

Overfitting

Khái niệm	Mô tả ngắn gọn	Tên khác
Underfitting	Mô hình quá đơn giản, không học được quy luật dữ liệu	High bias
Overfitting	Mô hình quá phức tạp, học thuộc lòng dữ liệu huấn luyện	High variance
Just right	Mô hình vừa đủ độ phức tạp, tổng quát tốt trên dữ liệu mới	Tốt/Khái quát tốt

- **High bias:** Mô hình có định kiến mạnh, bỏ qua quy luật trong dữ liệu.
- **High variance:** Mô hình quá nhạy cảm với từng điểm dữ liệu huấn luyện.

⇒ Cần **cân bằng bias-variance** để đạt hiệu suất tốt nhất.

3 cách chính để khắc phục overfitting

1. Thu thập thêm dữ liệu huấn luyện

- Là **cách hiệu quả nhất** để giảm overfitting nếu có thể thực hiện.
- Với nhiều dữ liệu hơn, mô hình sẽ học được quy luật tổng quát hơn và bớt bị ảnh hưởng bởi nhiễu.
- Tuy nhiên, trong thực tế, **không phải lúc nào cũng có thêm dữ liệu**, ví dụ như số lượng giao dịch nhà ở trong một khu vực có thể rất hạn chế.

2. Giảm số lượng đặc trưng (features) — *Feature Selection*

- Khi có **quá nhiều đặc trưng mà dữ liệu lại ít**, mô hình có thể bị overfit.
- Có thể chọn ra **những đặc trưng quan trọng nhất** (ví dụ: diện tích, số phòng ngủ, tuổi của ngôi nhà) thay vì dùng tất cả.
- Đây gọi là kỹ thuật **feature selection**.
- **Lưu ý:** Việc loại bỏ đặc trưng cũng đồng nghĩa với việc loại bỏ thông tin – có thể mất đi một phần thông tin hữu ích.

3. Regularization – Chính quy hóa tham số

- Không loại bỏ đặc trưng, mà **làm giảm độ lớn của các trọng số (w)**, giúp giảm ảnh hưởng của chúng.
- **Ý tưởng**: nếu một hệ số quá lớn thì có thể gây nên sự uốn lượn quá mức của hàm dự đoán.
- Regularization khuyến khích các trọng số **nhỏ lại gần 0**, nhưng không ép buộc bằng cách đặt bằng 0 như loại bỏ đặc trưng.
- **Thông thường chỉ regularize các trọng số w_1, w_2, \dots, w_n** , không cần thiết phải regularize b (bias).

Regularization

Mục đích:

- Làm **giảm độ lớn các trọng số w_1, w_2, \dots, w_n** nhằm tránh overfitting.
- **Trọng số nhỏ hơn \Rightarrow mô hình đơn giản hơn \Rightarrow ít overfit hơn.**

Áp dụng: Thêm vào hàm loss một giá trị **term regularization**:

$$J(w, b) = \frac{1}{2m} \sum_{i=1}^m (\hat{y}^{(i)} - y^{(i)})^2 + \frac{\lambda}{2m} \sum_{j=1}^n w_j^2$$

- λ : hệ số regularization — quyết định mức độ ưu tiên giữa:
 - **Fit tốt dữ liệu** (phần đầu).
 - **Giữ trọng số nhỏ** (phần sau).
- Thường **không regularize b** (bias) vì ảnh hưởng nhỏ.

Giải thích:

- Với tham số regularization, trong **mỗi vòng lặp gradient descent, trọng số w_j bị nhân với số nhỏ hơn 1** \Rightarrow khiến trọng số **"thu nhỏ dần"** \Rightarrow chống overfitting.