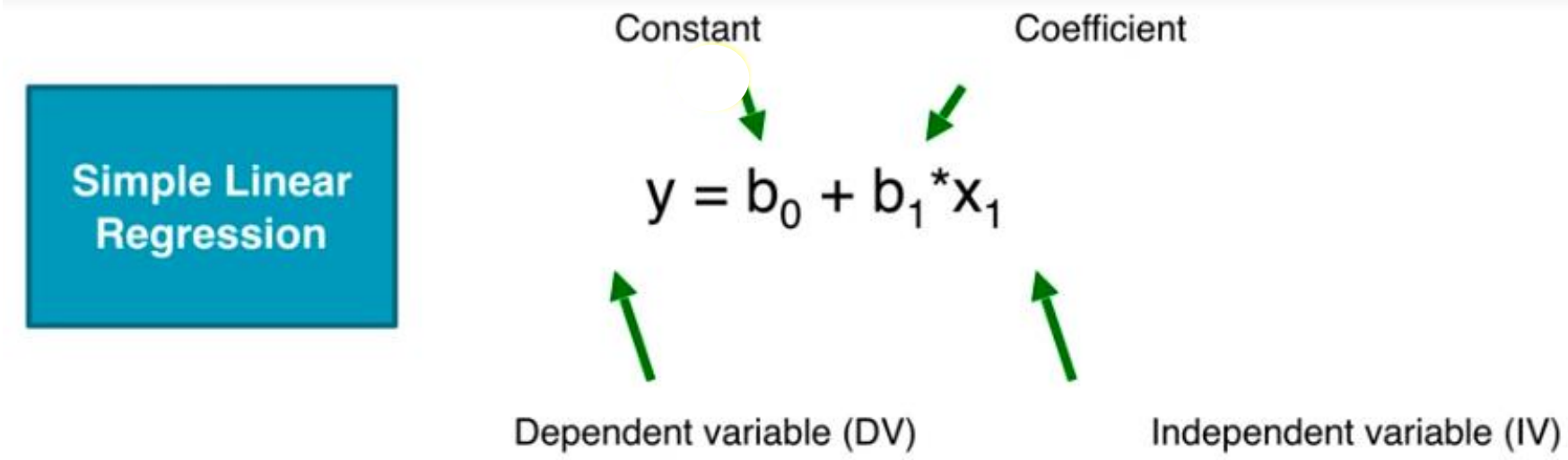
機器學習
Simple Linear Regression

授課老師：林彥廷

簡單線性回歸
Simple Linear Regression

- 一個應變數與一個自變數間之線性關係，此模型稱之為簡單線性迴歸模型
- 迴歸分析是將研究的變數區分為應變數及自變數，並建立應變數(Y)為自變數(X)之函數模型，然後再根據樣本所得的資料來估計函數模型的參數，其主要目的:
  - 是用來解釋資料過去的現象
  - 利用自變數(X)來預測應變數(Y)未來可能產生之數值。

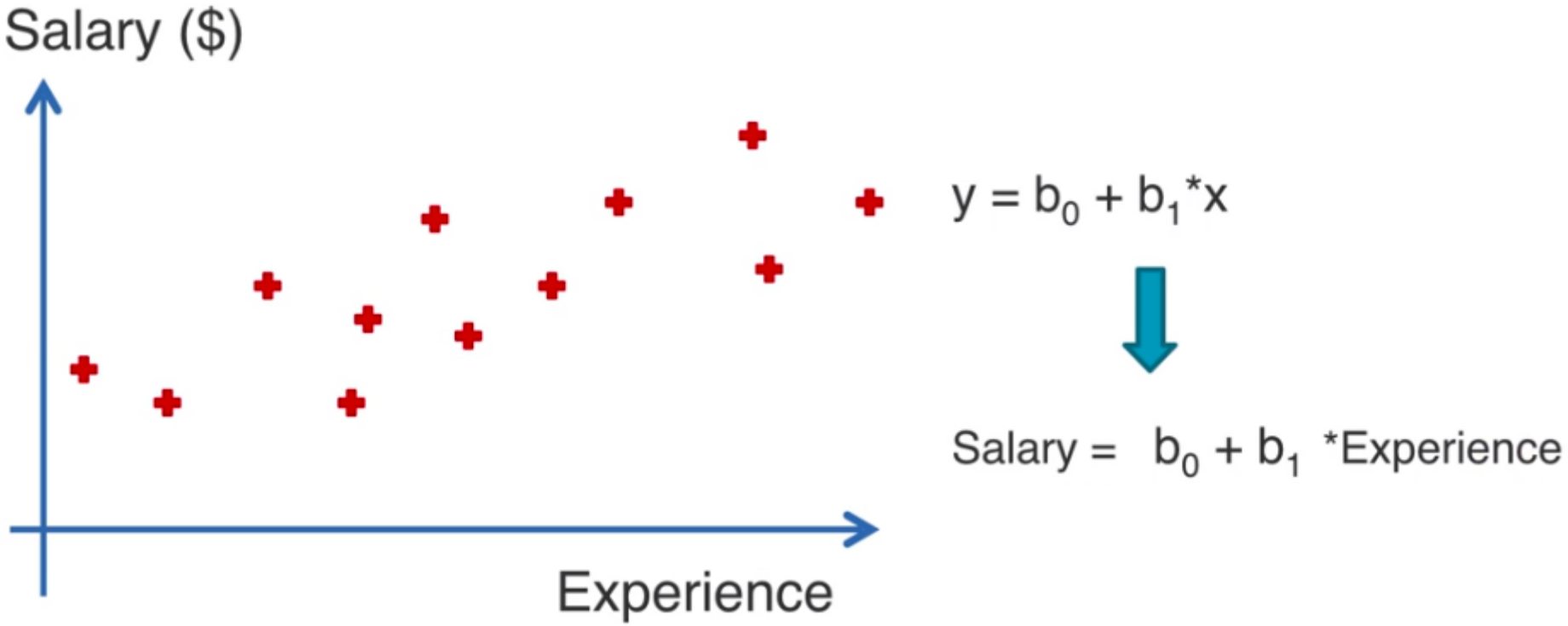
簡單線性回歸
Simple Linear Regression



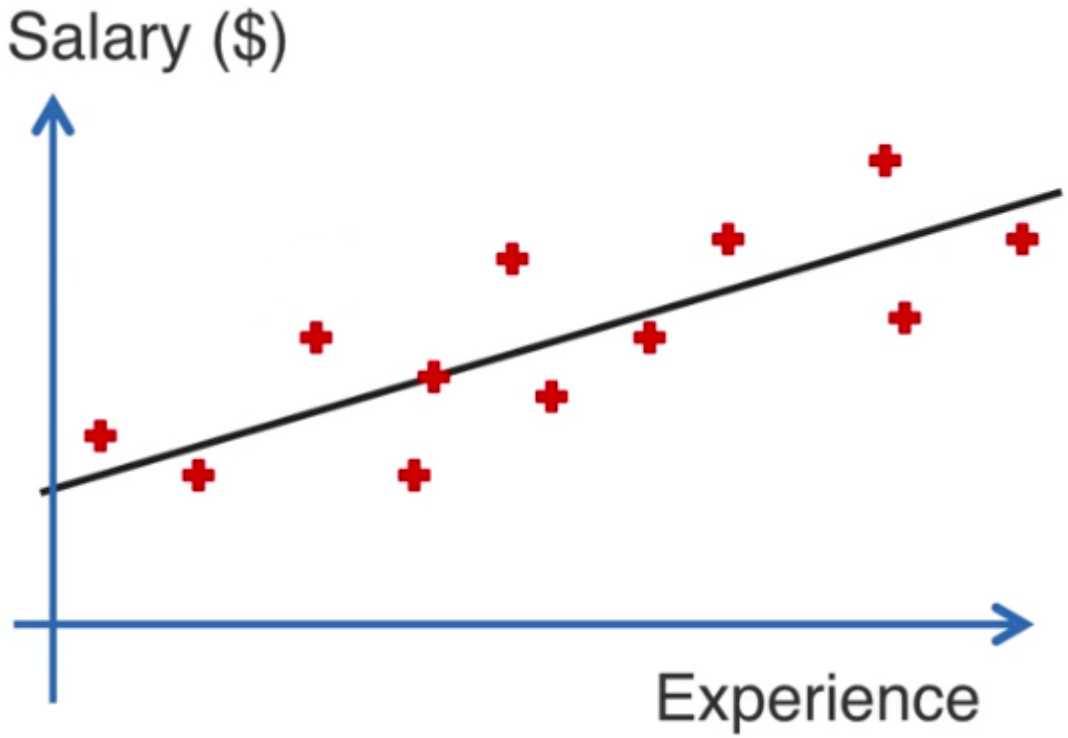
簡單線性回歸
Simple Linear Regression

1	YearsExpe	Salary
2	1.1	39343
3	1.3	46205
4	1.5	37731
5	2	43525
6	2.2	39891
7	2.9	56642
8	3	60150
9	3.2	54445
10	3.2	64445
11	3.7	57189
12	3.9	63218
13	4	55794
14	4	56957
15	4.1	57081
16	4.5	61111
17	4.9	67938
18	5.1	66029

簡單線性回歸
Simple Linear Regression



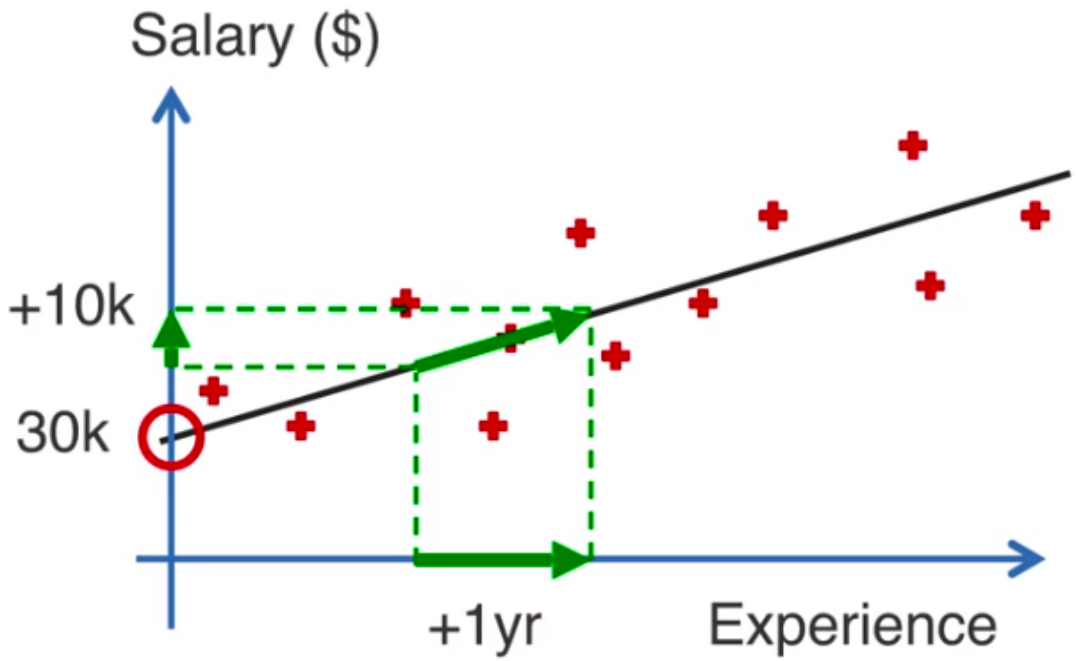
簡單線性回歸
Simple Linear Regression



$y = b_0 + b_1 * x$

Salary = $b_0 + b_1$ *Experience

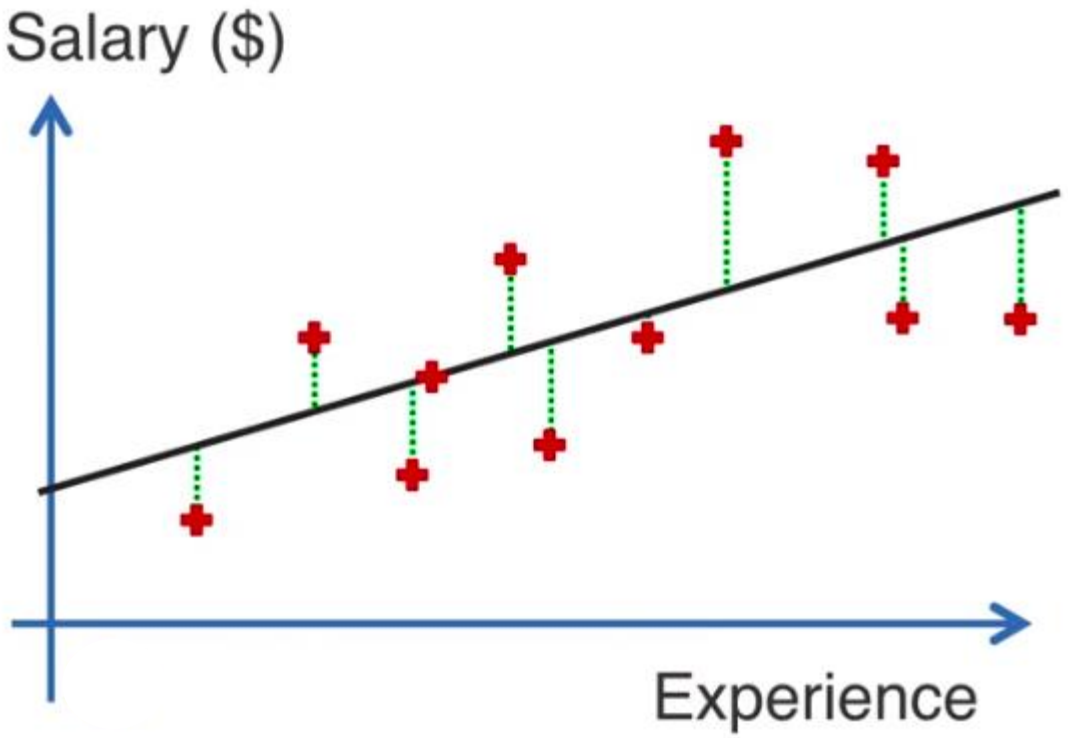
簡單線性回歸
Simple Linear Regression



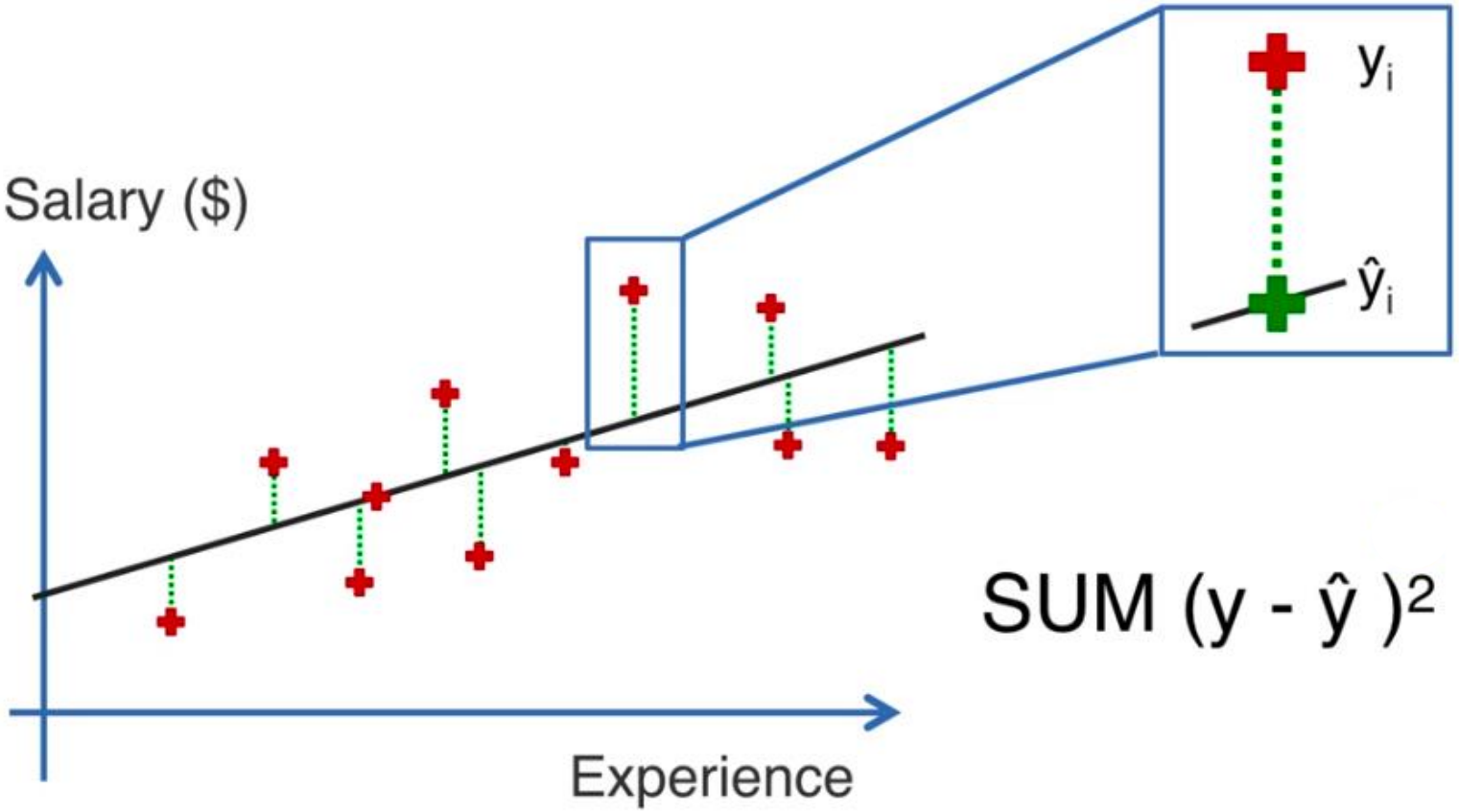
$$y = b_0 + b_1 * x$$

Salary = b_0 + b_1 * Experience

簡單線性回歸
Simple Linear Regression



簡單線性回歸
Simple Linear Regression



$$SUM\ (y - \hat{y})^2 \quad \text{-> min}$$

簡單線性回歸
Simple Linear Regression

- 簡單線性迴歸係數b_0、b_1之估計量

$$b_1 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$b_0 = \bar{y} - b_1\bar{x}$，其中 $x_1, x_2, \cdots, x_n$不全相等

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

簡單線性回歸
Simple Linear Regression

x	4	5	9	12
y	9	8	6	3

$$\bar{x} = (4+5+9+12)/4 = 7.5 \qquad \bar{y} = (9+8+6+3)/4 = 6.5$$

$$b_1 = \frac{(4-7.5)(9-6.5)+(5-7.5)(8-6.5)+(9-7.5)(6-6.5)+(12-7.5)(3-6.5)}{(4-7.5)^2+(5-7.5)^2+(9-7.5)^2+(12-7.5)^2} = -0.707$$

$$b_0 = 6.5-(-0.707)\times 7.5 = 11.803$$

由此可得此樣本迴歸直線爲 $\hat{y} = 11.803 - 0.707x$

THE END

ytlin@mail.nptu.edu.tw