



Khoa Công Nghệ Thông Tin  
Trường Đại Học Cần Thơ



## Đánh giá hiệu quả của giải thuật học

Đỗ Thanh Nghi  
*dtngchi@cit.ctu.edu.vn*

Cần Thơ  
02-12-2008

# Nội dung

---

- Phương pháp cổ điển
- ROC, LIFT, COST

# Nghi thức kiểm tra

---

- nếu dữ liệu có 1 tập học và 1 tập kiểm tra sẵn dùng
  - dùng dữ liệu học để xây dựng mô hình, dùng tập kiểm tra để đánh giá hiệu quả của giải thuật
- nếu dữ liệu không có 1 tập kiểm tra sẵn
  - sử dụng nghi thức k-fold : chia tập dữ liệu thành k phần (fold) bằng nhau, lặp lại k lần, mỗi lần sử dụng k-1 folds để học và 1 fold để kiểm tra, sau đó tính trung bình của k lần kiểm tra
  - nghi thức hold-out : lấy ngẫu nhiên 2/3 tập dữ liệu để học và 1/3 tập dữ liệu còn lại dùng cho kiểm tra, có thể lặp lại quá bước này k lần rồi tính giá trị trung bình

## Nghi thức kiểm tra

---

- nếu dữ liệu có số phần tử lớn hơn 300
  - sử dụng nghi thức k-fold với  $k = 10$
- nếu dữ liệu có số phần tử nhỏ hơn 300
  - sử dụng nghi thức leave-1-out (k-fold với  $k = \text{số phần tử}$ )

# Nội dung

---

- **Phương pháp cổ điển**
- ROC, LIFT, COST

## Confusion matrix (C) cho k lớp

---

dự đoán =>	1	...	k
1			
...			
k			

- ❑  $C[i, j]$ : số phần tử lớp **i** được giải thuật dự đoán là lớp **j**
- ❑  $C[i, i]$ : số phần tử phân lớp đúng
- ❑ Độ chính xác lớp **i**:  $C[i, i] / C[i, ]$
- ❑ Độ chính xác tổng thể:  $\sum C[i, i] / C$

- Phương pháp cổ điển
- ROC, LIFT, COST

## Confusion matrix (C) cho 2 lớp (+/-)

---

dự đoán =>	dương	âm
dương	TP	FN
âm	FP	TN

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

$$rec = \frac{tp}{tp + fn} = \frac{tp}{pos}$$

$$acc = \frac{tp + tn}{tp + fn + tn + fp} = \frac{tp + tn}{pos + neg}$$

$$F1 = \frac{2 \times prec \times rec}{prec + rec}$$

$$bep = \frac{prec + rec}{2}$$

# Dữ liệu không cân bằng

---

- nếu dữ liệu không cân bằng
  - lớp positive có số lượng rất nhỏ so với lớp negative
  - ví dụ : positive = 5%, negative = 95%
  - một giải thuật học có thể cho kết quả 95% độ chính xác khi phân loại, nhưng chúng ta có thể mất hoàn toàn lớp positive
- khả năng tách lớp positive từ lớp negative



# Nội dung

---

- Phương pháp cổ điển
- **ROC, LIFT, COST**

## ROC (Provost & Fawcett, 1997)

---

### ■ Receiver Operating Characteristic

- khả năng tách positive từ negative
- đồ thị 2 chiều : tpr, fpr
- $O(0, 1)$  : lý tưởng
- đường chéo ( $tpr = fpr$ ) : ngẫu nhiên
- sử dụng diện tích dưới đường ROC để đánh giá hiệu năng

## ROC (Provost & Fawcett, 1997)

---

### ■ Receiver Operating Characteristic

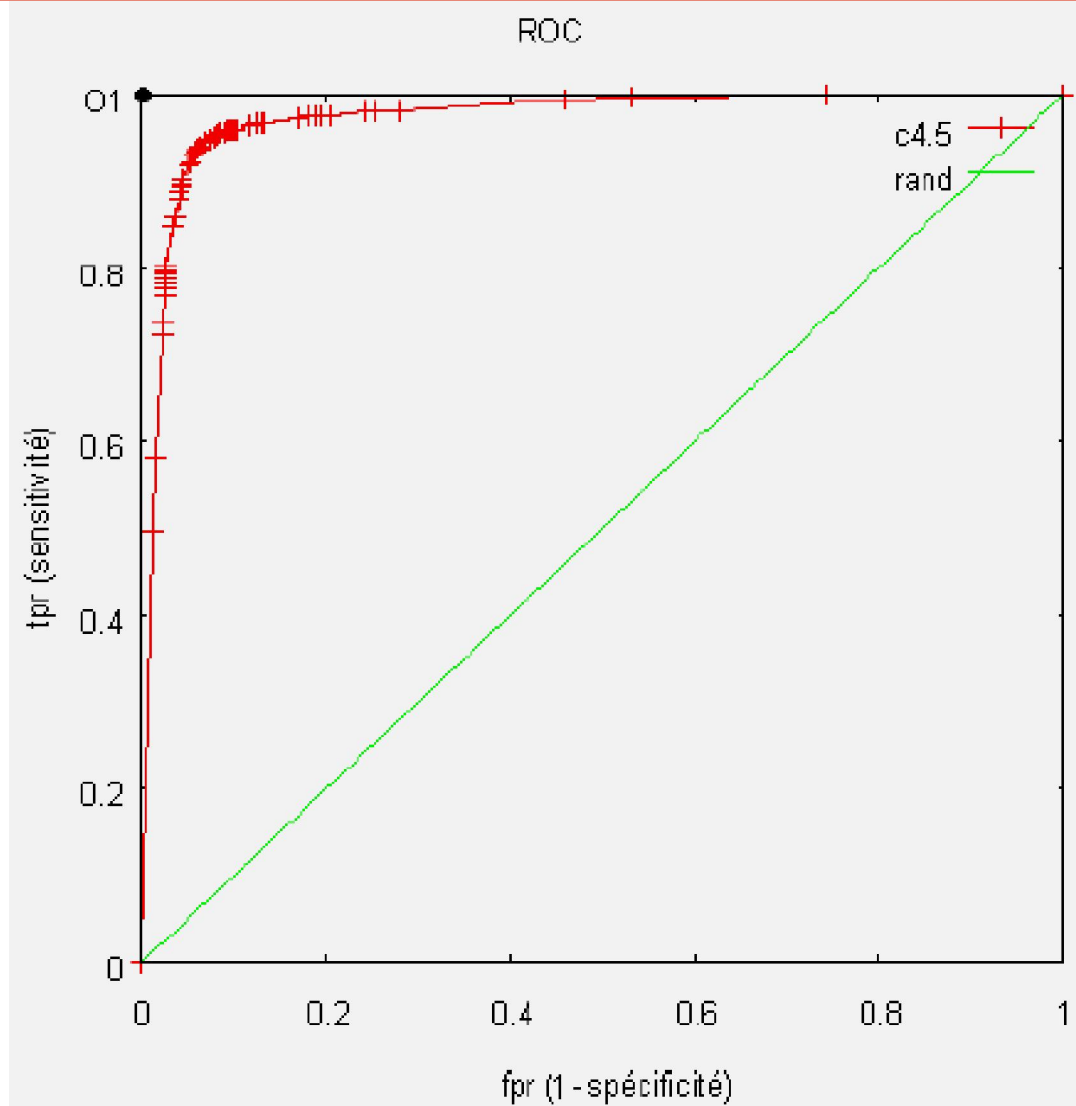
- khả năng tách positive từ negative
- đồ thị 2 chiều : tpr, fpr
- O(0, 1) : lý tưởng
- đường chéo (tpr = fpr) : ngẫu nhiên
- sử dụng diện tích dưới đường ROC để đánh giá hiệu năng

$$tpr = \frac{tp}{tp + fn} = \frac{tp}{pos}$$

$$fpr = \frac{fp}{fp + tn} = \frac{fp}{neg}$$

- Phương pháp cổ điển
- ROC, LIFT, COST

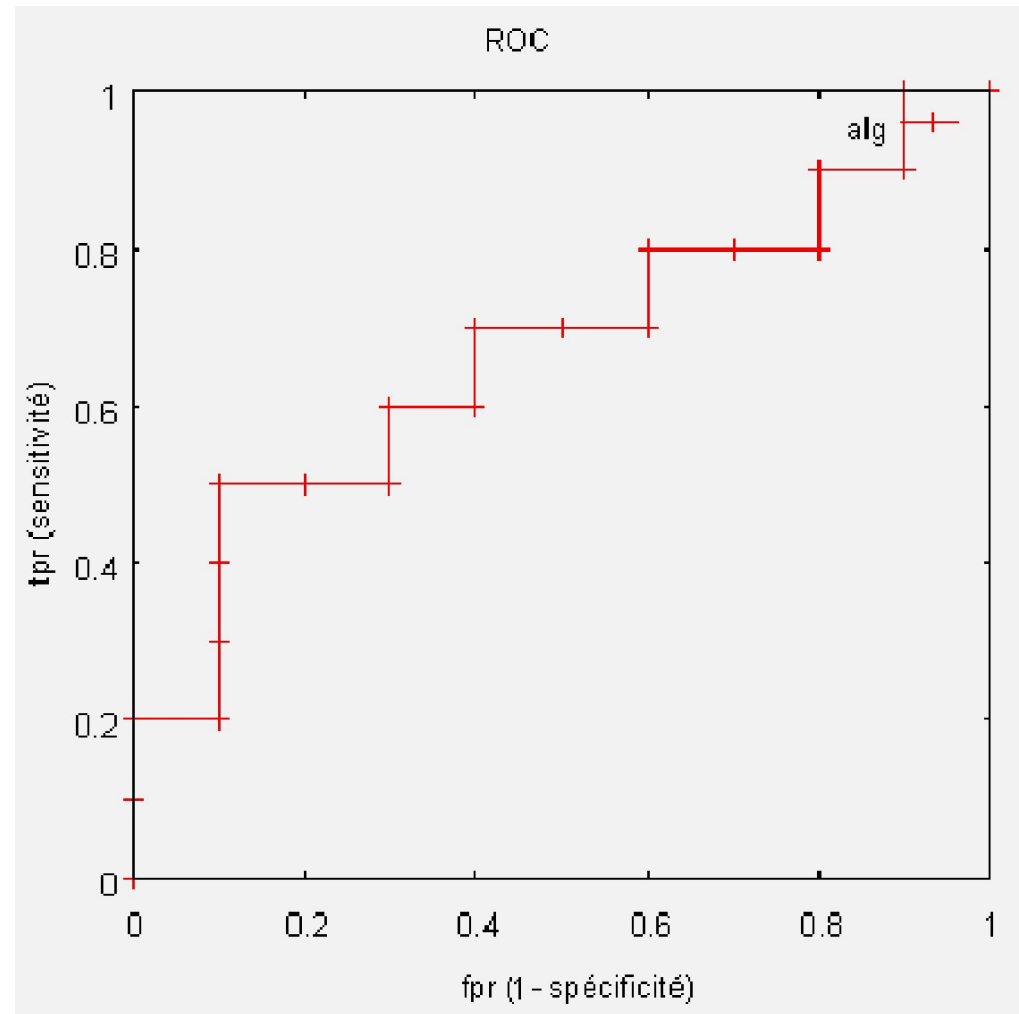
# ROC (Provost & Fawcett, 1997)



- Phương pháp cổ điển
- ROC, LIFT, COST

# ROC (Provost & Fawcett, 1997)

Order	Classes	Discriminant threshold
1	+1	0.9
2	+1	0.8
3	-1	0.7
4	+1	0.6
5	+1	0.55
6	+1	0.54
7	-1	0.53
8	-1	0.55
9	+1	0.51
10	-1	0.505
11	+1	0.4
12	-1	0.39
13	+1	0.38
14	-1	0.37
15	-1	0.36
16	-1	0.35
17	+1	0.34
18	-1	0.33
19	+1	0.3
20	-1	0.1



- Phương pháp cổ điển
- ROC, LIFT, COST

# LIFT

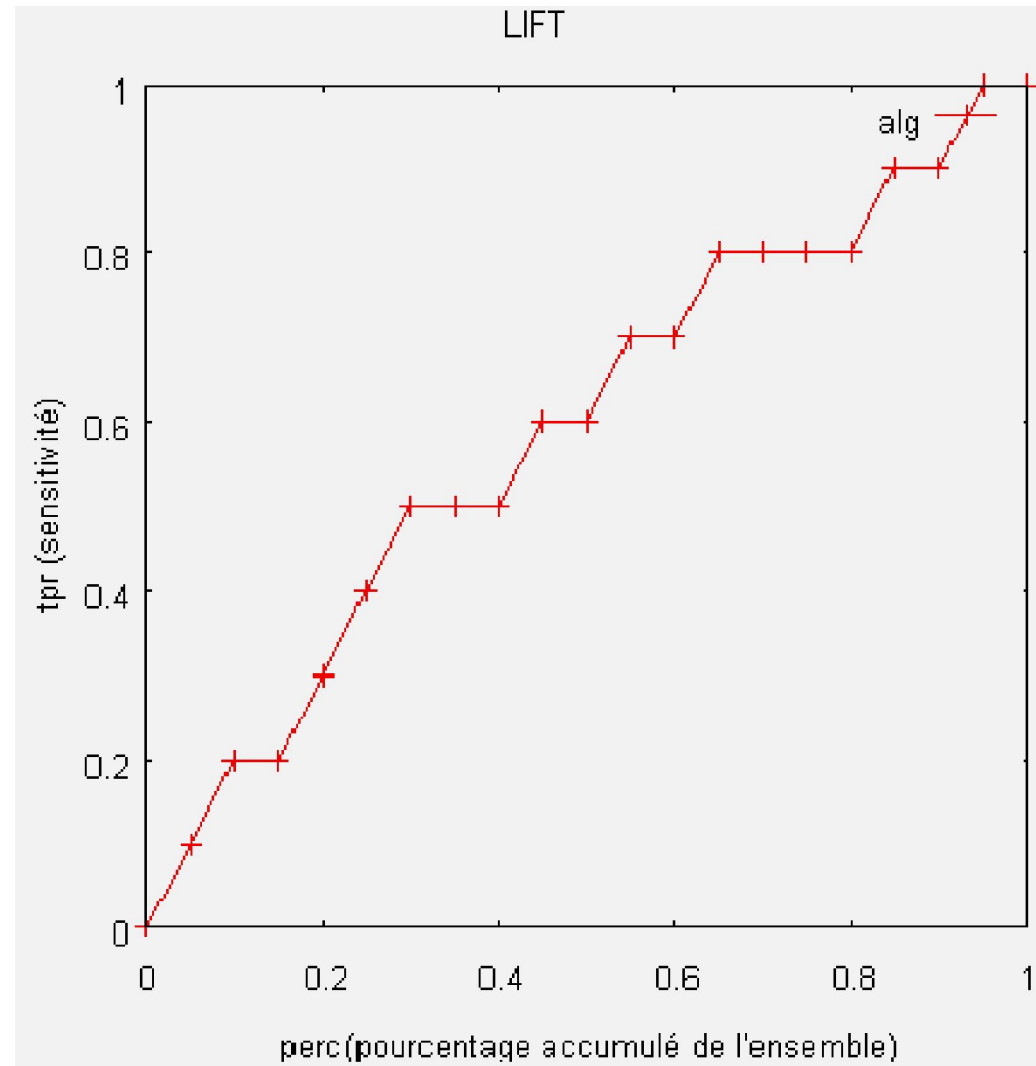
- đồ thị 2 chiều : tpr, perc. của tập dữ liệu

Order	Class	Discriminant threshold	% ensemble	% class +1
1	+1	0.9	5	10
2	+1	0.8	10	20
3	-1	0.7	15	20
4	+1	0.6	20	30
5	+1	0.55	25	40
6	+1	0.54	30	50
7	-1	0.53	35	50
8	-1	0.55	40	50
9	+1	0.51	45	60
10	-1	0.505	50	60
11	+1	0.4	55	70
12	-1	0.39	60	70
13	+1	0.38	65	80
14	-1	0.37	70	80
15	-1	0.36	75	80
16	-1	0.35	80	80
17	+1	0.34	85	90
18	-1	0.33	90	90
19	+1	0.3	95	100
20	-1	0.1	100	100

# LIFT

- đồ thị 2 chiều : tpr, perc. của tập dữ liệu

■ Phương pháp cổ điển  
■ ROC, LIFT, COST



## COST (Drummond et Holte, 2000)

---

### ■ COST

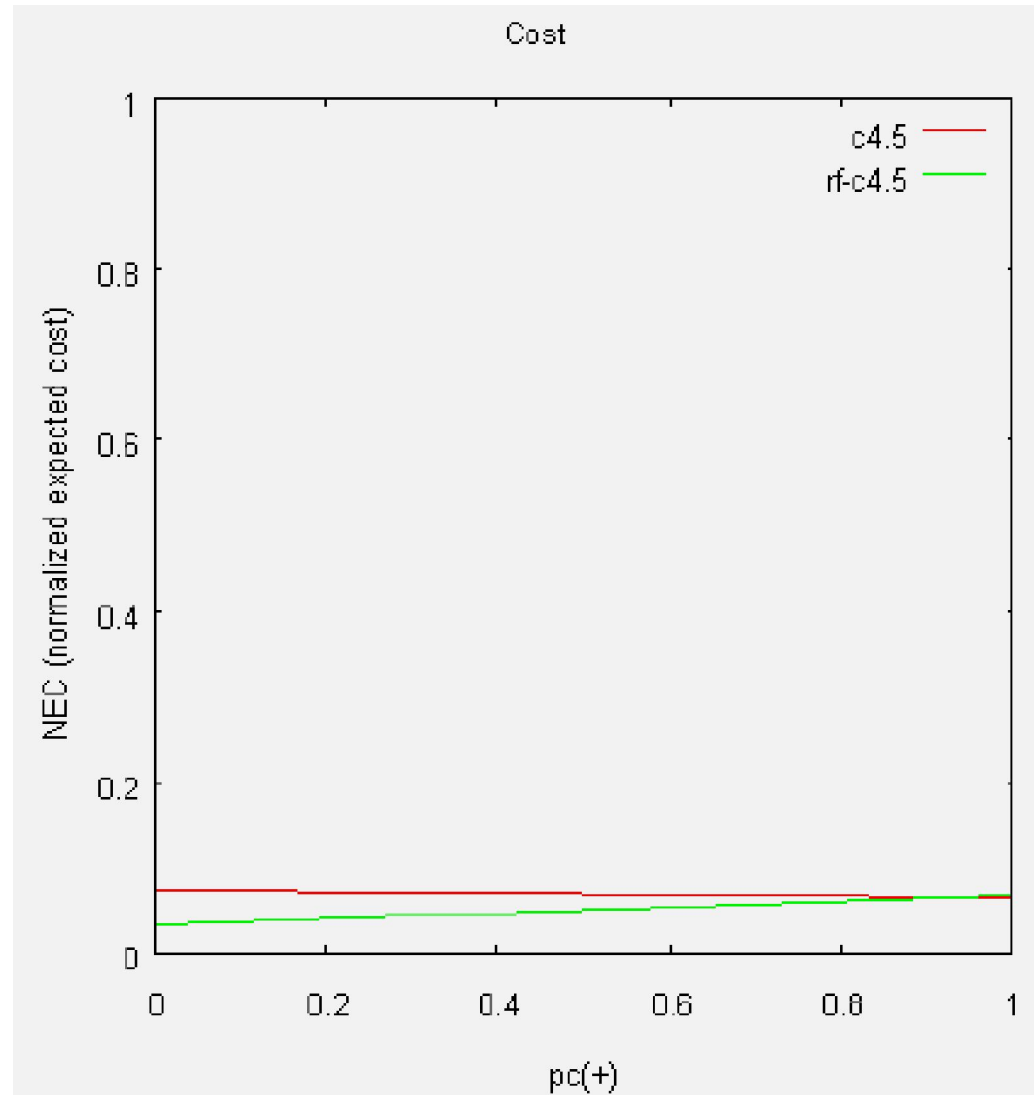
- được cho là hiệu quả hơn ROC ? ☺
- đồ thị 2 chiều : pc, nec
- vẽ COST qua 2 điểm : (0, fp) và (1, fn)
- dual với ROC

$$pc(+) = \frac{p(+)c(-/+)}{p(+)c(-/+) + (1 - p(+))c(+/-)}$$

$$nec = fn \times p(+) + fp(1 - pc(+))$$



# COST (Drummond et Holte, 2000)





Cám ơn !