

## D. BÀI TẬP

1. Sử dụng thuộc tính (lotsize, bedrooms, stories, garagepl) trong tập dữ liệu Housing.csv để dự báo giá nhà. Sử dụng nghi thức hold-out để huấn luyện và đánh giá mô hình. Sử dụng chỉ số MSE và RMSE để đánh giá mô hình (tham khảo trên internet công thức MSE)

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2.$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2}$$

`sklearn.metrics.mean_squared_error(y_true, y_pred,...)`  
[http://scikit-learn.org/stable/modules/generated/sklearn.metrics.mean\\_squared\\_error.html](http://scikit-learn.org/stable/modules/generated/sklearn.metrics.mean_squared_error.html)

2. Dựa vào giải thuật LR1 ở phần hướng dẫn thực hành, anh/chị hãy cài đặt giải thuật **LR2** với công thức cập nhật các tham số theta0 và theta1 như bên dưới

$$\theta_j := \theta_j + \alpha \sum_{i=1}^m (y^{(i)} - h_{\theta}(x^{(i)})) x_j^{(i)}$$

- Cho tập dữ liệu gồm 3 phần tử như bảng bên dưới, (giống phần thực hành)

<b>X</b>	1	2	4
<b>Y</b>	2	3	6

- Tìm hàm hồi quy  $h(x)$  với giá trị khởi tạo  $\theta_0=0$ ,  $\theta_1=1$ , tốc độ học: 0.2, số bước lặp: 2
- Dự báo giá trị  $y$  cho 3 phần tử sau:  $x=0$ ,  $x=3$ ,  $x=5$
- So sánh với kết quả đã thực hiện bởi giải thuật **LR1** ở phần hướng dẫn thực hành