1. Mô hình Simple Regressor:

**Chỉ số MAE trên tập huấn luyện (Train MAE = 0.038)**:

* Đây là một giá trị rất thấp, cho thấy mô hình hoạt động tốt trên dữ liệu huấn luyện và có thể dự đoán khá chính xác các giá trị mục tiêu trong tập này.

**Chỉ số MAE trên tập kiểm tra (Test MAE = 2.565)**:

* Giá trị này lớn hơn rất nhiều so với tập huấn luyện, cho thấy mô hình có thể đang **overfitting** dữ liệu huấn luyện. Điều này nghĩa là mô hình học quá chi tiết từ dữ liệu huấn luyện, dẫn đến khả năng khái quát kém trên dữ liệu mới (tập kiểm tra).
* **Hiện tượng overfitting**: Khoảng cách giữa MAE của tập huấn luyện và tập kiểm tra rất lớn, chỉ ra rằng mô hình không thể tổng quát hóa tốt. Điều này thường xảy ra khi mô hình quá đơn giản (underfitting) hoặc không được tinh chỉnh đúng cách.
* **Phân tích biểu đồ**:
  + **Dưới 100**: Trong khoảng giá trị nhỏ (dưới 100), các điểm gần như khớp với đường màu đỏ. Điều này cho thấy mô hình dự đoán tốt trong vùng giá trị thấp.
  + **Lớn hơn 100**: Khi giá trị thực tăng, nhiều điểm bắt đầu lệch xa đường tham chiếu, đặc biệt với các giá trị lớn hơn 150. Điều này cho thấy mô hình có xu hướng kém chính xác khi dự đoán các giá trị lớn.
  + Có hiện tượng **underestimation** (dự đoán thấp hơn giá trị thực) khi giá trị thực tăng lên, điều này phản ánh khả năng mô hình không đủ linh hoạt để bắt kịp xu hướng của dữ liệu ở vùng giá trị cao.
  + **Hiệu quả mô hình**:
    - Mô hình hoạt động khá tốt với các giá trị nhỏ và trung bình.
    - Tuy nhiên, nó gặp khó khăn trong việc dự đoán các giá trị lớn, dẫn đến sai số tăng cao (giải thích cho MAE lớn trên tập kiểm tra).

1. Mô hình Grid Search

**learning\_rate = 0.07**: Tốc độ học nhỏ, cho thấy mô hình học từng bước nhỏ để tối ưu hóa, giúp tránh overfitting.

**max\_depth = 7**: Chiều sâu tối đa của cây quyết định trong mô hình. Với giá trị 7, mô hình có khả năng học được các mối quan hệ phức tạp từ dữ liệu nhưng vẫn không quá phức tạp để gây overfitting.

**n\_estimators = 500**: Số lượng cây trong mô hình ensemble, cho thấy mô hình sử dụng khá nhiều cây để tăng khả năng dự đoán.

**Best score = -7.9249**: Đây là giá trị của hàm mục tiêu (negative mean squared error) được tối ưu trong Grid Search. Giá trị âm này xuất phát từ việc GridSearchCV sử dụng hàm mất mát để đánh giá, nhưng đảo dấu để chuyển từ hàm lỗi sang hàm đánh giá.

 Giá trị tương ứng của **MSE (Mean Squared Error)** là khoảng **7.92**, có thể chuyển đổi thành **RMSE = √7.92 ≈ 2.81**.

**Ý nghĩa**: RMSE là khoảng cách trung bình giữa dự đoán và giá trị thực tế. Giá trị này cho thấy mô hình dự đoán sai số trung bình khoảng 2.81 (phụ thuộc vào đơn vị của biến mục tiêu).

* **Tích cực**:
  + **Hàm mục tiêu**: Giá trị Best Score cho thấy mô hình tìm được tổ hợp tham số tốt nhất dựa trên Grid Search.
  + **Kiểm soát overfitting**: Với learning\_rate thấp và max\_depth không quá lớn, mô hình có khả năng khái quát hóa tốt hơn.
* **Hạn chế**:
  + Giá trị của **Best Score (-7.92)** chỉ đánh giá trên tập huấn luyện với cross-validation. Hiệu suất thực sự cần được kiểm tra trên tập kiểm tra để đánh giá tính tổng quát.
  + Nếu giá trị này được xem xét là cao (RMSE = 2.81 không đủ tốt), thì có thể cần tối ưu hóa thêm hoặc thử mô hình khác.
* Đánh giá biểu đồ của mô hình:
  + **Dự đoán chính xác ở mức thấp:**
    - Các điểm gần gốc tọa độ (giá trị nhỏ hơn 100) nằm khá sát với đường đỏ, cho thấy mô hình dự đoán tốt trong khoảng giá trị thấp.
  + **Sai số tăng ở giá trị lớn:**
    - Khi giá trị thực lớn hơn 150, các điểm bắt đầu lệch khỏi đường tham chiếu. Một số dự đoán quá thấp so với giá trị thực (underestimation). Điều này cho thấy mô hình không dự đoán chính xác các giá trị lớn, có thể do:
    - Tập dữ liệu bị lệch (giá trị lớn ít hơn).
    - Mô hình chưa đủ phức tạp để học được xu hướng ở vùng giá trị lớn.
  + **Hiệu suất tổng quát:**
    - Biểu đồ cho thấy mô hình hoạt động tốt ở các giá trị thấp và trung bình, nhưng gặp khó khăn với các giá trị cao hơn. Đây là dấu hiệu của việc mô hình có thể bị giới hạn bởi cấu trúc hoặc tham số được lựa chọn trong quá trình Grid Search.
* **Thành công:** Các thông số tối ưu từ Grid Search (learning\_rate = 0.07, max\_depth = 7, n\_estimators = 500) giúp mô hình hoạt động khá tốt trong vùng giá trị phổ biến (thấp và trung bình).
* **Hạn chế:**
  + Mô hình vẫn chưa xử lý tốt các giá trị bất thường hoặc vùng dữ liệu không đồng nhất (giá trị lớn).
  + Sai số này có thể phản ánh rằng mô hình cần được cải thiện thêm, đặc biệt khi dữ liệu có phân phối không đồng đều.

1. Mô hình Stacked Regressor

* **Ý nghĩa của chỉ số MAE (Mean Absolute Error):**
  + MAE đo lường sai số trung bình tuyệt đối giữa giá trị dự đoán và giá trị thực tế.
  + Với **MAE = 2.223** trên tập kiểm tra, mô hình dự đoán trung bình sai lệch khoảng 2.223 đơn vị so với giá trị thực tế. Điều này phản ánh hiệu suất khá tốt, đặc biệt nếu đơn vị biến mục tiêu không quá lớn.
* **So sánh với các mô hình trước:**
* **Simple Regressor:** MAE trên tập kiểm tra là **2.565**, lớn hơn so với Stacked Regressor. Điều này cho thấy Stacked Regressor hoạt động hiệu quả hơn, có thể nhờ khả năng kết hợp các mô hình con và tận dụng được sức mạnh tổng hợp.
* **Grid Search (RMSE = 2.81):** MAE của Stacked Regressor thấp hơn RMSE của mô hình Grid Search, cho thấy Stacked Regressor là lựa chọn tốt hơn trong trường hợp này.
* Hiệu quả của mô hình:
  + Stacked Regressor kết hợp nhiều mô hình (base models) khác nhau, giúp cải thiện độ chính xác và khả năng khái quát hóa trên dữ liệu kiểm tra. Điều này thường giúp giảm sai số trong các tình huống dữ liệu phức tạp hoặc phi tuyến.
* **Điểm mạnh:**
  + MAE = 2.223 trên tập kiểm tra là một kết quả tốt, cho thấy mô hình Stacked Regressor dự đoán chính xác hơn so với các mô hình trước đó.
  + Việc giảm sai số có thể do mô hình tận dụng được các đặc điểm tốt nhất của từng base model.
* **Hạn chế:**
  + Sai số 2.223 vẫn còn phụ thuộc vào đơn vị của biến mục tiêu. Nếu giá trị biến mục tiêu có độ lớn lớn hơn nhiều (ví dụ: từ 0-300), đây là sai số tương đối nhỏ. Nhưng nếu giá trị biến mục tiêu nhỏ hơn, đây vẫn là sai số cần cải thiện thêm.
  + Độ phức tạp của Stacked Regressor cũng có thể gây ra việc khó giải thích hơn so với các mô hình đơn giản.

1. Mô hình Neural Network với 3 hidden layers

**1. Ý nghĩa của chỉ số MAE:**

* **MAE trên tập huấn luyện (train): 1.948**
  + Mô hình dự đoán sai số trung bình khoảng **1.948 đơn vị** trên tập huấn luyện.
  + Đây là mức sai số chấp nhận được, cho thấy mô hình đã học tốt trên dữ liệu huấn luyện.
* **MAE trên tập kiểm tra (test): 1.602**
  + Mô hình có sai số trung bình thấp hơn trên tập kiểm tra so với tập huấn luyện, với mức sai số **1.602**.
  + Điều này phản ánh rằng mô hình có khả năng tổng quát hóa tốt và không bị overfitting.

**2. Nhận xét về hiệu suất mô hình:**

* **Hiệu quả của mô hình:**
  + Với MAE = 1.602 trên tập kiểm tra, mô hình Neural Network hoạt động tốt hơn so với:
    - **Stacked Regressor (MAE = 2.223).**
    - **Simple Regressor (MAE = 2.565).**
  + Neural Network với 3 hidden layers có vẻ đã tận dụng được cấu trúc phi tuyến để nắm bắt tốt hơn mối quan hệ phức tạp trong dữ liệu.
* **Khả năng khái quát hóa:**
  + Chênh lệch MAE giữa tập huấn luyện (1.948) và tập kiểm tra (1.602) là rất nhỏ, cho thấy mô hình không bị hiện tượng overfitting.
  + Mô hình dường như hoạt động ổn định cả trên tập huấn luyện và kiểm tra.

**3. Đánh giá cấu trúc Neural Network (3 hidden layers):**

* **Lợi thế của 3 hidden layers:**
  + Với 3 hidden layers, mô hình có khả năng học các mối quan hệ phi tuyến phức tạp hơn so với các mô hình tuyến tính hoặc các mô hình đơn giản khác.
  + Cấu trúc này phù hợp với các bài toán có dữ liệu phức tạp, nhiều đặc trưng hoặc tính chất phi tuyến mạnh.
* **Tuy nhiên:**
  + Số lượng hidden layers không nên quá nhiều nếu dữ liệu có kích thước nhỏ hoặc ít đặc trưng, vì có thể dẫn đến overfitting hoặc khó tối ưu.
  + Hiệu quả của mô hình cũng phụ thuộc vào các yếu tố như số lượng neuron trong từng layer, hàm kích hoạt (activation function), và phương pháp tối ưu hóa (optimizer).

1. Mô hình Neural Network với 5 hidden

**1. Ý nghĩa của chỉ số MAE:**

* **MAE trên tập huấn luyện (train): 3.909**
  + Sai số trung bình trên tập huấn luyện khá cao, cho thấy mô hình không học tốt dữ liệu huấn luyện.
* **MAE trên tập kiểm tra (test): 3.913**
  + Sai số trung bình trên tập kiểm tra gần bằng với tập huấn luyện, điều này cho thấy mô hình không bị **overfitting** nhưng cũng không hoạt động hiệu quả trong việc dự đoán dữ liệu kiểm tra.

**2. Nhận xét hiệu suất:**

1. **Hiệu suất kém so với các mô hình khác:**
   * So với các mô hình khác như:
     + **Neural Network (3 hidden layers)**: MAE = 1.602 (test).
     + **Stacked Regressor**: MAE = 2.223.
     + **Simple Regressor**: MAE = 2.565.
   * Neural Network với **5 hidden layers** có MAE lớn hơn đáng kể (3.913), tức là dự đoán kém chính xác hơn.
2. **Dấu hiệu underfitting:**
   * Cả MAE trên tập huấn luyện và kiểm tra đều cao, điều này cho thấy mô hình có thể bị **underfitting**:
     + Mô hình không đủ khả năng học được mối quan hệ trong dữ liệu, dù số lượng hidden layers đã tăng lên.
     + Có thể do cấu trúc hoặc siêu tham số của mô hình chưa được tối ưu.

**3. Nguyên nhân tiềm năng:**

1. **Quá phức tạp nhưng không hiệu quả:**
   * Với 5 hidden layers, mô hình có thể trở nên quá phức tạp đối với dữ liệu, dẫn đến khó khăn trong việc huấn luyện và tối ưu hóa.
   * Nếu dữ liệu không đủ lớn hoặc không đủ đa dạng, việc tăng số lượng layers có thể phản tác dụng.
2. **Cài đặt không tối ưu:**
   * Learning rate có thể quá cao hoặc quá thấp, dẫn đến việc huấn luyện không đạt được hiệu suất tốt.
   * Số lượng neuron trong mỗi hidden layer, batch size, hoặc số epoch có thể chưa được tinh chỉnh phù hợp.
3. **Gradient vanishing/exploding:**
   * Với nhiều hidden layers, vấn đề **gradient vanishing** hoặc **gradient exploding** có thể xảy ra, đặc biệt nếu không sử dụng các kỹ thuật như normalization hoặc hàm kích hoạt phù hợp.
4. Mô hình Neural Network, 10 hidden layers

**1. Ý nghĩa của chỉ số MAE:**

* **MAE trên tập huấn luyện (train): 15.826**
  + Sai số trung bình trên tập huấn luyện là rất lớn (**15.826**), cho thấy mô hình không học tốt dữ liệu huấn luyện. Điều này phản ánh rõ ràng mô hình không nắm bắt được mối quan hệ giữa các biến đầu vào và biến mục tiêu.
* **MAE trên tập kiểm tra (test): 18.462**
  + Sai số trên tập kiểm tra còn cao hơn tập huấn luyện (**18.462**), cho thấy mô hình không chỉ underfitting mà còn không có khả năng tổng quát hóa.

**2. Nhận xét hiệu suất:**

* So với các mô hình Neural Network với **3 hidden layers** (MAE test = **1.602**) và **5 hidden layers** (MAE test = **3.913**), mô hình **10 hidden layers** hoạt động kém hơn đáng kể.
* Sai số lớn này cho thấy việc tăng số lượng hidden layers không cải thiện được hiệu suất, thậm chí làm giảm khả năng học của mô hình.

**3. Nguyên nhân tiềm năng:**

1. **Quá phức tạp so với dữ liệu:**
   * Với 10 hidden layers, mô hình có rất nhiều tham số cần học. Nếu dữ liệu không đủ lớn hoặc không đủ phức tạp, mô hình sẽ gặp khó khăn trong việc tối ưu hóa và dễ bị underfitting.
2. **Gradient vanishing/exploding:**
   * Với số lượng layers lớn, vấn đề **gradient vanishing** hoặc **gradient exploding** rất dễ xảy ra, đặc biệt khi không sử dụng các kỹ thuật như batch normalization hoặc residual connections.
3. **Thiết lập siêu tham số không phù hợp:**
   * Learning rate có thể quá thấp hoặc quá cao, dẫn đến mô hình không hội tụ hoặc hội tụ kém.
   * Số epoch không đủ hoặc không có kỹ thuật điều chỉnh như early stopping.
4. **Dữ liệu không đủ để huấn luyện:**
   * Mô hình quá phức tạp yêu cầu lượng dữ liệu lớn để học, nếu dữ liệu không đủ hoặc không đa dạng, mô hình dễ bị học kém.