



Khoa Công Nghệ Thông Tin Trường Đại Học Cần Thơ



Đánh giá hiệu quả của giải thuật học

Nội dung

- Nghi thức kiểm tra
- Các chỉ số đánh giá

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 ?

Nghi thức kiểm tra

- 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 $k = \text{số phần tử}$)

Chỉ số đánh giá

Confusion matrix (C) cho k lớp

dự đoán =>	1	...	k
1			
...			
k			

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

Confusion matrix (C) cho k lớp

dự đoán =>	Setosa	vesicolor	virginica
Setosa	15	0	0
vesicolor	0	16	2
virginica	0	3	14

❑ Độ chính xác lớp **i**: $C[i,i] / C[i,]$

❑ Setosa = ?

❑ Vesicolor = ?

❑ Virginica = ?

❑ Độ chính xác tổng thể: $\sum C[i,i] / C = ?$

Confusion matrix (C) cho k lớp

dự đoán =>	Setosa	vesicolor	virginica
Setosa	15	0	0
vesicolor	0	16	2
virginica	0	3	14

- ❑ $C[i, j]$: số phần tử lớp **i** (dòng) được giải thuật dự đoán là lớp **j** (**cột**)
- ❑ $C[i, i]$: số phần tử phân lớp đúng
- ❑ Độ chính xác lớp **i**: $C[i, i] / C[i,]$
 - ❑ Setosa = 15/15
 - ❑ Vesicolor = 16/18
 - ❑ Virginica = 14/17
- ❑ Độ chính xác tổng thể: $\sum C[i, i] / C = 45/50$

Confusion matrix (C) cho 2 lớp (+/-)

Ma trận contingency

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

TP: true positive

tổng số phần tử lớp dương được giải thuật dự đoán lớp dương

TN: true negative

tổng số phần tử lớp âm được giải thuật dự đoán là lớp âm

FP: false positive

tổng số phần tử lớp âm được giải thuật dự đoán là lớp dương

FN: false negative

tổng số phần tử lớp dương được dự đoán là lớp âm

Confusion matrix (C) cho 2 lớp (+/-)

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

Precision

Recall

Accuracy

F1

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

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

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

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

dự đoán =>	dương	âm
dương	10 (TP)	5 (FN)
âm	8 (FP)	22 (TN)

Confusion matrix (C) cho 2 lớp (+/-)

dự đoán =>	dương	âm
dương	10 (TP)	5 (FN)
âm	8 (FP)	22 (TN)

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

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

$$\text{Precision} = 10/(10+8) = 0.56$$

$$\text{Recall} = 10/(10+5) = 0.67$$

$$\begin{aligned}\text{Accuracy} &= (10+22)/10+5+8+22 \\ &= 32/45 = 0.71\end{aligned}$$

$$\begin{aligned}\text{F1} &= 2 \times \text{precision} \times \text{recall} / \\ &(\text{prec} + \text{recall}) = 0.75/1.23 = 0.61\end{aligned}$$

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

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

Chỉ số đánh giá cho bài toán hồi quy

Đánh giá độ chính xác của các dự đoán

- Các chỉ số thường dùng: MSE – Mean Square Error, RMSE – Root Mean Square Error, MAE – Mean Absolute Error
- Đo lường mức độ sai số của các dự đoán. Các giá trị đo lường này bằng 0 khi hệ thống đạt được hiệu quả tốt nhất. Giá trị này càng cao thì hiệu quả của hệ thống càng thấp.
- MAE là chỉ số được sử dụng nhiều nhất vì khả năng giải thích trực tiếp của nó.

Chỉ số đánh giá cho bài toán hồi quy

Đánh giá độ chính xác của các dự đoán

- Tính chính xác của dự đoán được đo trên n quan sát trong đó p_i là giá trị dự đoán đánh giá của item i ,
 r_i là giá trị đánh giá thực tế của item i
- Mean Absolute Error (MAE) (sai số trung bình tuyệt đối) tính toán độ lệch giữa dự đoán xếp hạng và xếp hạng thực tế

$$MAE = \frac{1}{n} \sum_{i=1}^n |p_i - r_i|$$

Chỉ số đánh giá cho bài toán hồi quy

■ Đánh giá độ chính xác của các dự đoán

- Mean Square Error (MSE) (sai số bình phương trung bình).....

$$MSE = \frac{1}{n} \sum_{i=1}^n (p_i - r_i)^2$$

- Root Mean Square Error ($RMSE$) (sai số trung bình toàn phương) tương tự như MAE nhưng chú trọng tới những giá trị có độ lệch lớn

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (p_i - r_i)^2}$$

Ôn tập

Thông tin về các loại nấm được cho trong bảng dữ liệu sau, biết rằng các loại nấm từ A đến H là có thể ăn được hoặc không, hãy xác định

Example	<i>NotHeavy</i>	<i>Smelly</i>	<i>Spotted</i>	<i>Smooth</i>	<i>Edible</i>
<i>A</i>	1	0	0	0	1
<i>B</i>	1	0	1	0	1
<i>C</i>	0	1	0	1	1
<i>D</i>	0	0	0	1	0
<i>E</i>	1	1	1	0	0
<i>F</i>	1	0	1	1	0
<i>G</i>	1	0	0	1	0
<i>H</i>	0	1	0	0	0
<i>U</i>	0	1	1	1	?
<i>V</i>	1	1	0	1	?
<i>W</i>	1	1	0	0	?

1. Entropy của Edible
2. Thuộc tính nào sẽ được chọn làm gốc cho cây quyết định
3. Xây dựng cây qđ ID3 để phân loại U, V, W