

# Đánh giá hiệu quả mô hình phân lớp

Nguyễn Thanh Tùng, **Trần Thị Ngân**

Khoa Công nghệ thông tin – Đại học Thủy lợi

*tungnt@tlu.edu.vn*, [ngantt@tlu.edu.vn](mailto:ngantt@tlu.edu.vn)



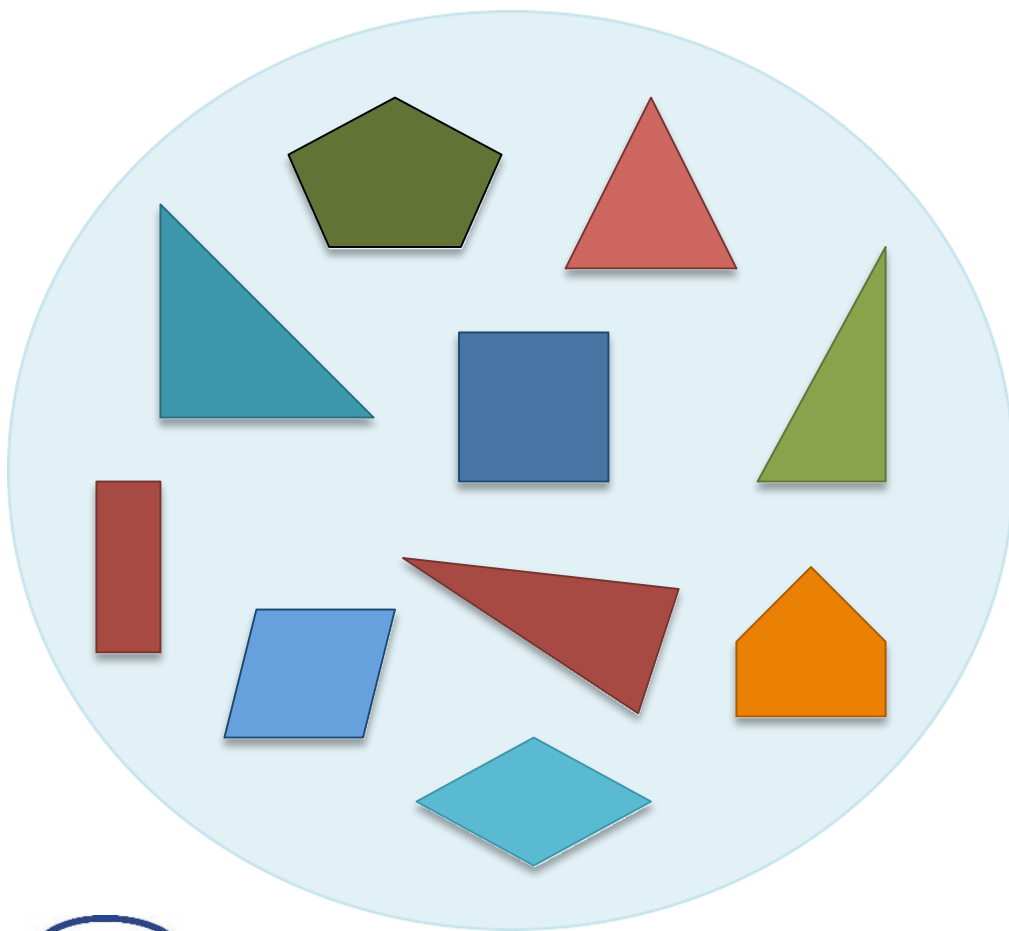
# Phân lớp

- Học có giám sát: Học từ các mẫu đã gán nhãn
- Biến đích có dạng rời rạc / hạng mục
- Mục tiêu: dự đoán biến đích có kiểu rời rạc
  - Gán mỗi mẫu cho 1 lớp
  - Một vài thuật toán phân lớp: **K-NN**, **CART**, Bagging, Random Forests

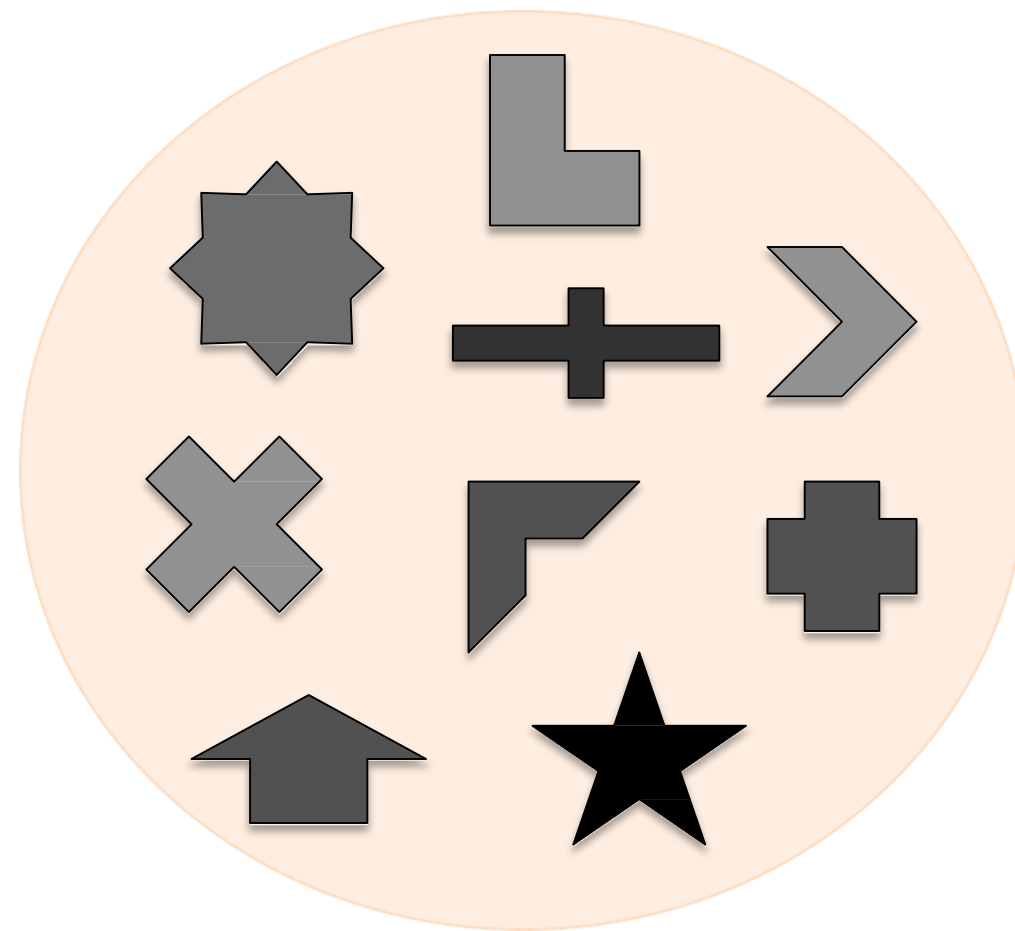


# Học từ mẫu đã gán nhãn

Lớp “+”



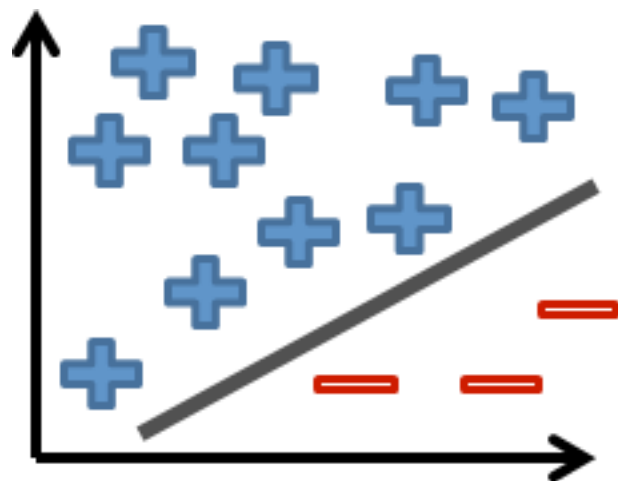
Lớp “-”



# Nhãn mất cân bằng

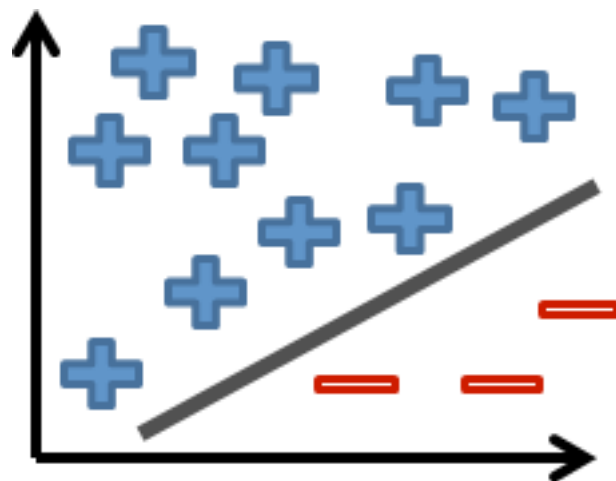
- Nhãn mất cân bằng (Imbalanced classes): lớp dương (+) xuất hiện với tần suất nhiều hơn lớp âm (-) trong tập dữ liệu huấn luyện

Vd: phát hiện gian lận, dữ liệu y học



# Nhấn mất cân bằng

- Tại sao đây là vấn đề?
  - Các thuật toán thực hiện tốt khi huấn luyện trên các mẫu trong mỗi lớp
  - Hiệu quả thấp trên các lớp có ít đại diện



# Đánh giá hiệu quả bộ phân lớp

- Hiệu năng của một mô hình thường được đánh giá dựa trên tập dữ liệu kiểm thử (test data).
- Giả sử  **$y_{pred}$**  là vector dự đoán đầu ra với mỗi phần tử là class được dự đoán của một điểm dữ liệu trong tập kiểm thử.
- Ta cần so sánh giữa vector dự đoán  **$y_{pred}$**  này với vector class thật của dữ liệu, được mô tả bởi vector  **$y_{true}$** .

# Đánh giá hiệu quả bộ phân lớp

- Ví dụ với bài toán có 3 lớp dữ liệu được gán nhãn là 0, 1, 2.
- Giả sử các class được đánh số từ 0 đến  $C-1$  trong trường hợp có  $C$  lớp dữ liệu. Có 10 điểm dữ liệu trong tập kiểm thử với các nhãn thực sự được mô tả bởi  
 $y_{\text{true}} = [0, 0, 0, 0, 1, 1, 1, 2, 2, 2]$ .
- Giả sử bộ phân lớp chúng ta đang cần đánh giá dự đoán nhãn cho các điểm này là  
 $y_{\text{pred}} = [0, 1, 0, 2, 1, 1, 0, 2, 1, 2]$ .

# Đánh giá hiệu quả bộ phân lớp

- Trong bài toán hồi quy, chúng ta đã dùng MSE, RMSE, MAE,  $R^2$  để đánh giá hiệu quả thuật toán.
- Với bài toán phân lớp, chúng ta cần độ đo để đánh giá hiệu quả của mô hình phân lớp

**Độ chính xác (Accuracy), Ma trận nhầm lẫn (Confusion matrix), Độ chính xác/Hồi tưởng (Precision/Recall), Độ nhạy/Độ đặc hiệu (Sensitivity/ Specificity), Đường cong ROC (ROC curve)**





# Đánh giá hiệu quả bộ phân lớp

- **Accuracy:** Cách đánh giá này đơn giản tính tỉ lệ giữa số điểm được dự đoán đúng và tổng số điểm trong tập dữ liệu kiểm thử.
- **Accuracy** chỉ cho ta biết được bao nhiêu phần trăm lượng dữ liệu được phân loại đúng.



# Đánh giá hiệu quả bộ phân lớp

## Ví dụ:

Trong ví dụ trước, ta có thể đếm được có 6 điểm dữ liệu được dự đoán đúng trên tổng số 10 điểm. Vậy ta kết luận độ chính xác của mô hình là 0.6 (hay 60%).



# Đánh giá hiệu quả bộ phân lớp

- **Confusion matrix:** Chỉ ra được cụ thể mỗi loại được phân loại như thế nào, lớp nào được phân loại đúng nhiều nhất, và dữ liệu thuộc lớp nào thường bị phân loại nhầm vào lớp khác.



# Đánh giá hiệu quả bộ phân lớp

**Confusion matrix:** Là một ma trận vuông với số chiều bằng số lượng lớp dữ liệu. Giá trị tại hàng thứ  $i$ , cột thứ  $j$  là số lượng điểm lẽ ra thuộc vào class  $i$  nhưng lại được dự đoán là thuộc vào class  $j$ . Như vậy, nhìn vào hàng thứ nhất (0), ta có thể thấy được rằng trong số bốn điểm thực sự thuộc lớp 0, chỉ có hai điểm được phân loại đúng, hai điểm còn lại bị phân loại nhầm vào lớp 1 và lớp 2.

Total: 10	Predicted as: 0	Predicted as: 1	Predicted as: 2	
True: 0	2	1	1	4
True: 1	1	2	0	3
True: 2	0	1	2	3

# Đánh giá hiệu quả bộ phân lớp

**Chú ý:** Việc biết một cột của confusion matrix này sẽ suy ra được cột còn lại vì tổng các hàng luôn bằng 1 và chỉ có hai lớp dữ liệu. **Với các bài toán có nhiều lớp dữ liệu**, ta có thể xây dựng bảng True/False Positive/Negative cho **mỗi lớp** nếu coi lớp đó là lớp *Positive*, các lớp còn lại gộp chung thành lớp *Negative*.



# Đánh giá hiệu quả bộ phân lớp

- True/False Positive/Negative
- Xét bài toán phân lớp **nhị phân**: Có 2 lớp (+) và (–)
- Trong hai lớp dữ liệu này có một lớp nghiêm trọng hơn lớp kia và cần được dự đoán chính xác.

# Đánh giá hiệu quả bộ phân lớp

- Ví dụ:
  - Ví dụ như việc xác định mail spam, việc nhầm mail quan trọng thành mail spam nguy hiểm hơn là bỏ sót mail spam.
  - Trong bài toán xác định có mìn dưới lòng đất hay không thì việc bỏ sót nghiêm trọng hơn việc báo động nhầm rất nhiều.



# Đánh giá hiệu quả bộ phân lớp

- Định nghĩa lớp dữ liệu quan trọng hơn cần được xác định đúng là lớp **Positive** (P-dương tính), lớp còn lại được gọi là **Negative** (N-âm tính).
- Định nghĩa True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN) dựa trên confusion matrix như sau:





		Lớp dự đoán (Predicted class)	
		+	-
Lớp thực (True class)	+	True Positive-TP	False Negative-FN Type II error
	-	(False Positive-FP) Type I error	True Negative-TN



# Đánh giá hiệu quả bộ phân lớp

- Bệnh nhân được chẩn đoán là có bệnh và thực tế đúng là như vậy: **True Positive (TP), dương tính thật.**
- Chẩn đoán là có bệnh nhưng thực tế không phải vậy: **False Positive (FP), dương tính giả.**
- Bệnh nhân được chẩn đoán là không có bệnh nhưng thực tế người này đang có bệnh. Tình huống này được gọi là **âm tính giả, kí hiệu là FN.**
- Bệnh nhân được chẩn đoán là không có bệnh và thực tế đúng như vậy. Tình huống này được gọi là **âm tính thật, kí hiệu là TN.**

# Đánh giá hiệu quả bộ phân lớp

		Actual Values	
		1	0
Predicted Values	1	<b>TRUE POSITIVE</b> 	<b>FALSE POSITIVE</b>  <b>TYPE 1 ERROR</b>
	0	<b>FALSE NEGATIVE</b>  <b>TYPE 2 ERROR</b>	<b>TRUE NEGATIVE</b> 

# Ví dụ

## True Positive (TP):

- Sự thật: Có chó sói (Positive)
- Cậu bé chăn cừu: Cảnh báo dân làng là có chó sói (True)
- Kết quả: Cậu bé là anh hùng của cả làng

## False Positive (FP):

Sự thật: không có chó sói

Cậu bé chăn cừu: Cảnh báo dân làng là có chó sói (Positive)

Kết quả: Cậu bé bị dân làng la mắng (Vì đưa ra cảnh báo giả - kết luận Positive của cậu bé là sai – False Positive)

## False Negative (FN):

Sự thật: Có chó sói.

- Cậu bé: nói với dân làng là không có chó sói (Negative)
- Kết quả: Chó sói ăn hết bầy cừu của làng (Vì ko cảnh báo nên kết luận Negative của cậu bé là sai– False Negative)

## True Negative (TN):

- Sự thật: không có chó sói (Negative).
- Cậu nhé: nói với dân làng là không có chó sói
- Kết quả: Đàn cừu của dân làng không bị đe dọa

# Đánh giá hiệu quả bộ phân lớp

<div><div>True positive rate (TPR) <i>(recall, sensitivity)</i></div><div><div><div><div></div><div>Predicted class</div><div><div><div>+</div><div>-</div></div><div><div>True class</div><div><div><div>+</div><div>-</div></div></div><div><div>TP</div><div>FN</div><div>FP</div><div>TN</div></div></div></div></div><div><math display="block">TPR = \frac{TP}{TP + FN}</math></div></div></div></div>	<div><div>False negative rate (FNR)</div><div><div><div><div></div><div>Predicted class</div><div><div><div>+</div><div>-</div></div><div><div>True class</div><div><div><div>+</div><div>-</div></div></div><div><div>TP</div><div>FN</div><div>FP</div><div>TN</div></div></div></div></div><div><math display="block">FNR = \frac{FN}{TP + FN}</math></div></div></div></div>
<div><div>False positive rate (FPR)</div><div><div><div><div></div><div>Predicted class</div><div><div><div>+</div><div>-</div></div><div><div>True class</div><div><div><div>+</div><div>-</div></div></div><div><div>TP</div><div>FN</div><div>FP</div><div>TN</div></div></div></div></div><div><math display="block">FPR = \frac{FP}{FP + TN}</math></div></div></div></div>	<div><div>True negative rate (SPC) <i>(specificity)</i></div><div><div><div><div></div><div>Predicted class</div><div><div><div>+</div><div>-</div></div><div><div>True class</div><div><div><div>+</div><div>-</div></div></div><div><div>TP</div><div>FN</div><div>FP</div><div>TN</div></div></div></div></div><div><math display="block">SPC = \frac{TN}{FP + TN}</math></div></div></div></div>

# Đánh giá hiệu quả bộ phân lớp

- **False Positive Rate** còn được gọi là False Alarm Rate (tỉ lệ báo động nhầm)
- **False Negative Rate** còn được gọi là Miss Detection Rate (tỉ lệ bỏ sót).
- Trong bài toán dò mìn, thà báo nhầm còn hơn bỏ sót, tức là ta có thể chấp nhận **False Alarm Rate cao** để đạt được **Miss Detection Rate thấp**.



# Đánh giá hiệu quả bộ phân lớp

True positive rate (TPR)

(recall, sensitivity)

Predicted class

+

-

	+	-
True class +	TP	FN
True class -	FP	TN

$$TPR = \frac{TP}{TP + FN}$$

Positive predictive value (PPV) (*precision*)

Predicted class

+

-

	+	-
True class +	TP	FN
True class -	FP	TN

$$PPV = \frac{TP}{TP + FP}$$

False positive rate (FPR)

Predicted class

+

-

	+	-
True class +	TP	FN
True class -	FP	TN

$$FPR = \frac{FP}{FP + TN}$$

True negative rate (SPC)

(*specificity*)

Predicted class

+

-

	+	-
True class +	TP	FN
True class -	FP	TN

$$SPC = \frac{TN}{FP + TN}$$

ROC curve

Precision/recall



# Precision/recall

$$\text{Precision} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalsePositive}}$$

$$\text{Recall} = \frac{\text{TruePositive}}{\text{TruePositive} + \text{FalseNegative}}$$

# Ví dụ: Tính các tỷ lệ TP, TN, FP, FN và các độ đo Accuracy, Precision, Recall

		Actual	
Predict		Spam	Not Spam
	Spam	8	32
	Not Spam	2	8



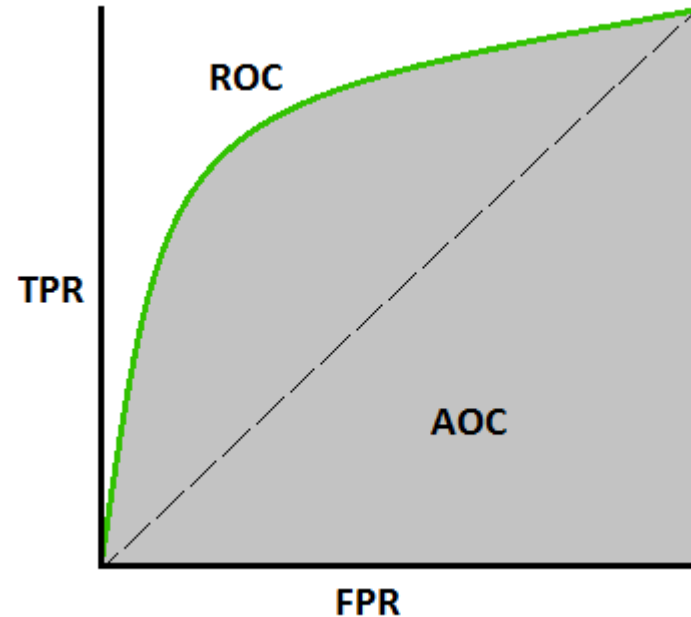
# Accuracy, Precision vs Recall

- **Accuracy:** đo tỷ lệ phần trăm các email được phân lớp đúng
- **Precision:** đo tỷ lệ phần trăm của các email được gắn cờ là thư rác được phân loại chính xác
- **Recall:** đo lường tỷ lệ phần trăm của các email spam thực tế đã được phân loại chính xác



# Đánh giá hiệu quả bộ phân lớp

- Đường cong ROC (receiver operating characteristic) là đường cong được vẽ với TPR so với FPR trong đó TPR nằm trên trục y và FPR nằm trên trục x

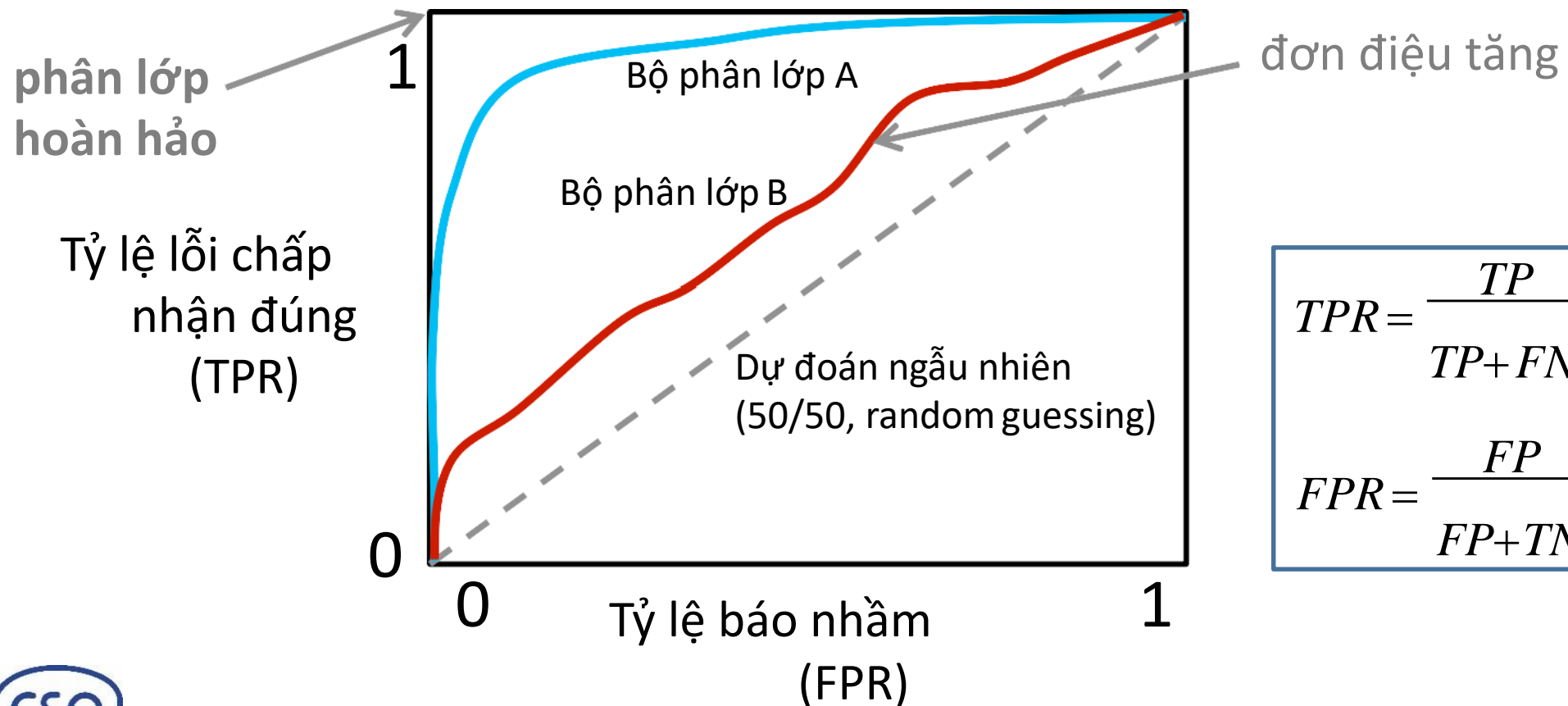


# Đánh giá hiệu quả bộ phân lớp

- **AUC (Area Under The Curve) or AUROC:**  
AUC càng cao, mô hình càng tốt khi dự đoán 0 là 0 và 1 là 1.

# Đánh giá hiệu quả bộ phân lớp

- Đường cong ROC (receiver operating characteristic)



$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{FP + TN}$$

# Đánh giá hiệu quả bộ phân lớp

- Nhược điểm của đường cong ROC
  - ROC không biểu thị đúng độ mất cân bằng các mẫu trong lớp thực
  - vd: Xét bộ dữ liệu có 1% mẫu thuộc lớp “+” và 99% mẫu thuộc lớp “–”
    - Giả sử ta nhận được kết quả phân lớp như sau:  
TPR = 0.9 và FPR = 0.12
    - TPR và FPR không hiểu thị được theo tính chất của đường cong ROC

		Predicted class	
		+	-
True class	+	90	10
	-	1188	8712

# Đánh giá hiệu quả bộ phân lớp

- Độ chính xác/Triệu hồi (Precision/recall)
  - Độ chính xác (Positive predictive value):  $PPV = \frac{TP}{TP+FP}$ 
    - Tỷ lệ phần trăm của số mẫu thuộc lớp (+) được dự đoán đúng trên số mẫu thực là (+)
  - Recall (True positive rate):  $TPR = \frac{TP}{TP+FN} = \frac{TP}{P}$ 
    - Tỷ lệ các mẫu (+) phân lớp chính xác lớp (+)
  - Recall và precision tỷ lệ nghịch với nhau
  - Với bộ phân lớp hoàn hảo, Recall = 1, Precision = 1
  - VD phân lớp mất cân bằng: Recall = 0.9, Precision = 0.07

		Predicted class	
		+	-
True class	+	90	10
	-	1188	8712