

以機器學習方法預測未來筆記型電腦售價

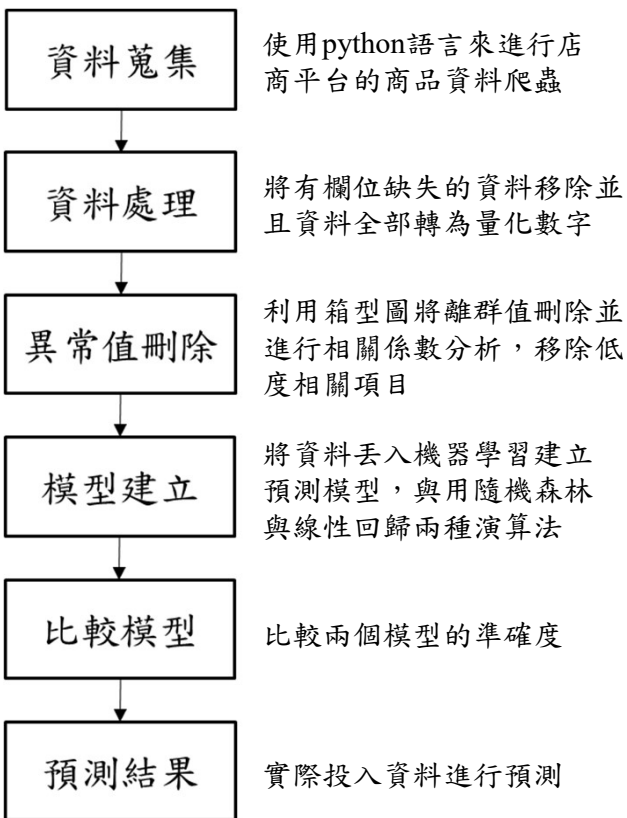
指導教授：鄭元杰 教授 學生：黃楷竣、黃自瑄

研究動機與目的

當製造商在訂定新產品價格的時候，商品的價格需要符合現今市面上消費者市場的價格要制定最佳的價格定價策略，來更好地滿足市場需求。

為了讓廠商的新產品價格符合現在消費者市場的售價，我們會使用商品特徵來預測商品的售價，避免將售價訂得太高，導致消費者的購買慾望降低。

研究流程



研究方法

• 線性回歸

線性回歸是一種常見的建模技術，用於建立變數之間的線性關係。它用於假設自變量（特徵）和因變量（目標變量）之間存在線性關係。

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$$

Y = 目標變量

X_i = 特徵值

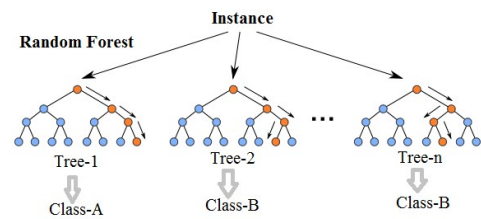
β_i = 線性回歸常數

ε = 誤差項

• 隨機森林

隨機森林是一種集成學習方法，在模型中，決策樹通過對特徵進行多次判斷和分割，生成樹狀結構，以預測目標變量的值。

Random Forest Simplified

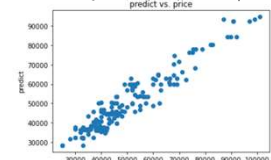


研究結果

線性迴歸



隨機森林迴歸



這兩張散佈圖為測試資料集的預測情況，橫軸為實際價格，縱軸為預測的價格。

均方誤差 (MSE) : 63760994.935319446
R-squared: 0.7371412120247873

均方誤差 (MSE) : 23179547.10926855
R-squared: 0.912563248553299

比較兩種模型的均方誤差與R平方，隨機森林迴歸的表現皆較為優秀，因此使用該模型進行預測。

結論

本研究使用了線性規劃以及隨機森林兩種方法來建立機器學習預測模型，從上述的研究過程中可以發現，相較於線性規劃預測模型，以隨機森林演算法所建立的預測模型其模型解釋能力較佳，所以我們會選擇使用隨機森林的預測模型作進一步的預測。

結論

1. 增加資料蒐集筆數，並且在每個價格區間都要有一定的資料量。
2. 可以從不同的平台上蒐集資料，減少不同平台會產生之變異性。
3. 模型的參數應該要定期更新，或是加上一個可變動的參數防止預測誤差。