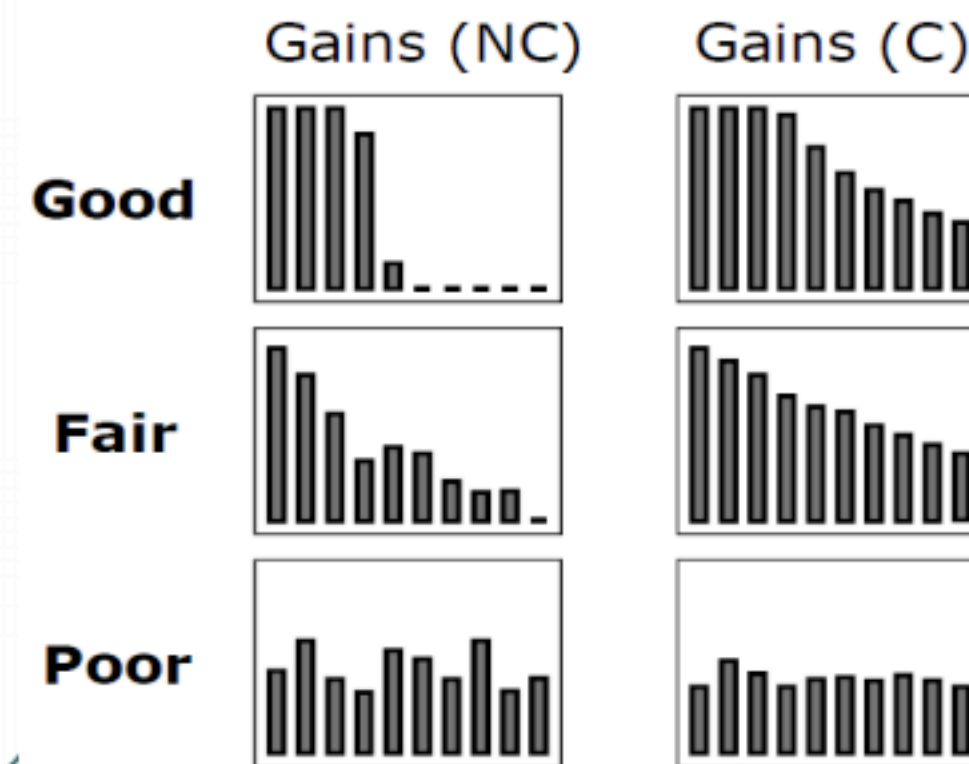
Score Table		Sorted by Score		Gain Table		Lift Table	
Target	Score	Target	Score	Count%	Target%	Count%	Lift
0	235	1	880	10	36	10	3.6
1	724	1	724	20	54	20	2.7
1	556	1	676	30	66	30	2.2
0	345	1	556	40	76	40	1.9
0	480	0	480	50	85	50	1.7
1	676	0	368	60	90	60	1.5
0	195	0	345	70	94	70	1.3
1	880	0	235	80	98	80	1.2
0	368	0	195	90	100	90	1.1
...	100	100	100	1

향상도 테이블(Lift Table)

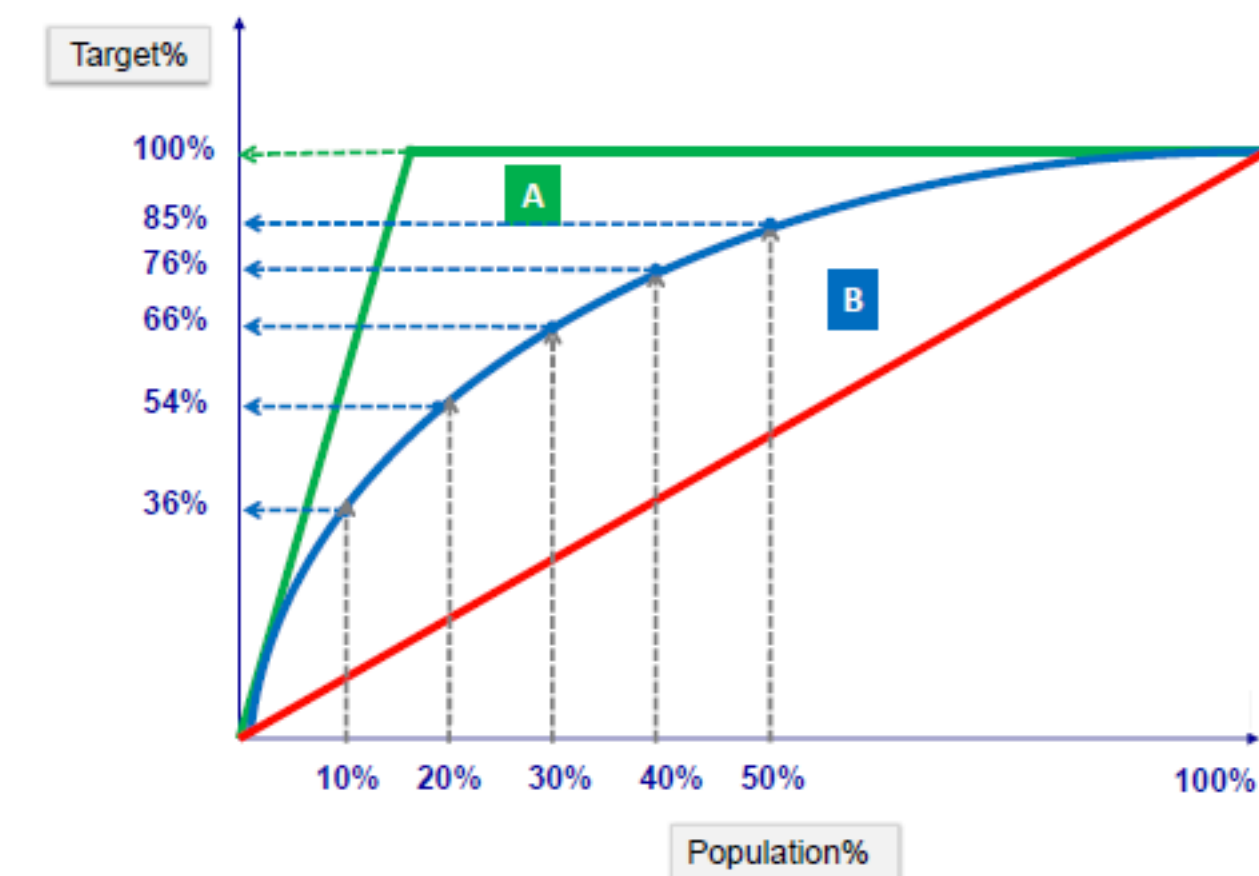
« Example » n = 2000, 1 = 381

Baseline = $381/2000 = 19\%$

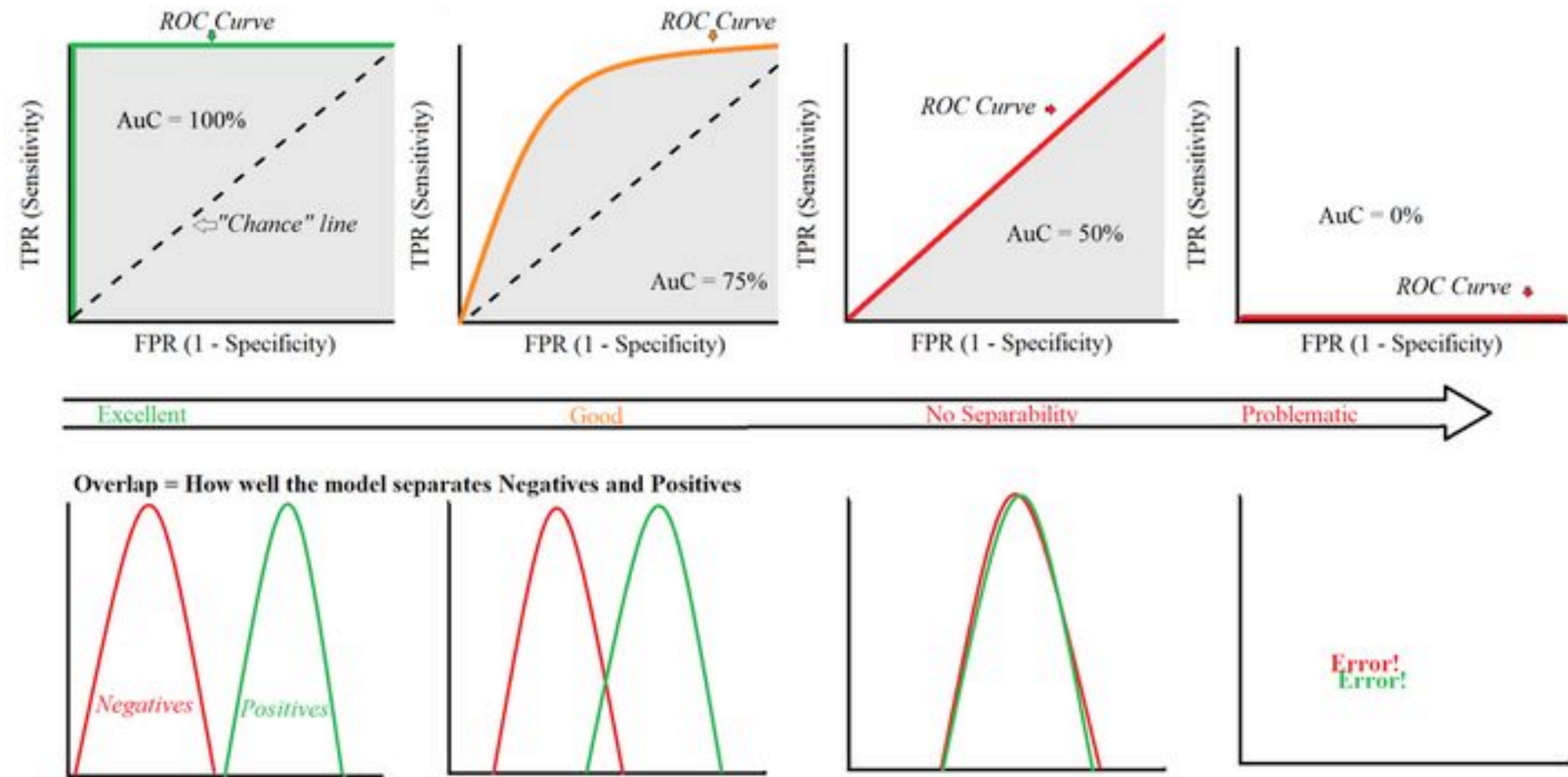
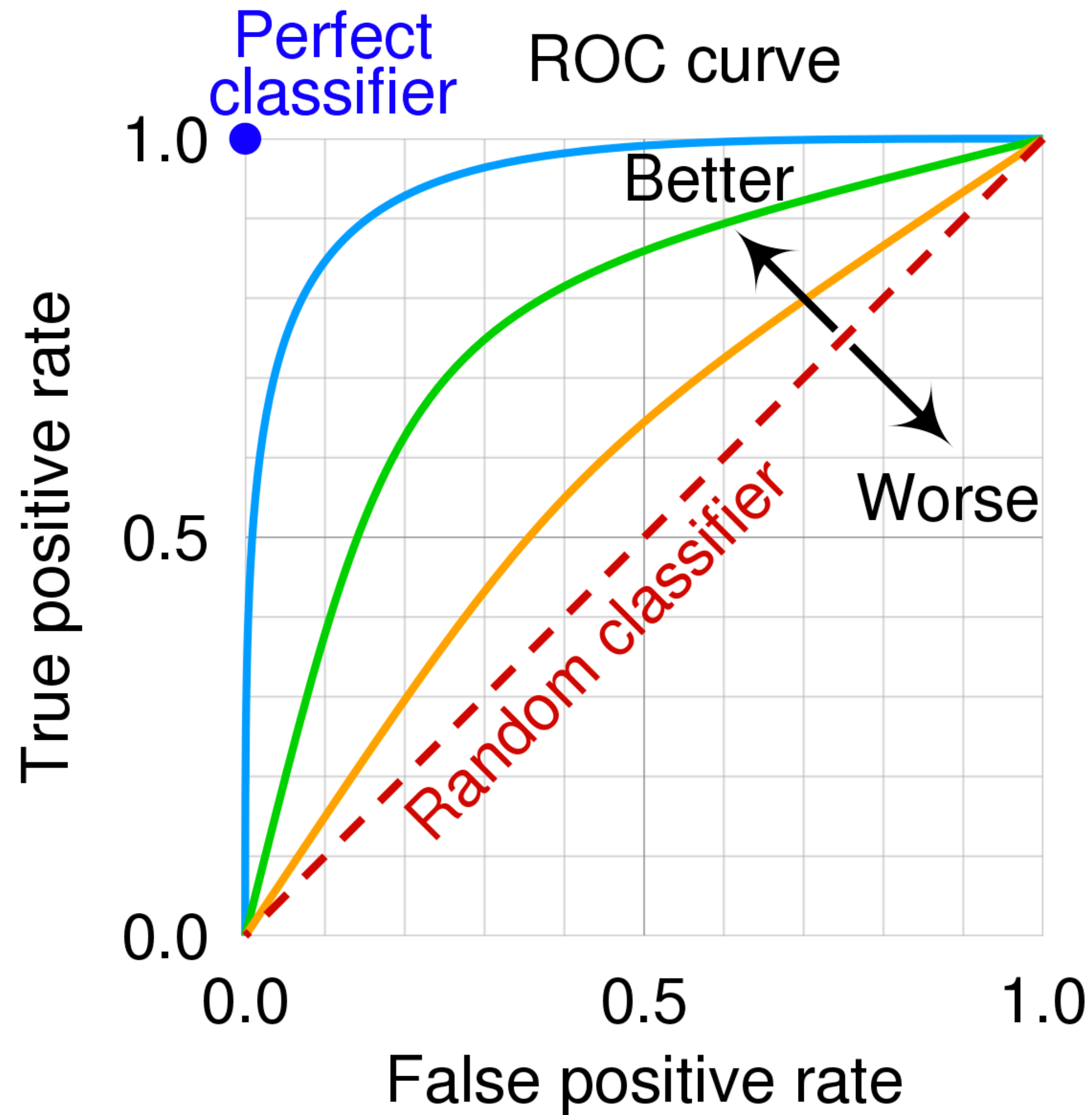
Decile	Y=1	%Captured	%Response	Lift
1	174	$174/381=45.6$	$174/200=87.0$	$87.0/19=4.57$
2	110	$110/381=28.8$	$110/200=55.0$	$55.0/19=2.89$
3	38	$38/381=9.9$	$38/200=19.0$	$19.0/19=1.00$
4	14	$14/381=3.6$	$14/200=7.0$	$7.0/19=0.36$
5	11	$11/381=2.8$	$11/200=5.5$	$5.5/19=0.28$
6	10	$10/381=2.6$	$10/200=5.0$	$5.0/19=0.26$
7	7	$7/381=1.8$	$7/200=3.5$	$3.5/19=0.18$
8	10	$10/381=2.6$	$10/200=5.0$	$5.0/19=0.26$
9	3	$3/381=0.7$	$3/200=1.5$	$1.5/19=0.07$
10	4	$4/381=1.0$	$4/200=2.0$	$2.0/19=0.10$



Gain Chart



ROC곡선-AUC



		Predicted condition			
Actual condition	Total population = P + N	Positive (PP)	Negative (PN)	Informedness, bookmaker informedness (BM) = TPR + TNR - 1	Prevalence threshold (PT) $= \frac{\sqrt{TPR \times FPR} - FPR}{TPR - FPR}$
	Positive (P)	True positive (TP), hit	False negative (FN), type II error, miss, underestimation	True positive rate (TPR), recall, sensitivity (SEN), probability of detection, hit rate, power $= \frac{TP}{P} = 1 - FNR$	False negative rate (FNR), miss rate $= \frac{FN}{P} = 1 - TPR$
	Negative (N)	False positive (FP), type I error, false alarm, overestimation	True negative (TN), correct rejection	False positive rate (FPR), probability of false alarm, fall-out $= \frac{FP}{N} = 1 - TNR$	True negative rate (TNR), specificity (SPC), selectivity $= \frac{TN}{N} = 1 - FPR$
	Prevalence $= \frac{P}{P + N}$	Positive predictive value (PPV), precision $= \frac{TP}{PP} = 1 - FDR$	False omission rate (FOR) $= \frac{FN}{PN} = 1 - NPV$	Positive likelihood ratio (LR+) $= \frac{TPR}{FPR}$	Negative likelihood ratio (LR-) $= \frac{FNR}{TNR}$
	Accuracy (ACC) $= \frac{TP + TN}{P + N}$	False discovery rate (FDR) $= \frac{FP}{PP} = 1 - PPV$	Negative predictive value (NPV) = $\frac{TN}{PN}$ = 1 - FOR	Markedness (MK), deltaP (Δp) = PPV + NPV - 1	Diagnostic odds ratio (DOR) = $\frac{LR+}{LR-}$
	Balanced accuracy (BA) $= \frac{TPR + TNR}{2}$	F ₁ score $= \frac{2PPV \times TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN}$	Fowlkes–Mallows index (FM) $= \sqrt{PPV \times TPR}$	Matthews correlation coefficient (MCC) $= \sqrt{TPR \times TNR \times PPV \times NPV} - \sqrt{FNR \times FPR \times FOR \times FDR}$	Threat score (TS), critical success index (CSI), Jaccard index $= \frac{TP}{TP + FN + FP}$