

**Bài 4: HUẤN LUYỆN MÔ HÌNH VÀ LỖI***Hướng dẫn 2: Các độ đo đánh giá hiệu năng của mô hình hồi quy – Metrics***1. Giá trị trung bình tổng bình phương sai số (Mean Squared Error – MSE)**

Cho  $y \in R^m$  và  $\hat{y} \in R^m$  lần lượt là vector “nhân lớp” thực tế và vector chứa giá trị dự đoán của mô hình  $h_w(X)$  với  $w \in R^n$  và  $X \in R^{m \times n}$ . Để đo độ khác biệt giữa giá trị thực tế ( $y$ ) và giá trị dự đoán ( $\hat{y}$ ), phương pháp tính giá trị trung bình tổng bình phương sai số -  $MSE$  – được sử dụng:

$$MSE(y, \hat{y}) = \frac{1}{m} \sum_{i=0}^{m-1} (y_i - \hat{y}_i)^2$$

Công thức trên được viết lại theo cú pháp vector/ma trận để phù hợp với lập trình Python/Numpy như sau:

$$MSE(y, \hat{y}) = \frac{1}{m} \sum (y - \hat{y})^2$$

**2. Lập trình tính MSE**

Công thức Toán	Mã lệnh Python	Sử dụng thư viện sklearn.metrics
Độ đo MSE: $MSE(y, \hat{y}) = \frac{1}{m} \sum_{i=0}^{m-1} (y_i - \hat{y}_i)^2$ Viết lại: $MSE(y, \hat{y}) = \frac{1}{m} \sum (y - \hat{y})^2$	<pre>def mse(y, y_hat):     m = y.shape[0]     result =     (1/m)*np.sum(np.square(y - y_hat))     return result</pre>	<pre>from sklearn.metrics import mean_squared_error  ... print('\t\tSử dụng sklearn MSE: ', mean_squared_error(y_test, y_hat))</pre>

Tham khảo chương trình mẫu: trong classroom

### 3. Các độ đo phổ biến đánh giá hiệu năng mô hình hồi quy (regression models)

Ngoài độ đo MSE, còn có các độ đo khác dùng để đánh giá hiệu năng của mô hình hồi quy ([tham khảo link](#)). Hãy xây dựng bảng như trên, với yêu cầu:

- Liệt kê tất cả các độ đo phổ biến của mô hình hồi quy theo [tài liệu tham khảo](#);
- Cột 1: Viết công thức Toán gốc và công thức Toán chuyển đổi theo cú pháp vector/ma trận;
- Cột 2: Tự lập trình Python theo công thức Toán đã chuyển đổi;
- Cột 3: Sử dụng thư viện sklearn.metrics

Công thức Toán	Mã lệnh Python	Sử dụng sklearn.metrics
Độ đo $R^2$ :		
Độ đo $MAE$ :		
Độ đo $MSLE$ :		
Độ đo $MAPE$ :		