

KHOA CNTT & TRUYỀN THÔNG  
BM KHOA HỌC MÁY TÍNH

---

# Đánh giá hiệu quả của giải thuật học

PGS. TS. Đỗ Thanh Nghị  
TS. Trần Nguyễn Minh Thư  
tnmthu@ctu.edu.vn

1

1

## Nội dung

---

- Nghi thức kiểm tra
- Các chỉ số đánh giá

2

2

## 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?

3

3

## 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

4

4

## 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  $k = \text{số phần tử}$ )  
 $\Rightarrow$  Vd leave 1 out

5

5

## ➤ Chỉ số đánh giá

---

6

6