

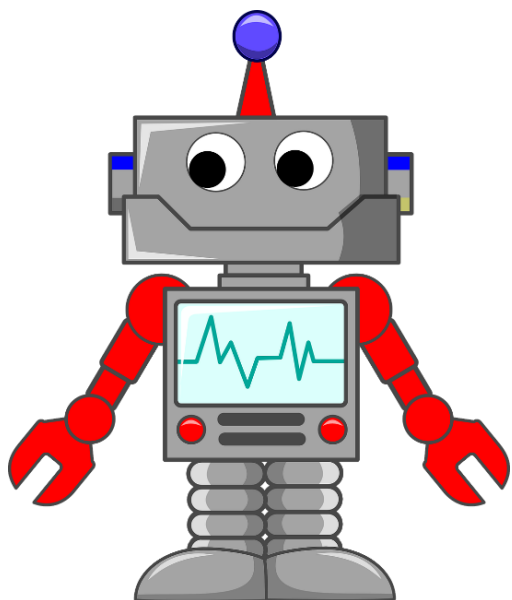


CS116 – LẬP TRÌNH PYTHON CHO MÁY HỌC

BÀI 09

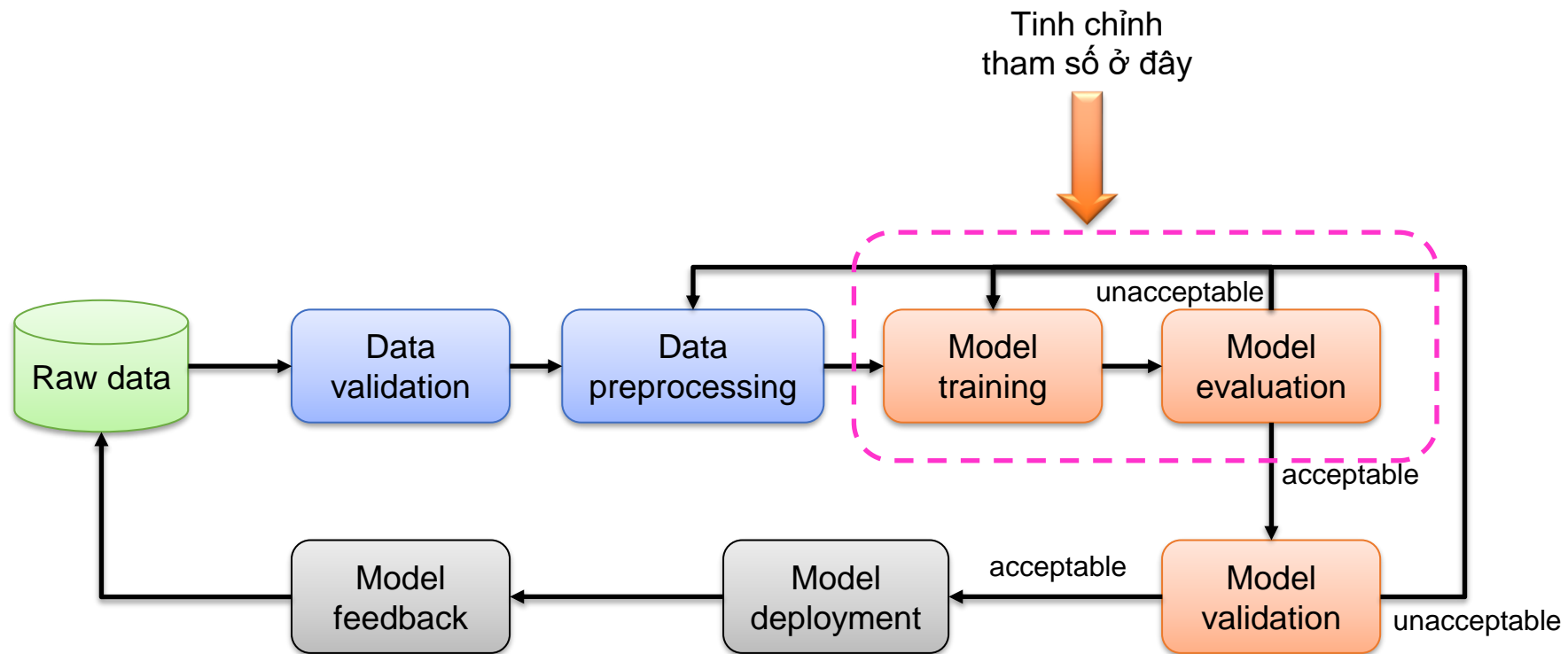
TINH CHỈNH THAM SỐ - PARAMETER TUNING

TS. Nguyễn Vinh Tiệp





Vị trí của bài hôm nay





NỘI DUNG

1. TẠI SAO CẦN PHẢI TÍNH CHỈNH THAM SỐ

2. PHƯƠNG PHÁP GRID SEARCH vét cạn

3. PHƯƠNG PHÁP RANDOM SEARCH ngẫu nhiên

4. PHƯƠNG PHÁP BAYESIAN OPTIMIZATION
tối ưu hóa bayes



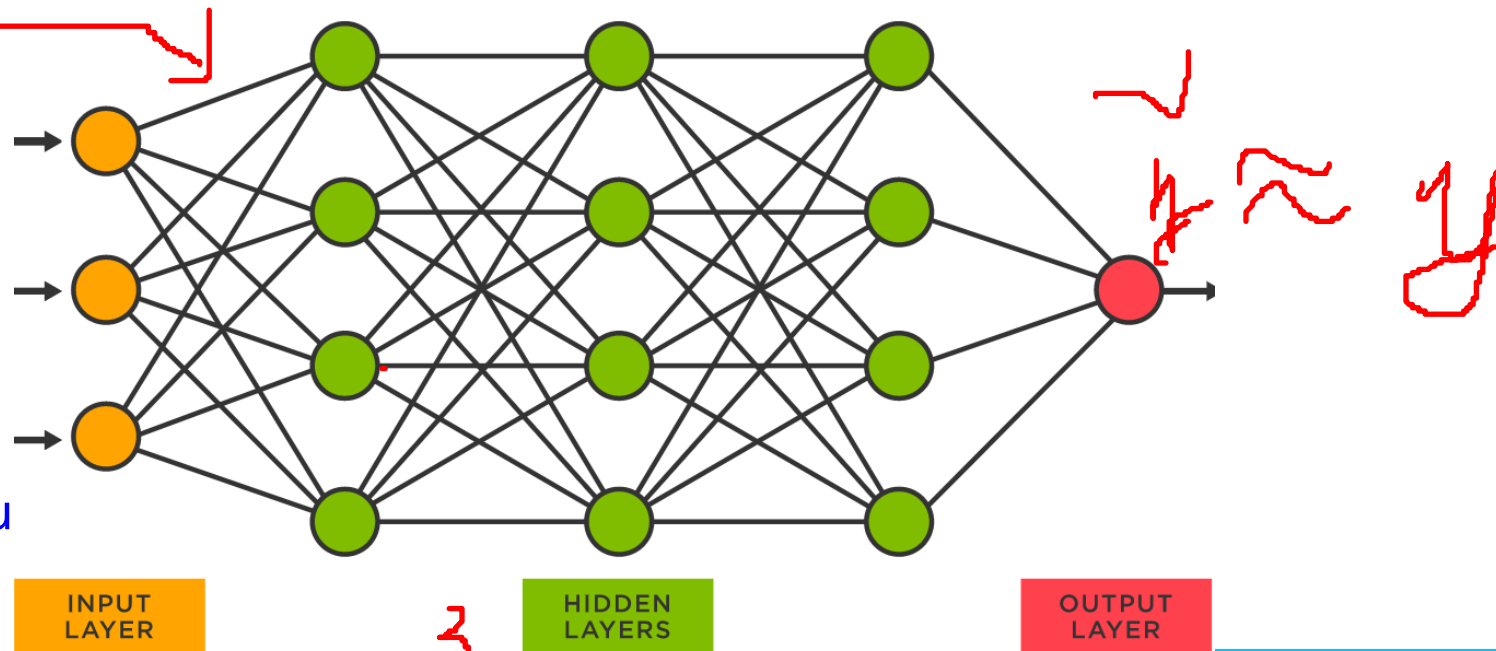
Tham số trong máy học

- **Tham số mô hình (parameter):** là các biến số mô hình học từ dữ liệu train
 - Ví dụ: trọng số, bias của một Neural Network
- **Siêu tham số (hyper-parameter):** là các cấu hình của mô hình trước khi train
 - Được thiết lập thủ công hoặc thông qua quá trình tinh chỉnh siêu tham số
 - Ví dụ: số lớp, số neuron của một lớp trong Neural Network

KNNClassifier với k
là siêu tham số

Logistic Reg thì learning
rate là siêu tham số

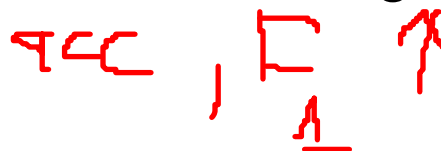
NaiveBayesian classifier
thì không có tham số và siêu
tham số





Tại sao cần phải tinh chỉnh tham số

- **Hiệu suất mô hình:** tinh chỉnh siêu tham số giúp mô hình đạt được độ chính xác cao hơn, hiệu quả hơn



- **Overfitting và underfitting:** việc chọn siêu tham số không phù hợp có thể dẫn đến overfitting hoặc underfitting
dataset có dạng hàm phức tạp mà siêu tham số cho ít thì underfitting
dữ liệu đơn giản: ví dụ tuyến tính, bậc 1 hay bậc 2 mà cho quá nhiều siêu tham số khiến phức tạp hơn, layer quá nhiều hay neuron quá lớn => overfitting

- **Tài nguyên tính toán:** siêu tham số ảnh hưởng đến thời gian và tài nguyên tính toán cần thiết để huấn luyện mô hình

- **Thích ứng với dữ liệu:** tinh chỉnh siêu tham số giúp mô hình được điều chỉnh để phù hợp nhất với đặc trưng riêng của từng dữ liệu



NỘI DUNG

1. TẠI SAO CẦN PHẢI TÍNH CHỈNH THAM SỐ

2. PHƯƠNG PHÁP GRID SEARCH

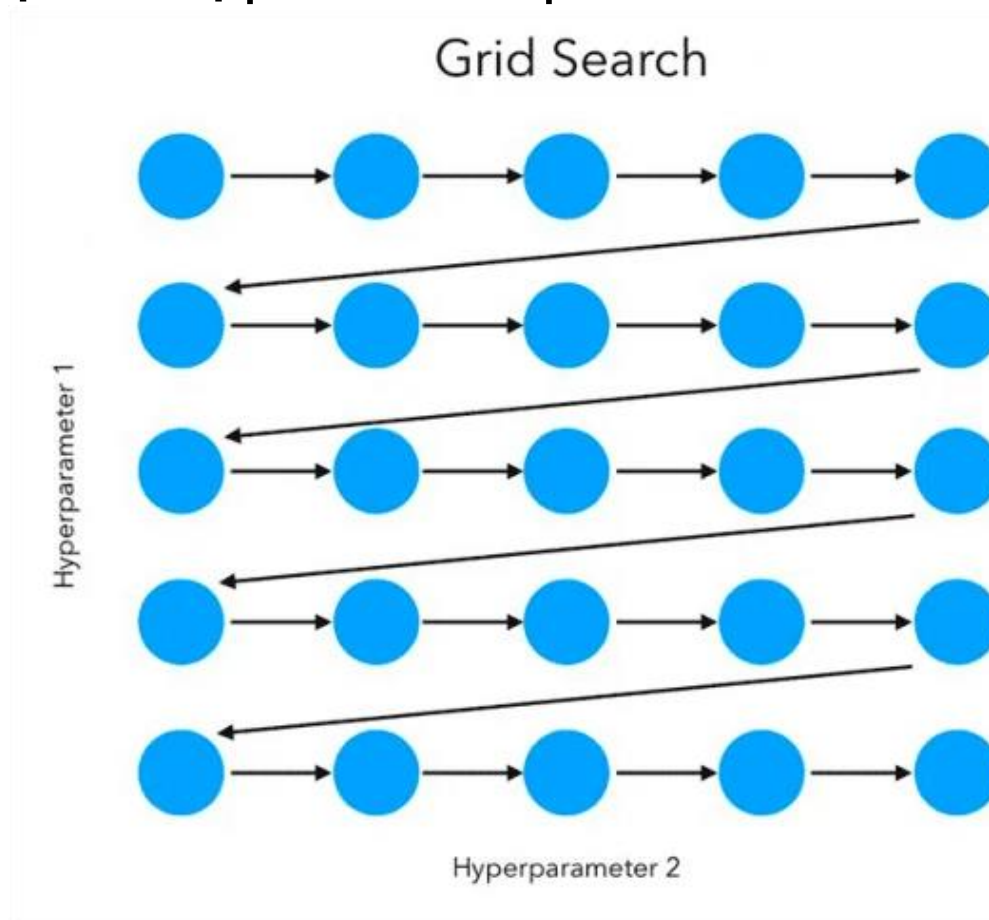
3. PHƯƠNG PHÁP RANDOM SEARCH

4. PHƯƠNG PHÁP BAYESIAN OPTIMIZATION



Phương pháp Grid Search

- **Ý tưởng:** “vét cạn” các tổ hợp tham số có khả năng xảy ra, sau đó tiến hành thử và chọn tổ hợp cho kết quả tốt nhất

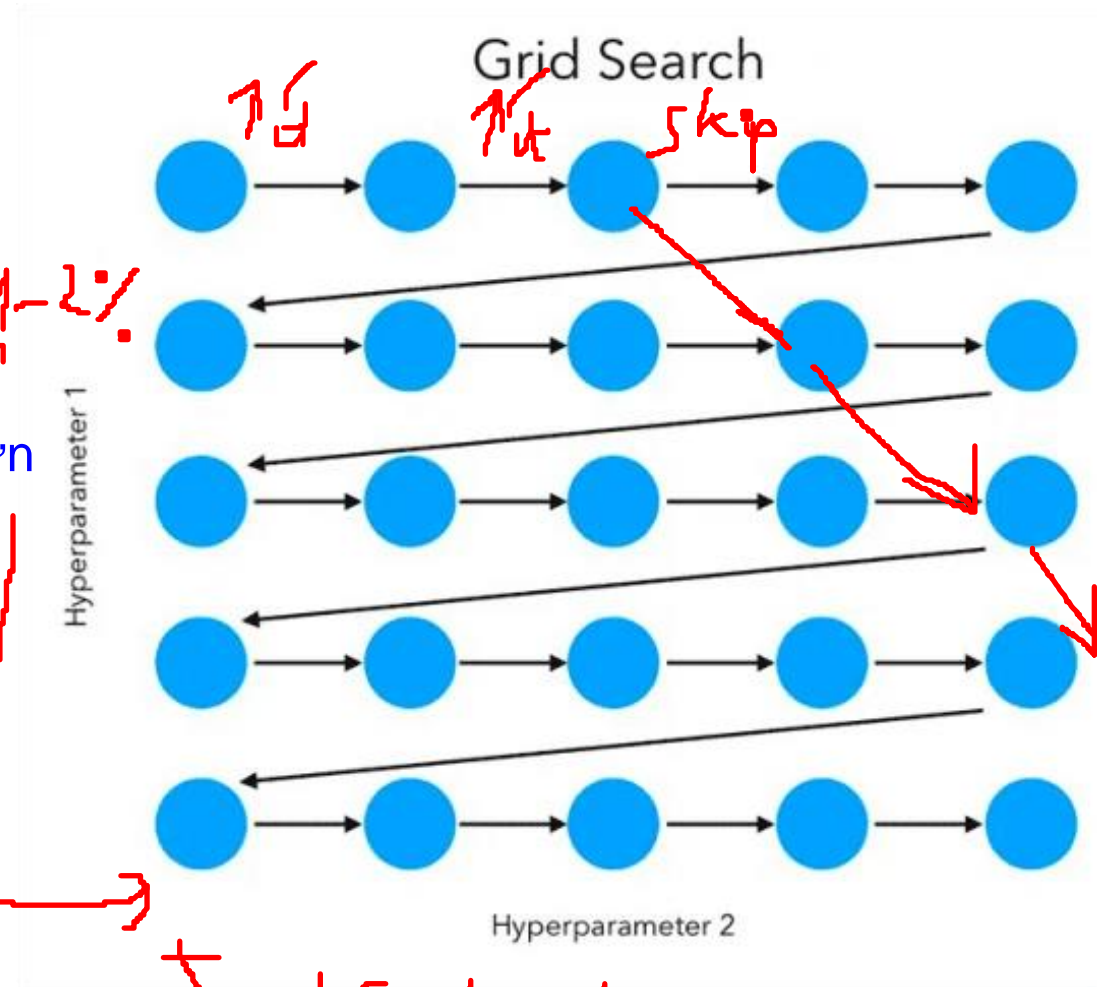


ta có 2 siêu tham số và thử từng tham số dãy từ thấp nhất đến lớn nhất (nếu có vô cùng thì chỉ lấy chia khoảng nhỏ hơn), ứng với mỗi node cấu hình thì xây dựng mô hình -> train -> đánh giá trên tập valid => so sánh tất cả cấu hình và lấy cái tối ưu nhất rồi train lại theo tham số đó tiếp tục train trên dataset lớn hơn



Phương pháp Grid Search

- **Ưu điểm:**
 - Thuật toán đơn giản
 - Tìm kiếm toàn diện không gian tham số
- **Khuyết điểm:** có siêu tham số quan trọng hơn
 - Đánh đồng vai trò của các siêu tham số
 - Không gian tham số lớn khiến chi phí tính toán lớn
 - Không hiệu quả do không kế thừa kết quả tìm kiếm trước đó



Nên tập trung tần suất lấy mẫu trên
siêu tham số quan trọng thôi



NỘI DUNG

1. TẠI SAO CẦN PHẢI TÍNH CHỈNH THAM SỐ

2. PHƯƠNG PHÁP GRID SEARCH

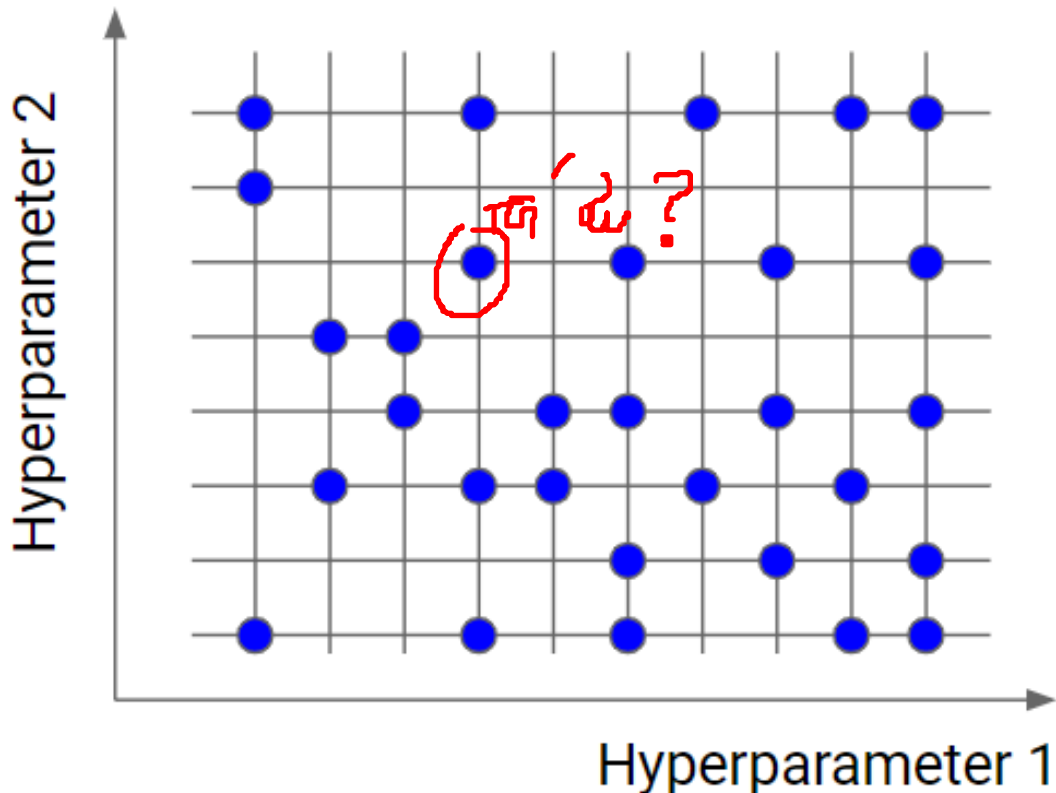
3. PHƯƠNG PHÁP RANDOM SEARCH

4. PHƯƠNG PHÁP BAYESIAN OPTIMIZATION

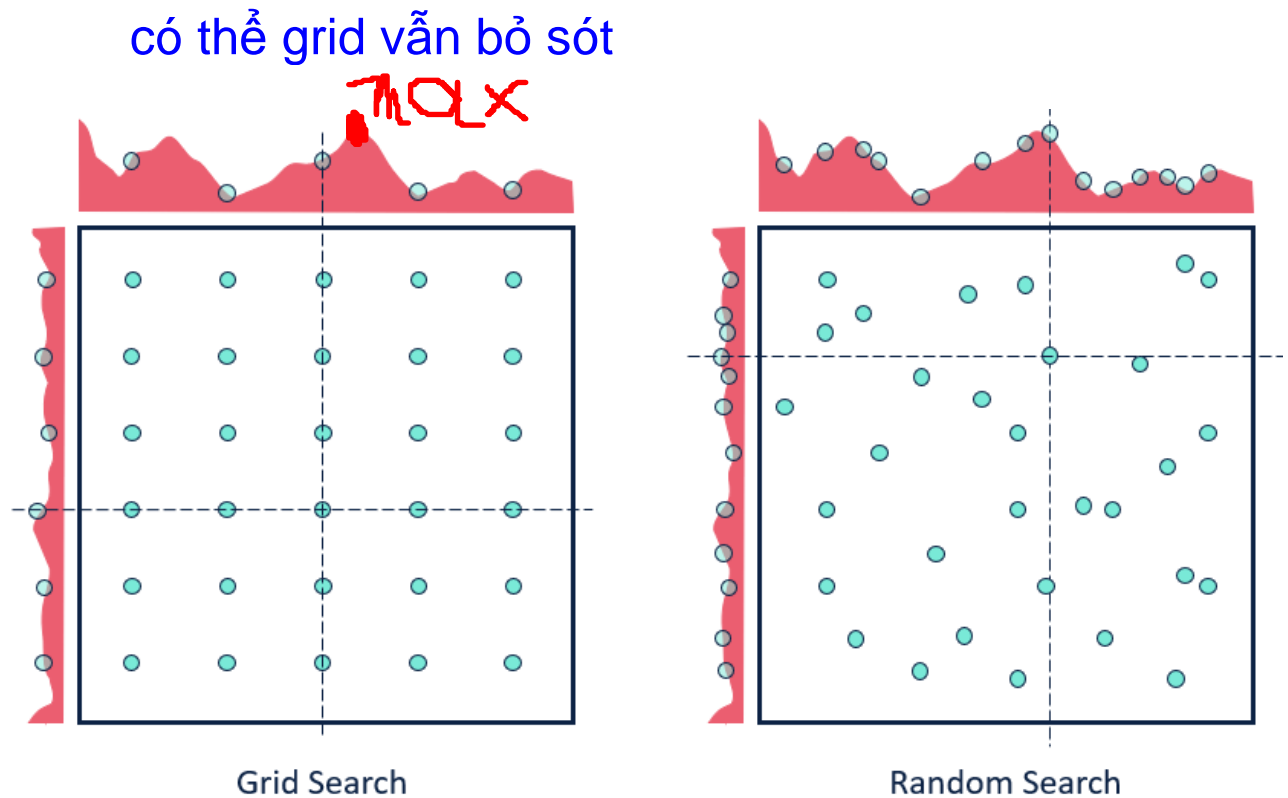


Phương pháp Random Search

- Ý tưởng: chọn ngẫu nhiên các tổ hợp hyperparameter, thay vì lấy mẫu đều, sau đó thử lần lượt và chọn tổ hợp cho mô hình tốt nhất



Phương pháp Random search

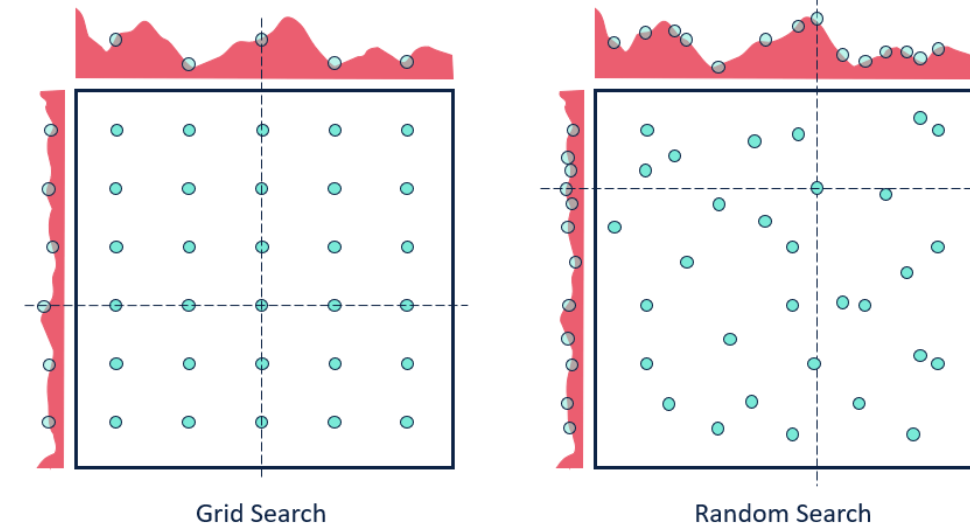


So sánh Grid Search với Random Search



Phương pháp Random Search

- **Ưu điểm:**
 - Đơn giản, dễ cài đặt,
 - Hiệu quả với không gian tìm kiếm lớn
- **Khuyết điểm:**
 - Nếu không gian tìm kiếm lớn, khó tìm được điểm tối ưu toàn cục
 - Các lần thử độc lập, không kế thừa thông tin các lần thử trước
 - Kết quả không nhất quán do yếu tố ngẫu nhiên



So sánh Grid Search với Random Search



NỘI DUNG

1. TẠI SAO CẦN PHẢI TÍNH CHỈNH THAM SỐ

2. PHƯƠNG PHÁP GRID SEARCH

3. PHƯƠNG PHÁP RANDOM SEARCH

4. PHƯƠNG PHÁP BAYESIAN OPTIMIZATION



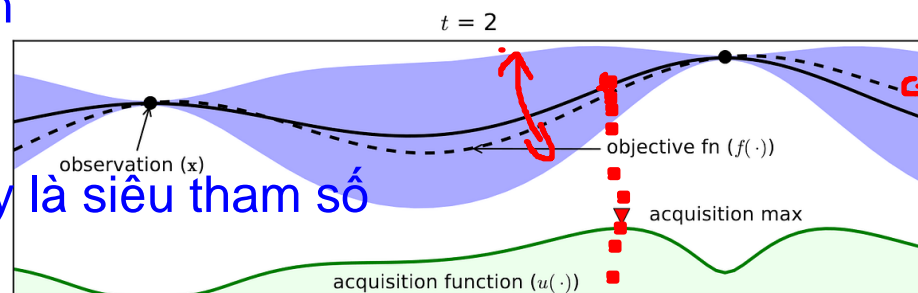
Phương pháp Bayesian Optimization

có tính kế thừa

- Ý tưởng: chiến lược tuần tự để tìm điểm tối ưu toàn cục của các mô hình dạng black-box cho input trả ra output không quan tâm bên trong có gì

exploit: khai phá, khai thác thông tin của lần trước đó
exploration: khám phá bước nhảy ngẫu nhiên

quan sát này là siêu tham số

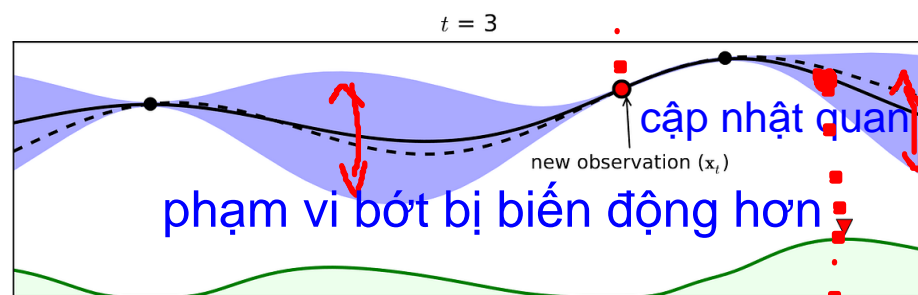


đường --- là hàm mục tiêu

thực tế ước lượng

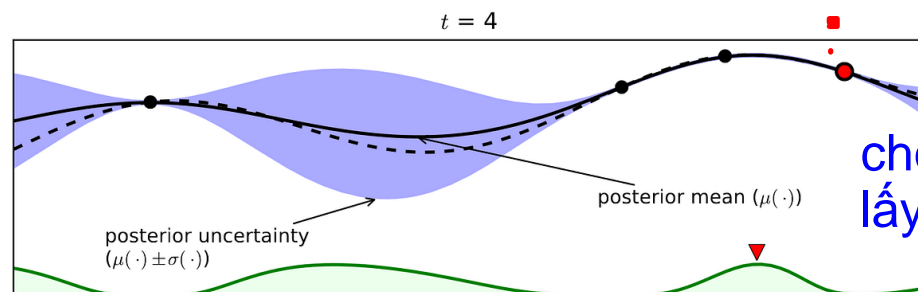
muốn cho ước lượng về sát thực tế

hàm hữu dụng: exploit và explore



cập nhật quan sát mới và cập nhật hàm ước lượng mới

phạm vi bớt bị biến động hơn



cho biết nên thử chỗ nào thì có thể lấy ra siêu tham số tốt



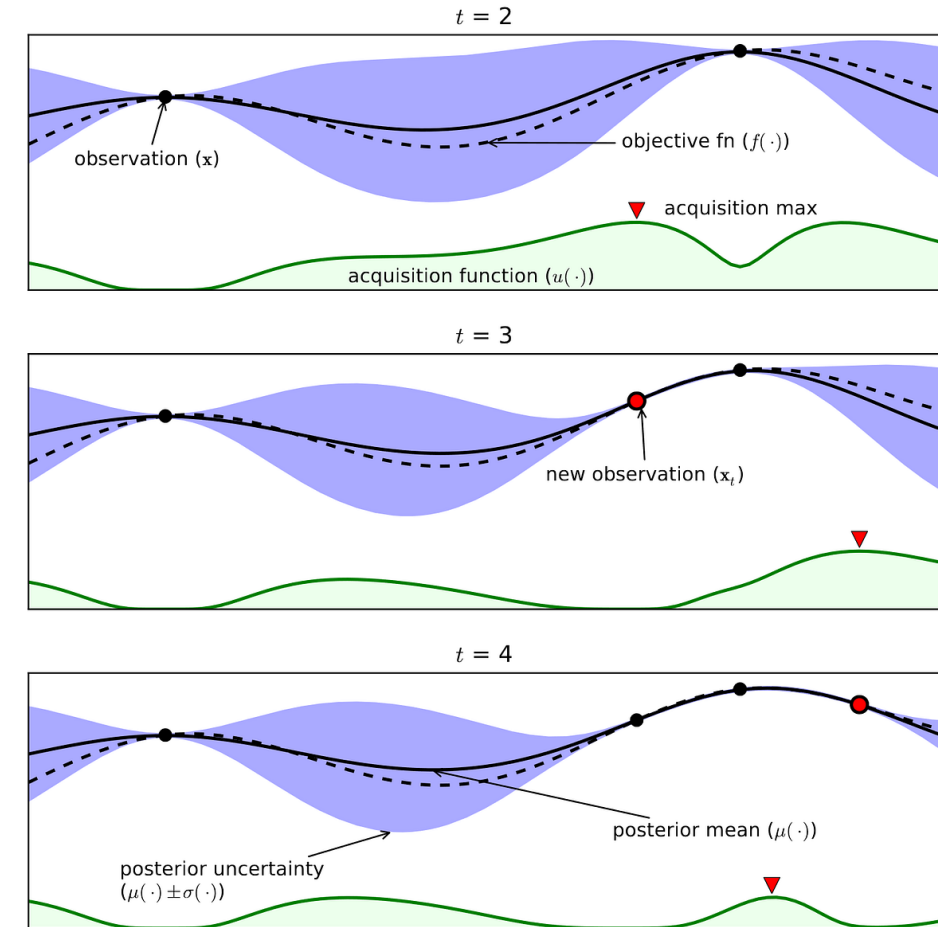
Phương pháp Bayesian Optimization

- **Ưu điểm:**

- Hiệu quả, có kế thừa thông tin các lần thử trước đó để tìm hiệu quả hơn
- Giảm bớt số lần thử nghiệm
- Phù hợp với không gian tham số lớn

- **Khuyết điểm:**

- Phương pháp phức tạp hơn so với RS, GS
- Việc chọn hàm acquisition ảnh hưởng lớn đến hiệu quả của phương pháp





So sánh các phương pháp

	PP Grid Search	PP Random Search	PP Bayesian Optimization
Ưu điểm	<ul style="list-style-type: none">- Đơn giản, dễ cài- Tìm kiếm toàn diện	<ul style="list-style-type: none">- Đơn giản, dễ cài- Hiệu quả với không gian tham số lớn	<ul style="list-style-type: none">- Hiệu quả do kế thừa được các lần thử trước đó- Hiệu quả với không gian tham số lớn
Khuyết điểm	<ul style="list-style-type: none">- Không gian tìm kiếm quá lớn → chi phí tính toán lớn- Không hiệu quả do không kế thừa lần thử trước để cải thiện tìm kiếm	<ul style="list-style-type: none">- Không tìm kiếm toàn diện- Không hiệu quả do không kế thừa lần thử trước để cải thiện tìm kiếm	<ul style="list-style-type: none">- Sử dụng lý thuyết xác suất, phức tạp hơn- Việc chọn hàm acquisition ảnh hưởng lớn đến hiệu quả của phương pháp



BÀI QUIZ VÀ HỎI ĐÁP