

# 1. Model không chính xác (Inaccurate Model) là gì?

- **Model không chính xác** xảy ra khi các chuyển tiếp (transition) mà model dự đoán khác với những gì thực sự xảy ra trong môi trường.
  - Có 2 dạng model không chính xác phổ biến:
    - **Model không đầy đủ (Incomplete model):**  
Ở đầu quá trình học, agent chưa thử nhiều hành động ở nhiều trạng thái, nên model chưa biết transition ở các cặp (state, action) đó → thiếu dữ liệu, model chưa có thông tin về nhiều chuyển tiếp.
    - **Model lỗi thời (Outdated model):**  
Môi trường thay đổi, hành động trong một trạng thái không còn dẫn đến kết quả như trước nữa. Model vẫn dự đoán theo dữ liệu cũ, nên sai lệch so với thực tế.
- 

# 2. Ảnh hưởng của việc lập kế hoạch với model không chính xác

- Khi model **không đầy đủ (incomplete)**:
    - Vì thiếu các chuyển tiếp, agent không thể mô phỏng các bước đi chưa từng gặp → không thể dùng để planning cho các cặp trạng thái-hành động chưa biết.
    - Nhưng khi agent càng tương tác, model càng được hoàn thiện, và có thể lập kế hoạch dựa trên các transition đã biết.
  - Khi model **bị lỗi thời (environment changes)**:
    - Nếu agent dùng các transition sai lệch trong model để planning, việc cập nhật giá trị Q hay chính sách sẽ dựa trên thông tin sai → có thể làm giảm hiệu quả chính sách trong môi trường thực.
    - Planning lúc này cải thiện chính sách theo model, không phải theo môi trường thật, dẫn đến chính sách kém đi.
-

### 3. Làm sao Dyna-Q vẫn lập kế hoạch thành công với model không đầy đủ?

- Dyna-Q chỉ lập kế hoạch trên các cặp (state, action) mà agent **đã từng trải nghiệm** (đã từng thấy transition).
- Khi mới bắt đầu học, có thể Dyna-Q sẽ lặp lại nhiều lần các transition đã biết, vì chưa có nhiều dữ liệu khác.
- Dần dần, khi agent trải nghiệm nhiều hơn, kế hoạch được thực hiện trên nhiều trạng thái và hành động hơn, giúp cải thiện chính sách toàn diện.
- Do vậy, Dyna-Q có thể **lập kế hoạch hiệu quả ngay cả khi model chưa đầy đủ**, vì nó chỉ chọn các transition đã biết để mô phỏng.