

Bias variance tradeoff Overfitting & Underfitting

Mentor: PĐ Hải

Người trình bày: Nguyễn Minh Hiếu

PAYT CLUB PTIT

September 2025

- 1 Overfitting & underfitting
 - 1.1: Overfitting & underfitting
 - 1.2: Một số phương pháp

- 2 Bias-variance tradeoff (Sự đánh đổi giữa độ chệch và phương sai)
 - 2.1: Tìm hiểu về Bias
 - 2.2: Tìm hiểu về Variance
 - 2.3: Tìm hiểu về Bias variance tradeoff

1: Overfitting & Underfitting

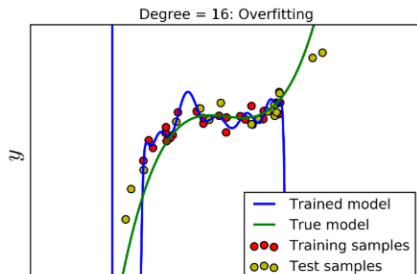
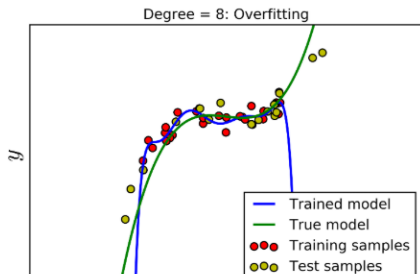
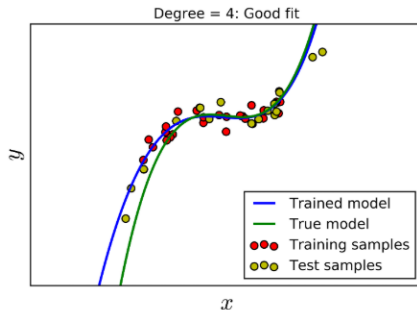
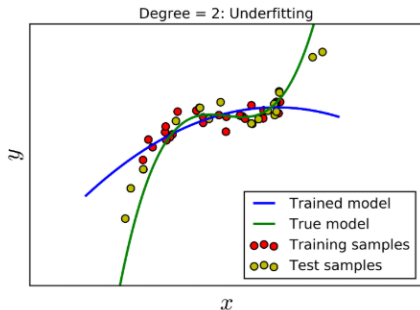
1.1: Overfitting & underfitting - Hiện tượng overfitting

- **Định nghĩa:** Mô hình học "**quá tốt**" từ dữ liệu huấn luyện, bao gồm cả nhiễu và ngoại lệ, dẫn đến khả năng dự đoán kém trên dữ liệu mới.
- **Nguyên nhân:**
 - ▶ Mô hình quá phức tạp.
 - ▶ Dữ liệu huấn luyện không đủ.

Hiện tượng underfitting

- **Định nghĩa:** Mô hình quá đơn giản, không nắm bắt được xu hướng cơ bản của dữ liệu.
- **Nguyên nhân:**
 - ▶ Mô hình quá đơn giản.
 - ▶ Dữ liệu huấn luyện không đủ.
 - ▶ Đặc trưng (features) không đủ tốt.
 - ▶ Dữ liệu chưa được chuẩn hóa (scaling).

Minh họa



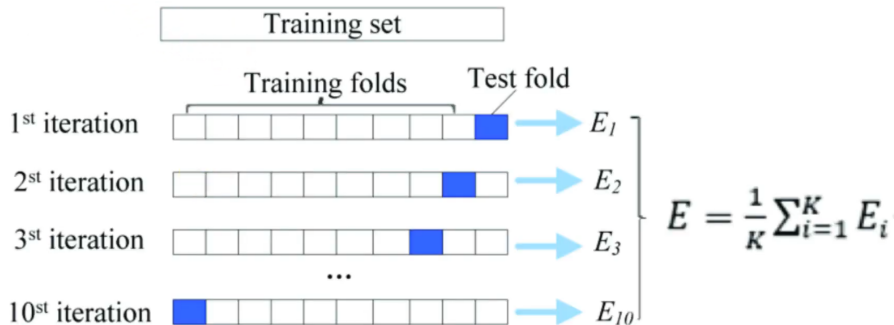
1.2: Một số phương pháp - validation

- Là trích từ tập training data ra một tập con nhỏ (validation set) và thực hiện việc đánh giá mô hình trên tập con nhỏ này.
 - Nếu accuracy train \gg accuracy validation \Rightarrow Overfitting
 - Nếu cả 2 thấp \Rightarrow Underfitting
 - Nhược điểm:
 - ▶ Phụ thuộc vào một lần chia dữ liệu duy nhất.
 - ▶ Việc chia dữ liệu Validation cố định có thể vô tình bỏ qua các mẫu dữ liệu quan trọng nằm trong tập Training.
- \Rightarrow Đánh giá không ổn định và có thể không đại diện cho toàn bộ tập dữ liệu.

Cross - validation (K-fold cross validation)

- Bước 1: Chia ngẫu nhiên toàn bộ dữ liệu thành k phần (folds) bằng nhau.
- Bước 2: Lặp lại k lần:
 - ▶ Lấy 1 fold làm tập Validation.
 - ▶ Lấy $k-1$ folds còn lại gộp lại làm tập Training.
 - ▶ Huấn luyện mô hình và đánh giá trên tập Validation.
 - ▶ Lưu lại kết quả đánh giá.
- Bước 3: Lấy trung bình cộng của k kết quả đánh giá đó.

Cross - validation (K-fold cross validation)



Regularization

- Ý tưởng là thêm vào hàm mất mát 1 số hạng nữa. Số hạng này thường dùng để đánh giá độ phức tạp của mô hình. Số hạng này càng lớn thì mô hình càng phức tạp
- L2 Regularization (Ridge Regression):

$$w^* = \arg \min_w \left(\sum_{i=1}^M (y_i - A_i w)^2 + \lambda \sum_{j=0}^n w_j^2 \right)$$

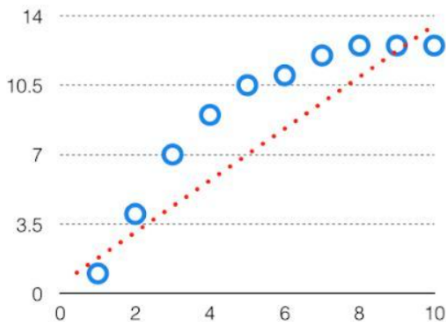
- L1 Regularization (Lasso Regression):

$$w^* = \arg \min_w \left(\sum_{i=1}^M (y_i - A_i w)^2 + \lambda \sum_{j=0}^n |w_j| \right)$$

2: Bias variance tradeoff

2.1: Tìm hiểu về Bias

- Bias (Độ chệch) là lỗi xảy ra do sự **không đủ khả năng** của mô hình, dẫn đến **sự khác biệt có hệ thống** giữa giá trị dự đoán của mô hình và giá trị thực tế.
- Bias cao dẫn tới **lỗi lớn** trên cả training data và testing data (Underfitting).
- Mô hình có bias cao là do mô hình quá đơn giản, ví dụ dự đoán dạng 1 đường thẳng (linear) không thể fit với dữ liệu trong dataset.



2.1: Công thức tính Bias

- **Công thức định lượng độ chệch:**

$$\text{Bias}(\hat{Y}) = E(\hat{Y}) - Y$$

Trong đó:

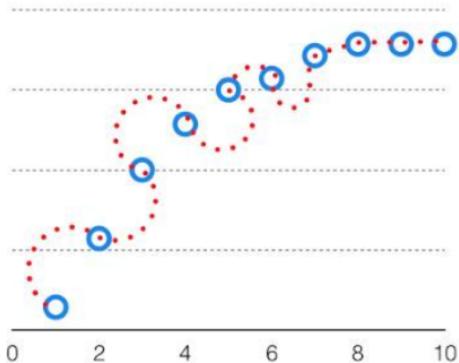
- ▶ Y : Giá trị thực tế.
 - ▶ \hat{Y} : Giá trị do mô hình dự đoán.
 - ▶ $E(\hat{Y})$: Kỳ vọng của các dự đoán trên nhiều tập dữ liệu.
- **Ý nghĩa:** Đo lường mức độ "lệch" có hệ thống của các dự đoán so với giá trị thực.

2.1: Cách giảm high bias

- Sử dụng mô hình phức tạp hơn
- Tăng features, tăng số lượng tập training data
- Điều chỉnh regularization

2.2: Tìm hiểu về Variance

- Phương sai (Variance) là độ biến thiên của dự đoán mô hình cho một điểm dữ liệu.
- Mô hình phương sai cao dẫn tới hiệu suất rất tốt ở training data nhưng lại rất tệ ở testing data (Overfitting)



2.2: Công thức tính Variance

Công thức tính Variance:

$$\text{Variance} = E[(\hat{Y} - E[\hat{Y}])^2]$$

- $E[\hat{Y}]$: Giá trị kỳ vọng (trung bình) của các dự đoán
- \hat{Y} : Giá trị dự đoán của mô hình

2.2: Cách giảm high Variance

Một số cách giảm high variance:

- Cross - validation
- Regularization

2.3: Tìm hiểu về Bias variance tradeoff

Mục tiêu: Tối thiểu hóa **tổng lỗi** của mô hình

$$\text{Total Error} = \text{Bias}^2 + \text{Variance} + \text{Irreducible Error}$$

$$\begin{aligned}\text{MSE} &= (Y - \hat{Y})^2 \\ &= (Y - E(\hat{Y}) + E(\hat{Y}) - \hat{Y})^2 \\ &= (Y - E(\hat{Y}))^2 + (E(\hat{Y}) - \hat{Y})^2 + 2(Y - E(\hat{Y}))(E(\hat{Y}) - \hat{Y})\end{aligned}$$

2.3: Tìm hiểu về Bias variance tradeoff

$$\begin{aligned}E[(Y - \hat{Y})^2] &= E[(Y - E(\hat{Y}))^2 + (E(\hat{Y}) - \hat{Y})^2 + 2(Y - E(\hat{Y}))(E(\hat{Y}) - \hat{Y})] \\&= E[(Y - E(\hat{Y}))^2] + E[(E(\hat{Y}) - \hat{Y})^2] + 2E[(Y - E(\hat{Y}))(E(\hat{Y}) - \hat{Y})] \\&= E[(Y - E(\hat{Y}))^2] + E[(E(\hat{Y}) - \hat{Y})^2] + 2(Y - E(\hat{Y}))E[(E(\hat{Y}) - \hat{Y})] \\&= E[(Y - E(\hat{Y}))^2] + E[(E(\hat{Y}) - \hat{Y})^2] + 2(Y - E(\hat{Y}))[E(E(\hat{Y})) - E[\hat{Y}]] \\&= E[(Y - E(\hat{Y}))^2] + E[(E(\hat{Y}) - \hat{Y})^2] + 2(Y - E(\hat{Y}))[E(\hat{Y})] - E[\hat{Y}] \\&= E[(Y - E(\hat{Y}))^2] + E[(E(\hat{Y}) - \hat{Y})^2] + 2(Y - E(\hat{Y}))[0] \\&= E[(Y - E(\hat{Y}))^2] + E[(E(\hat{Y}) - \hat{Y})^2] + 0 \\&= [\text{Bias}^2] + \text{Variance}\end{aligned}$$

2.3: Biểu đồ tương quan bias - variance tradeoff

