

二手車價預測(II)

檔案 `car-resale-price.csv` 為過去二手車行情價資訊，其前五筆資訊如下：

	km_driven	new_price	year	engine_power	resale_price
0	16862	6.07	2019	81.80	6.06
1	18335	4.61	2018	67.00	4.50
2	15379	7.17	2019	83.83	6.97
3	28371	4.98	2017	73.00	4.74
4	32539	8.63	2018	88.50	8.39

其中

km_driven：行駛里程(kilometeres)

new_price：新車價(千美元 USD)

engine power：引擎馬力(bHp)

year：西元年

resale_price：二手車(千美元 USD)

前四欄位將擇二作為因變數 (independent variables)，resale price 為應變數 (dependent variable)；所選擇之因變數應使得，以 scikit learn 進行線性迴歸後，以測試集求算得出之判定係數值為最高。測資將提供 test_size 與 random_state；請以此做為切割訓練集 (Training set) 與測試集 (Test set) 之參數；迴歸完成後，請輸出線性預測函數；以測試集求得之判定係數 (coefficient of determination) 與均方誤差 (mean squared error)。

輸入檔案

`car-resale-price.csv`

輸入

測資計兩行，第一行為 test_size，第二行為 random_state。

輸出

線性預測函數，判定係數與均方誤差；浮點輸出至小數第五位，格式請見範例輸出。

範例輸入

0.3

32

範例輸出

預測函數：

$0.57425 \cdot \text{new_price} + 1.31071 \cdot \text{year} - 2641.22129$

判定係數：0.86886

均方誤差：18.05115