

利用簡單機器學習模型 預測股市成長

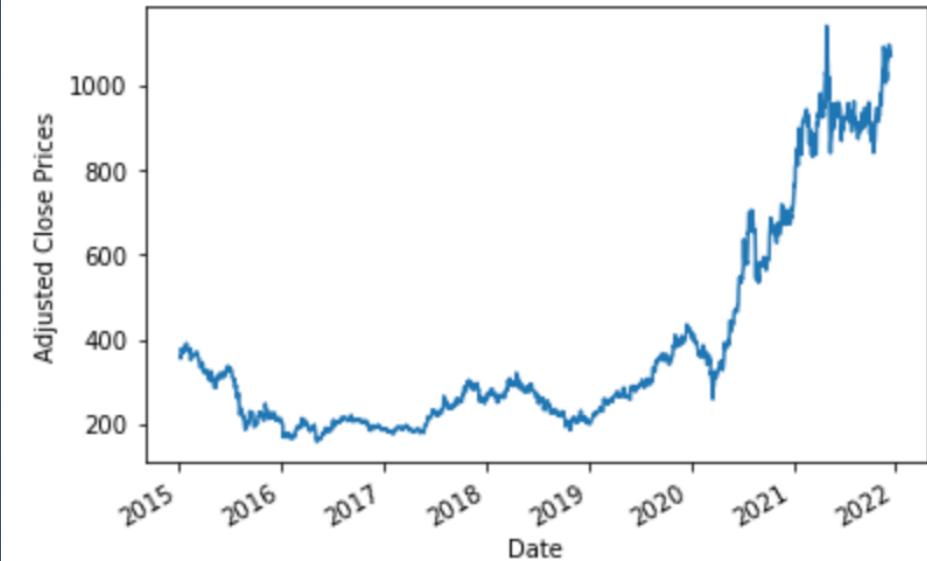
Using Basic Machine Learning Model to Predict Stock
Price Change

應數14級 林楷傑

收集資料 Collecting Data

- 預測公司：聯發科2454
- Time Interval of Training Data :
2015/01/04 ~ 2021/05/06
- Time Interval of Test Data : 2021/05/07
~ 2021/12/07
- 資料來源：Yahoo Finance API
- 機器學習套件：scikit-learn
- Git Hub:
<https://github.com/KJLdefeated/Stock-Prediction-with-Sveral-Machine-Learning-Model>

```
stock_data['Adj Close'].plot()  
plt.ylabel("Adjusted Close Prices")  
plt.show()
```



特徵工程 Feature Engineering

- 抽取一些對於預測結果有用的資訊
- 相對強弱指數(RSI)：藉由比較價格升降運動，以表達價格強度的技術分析工具。
- 移動平均(MA)：用來觀察價格波動趨勢的技術指標，可確認現有的價格趨勢，判斷未來潛在的價格走勢。
- 欲預測目標：五天以後的價格百分比變化

Date	Open	High	Low	Close	Adj Close	Volume
2015-01-05	460.0	467.0	460.0	462.0	357.540588	2459000
2015-01-06	460.0	462.0	458.0	458.0	354.444946	5915000
2015-01-07	468.0	487.0	466.5	486.5	376.501068	13046000
2015-01-08	492.5	494.5	484.5	485.0	375.340210	8367000
2015-01-09	492.0	495.0	488.0	488.0	377.661896	7473000
...
2021-12-08	1085.0	1110.0	1080.0	1095.0	1095.000000	7323171
2021-12-09	1095.0	1100.0	1075.0	1085.0	1085.000000	2463930
2021-12-10	1075.0	1085.0	1065.0	1075.0	1075.000000	2943248
2021-12-13	1070.0	1095.0	1070.0	1075.0	1075.000000	3276917
2021-12-14	1070.0	1075.0	1050.0	1070.0	1070.000000	4130770



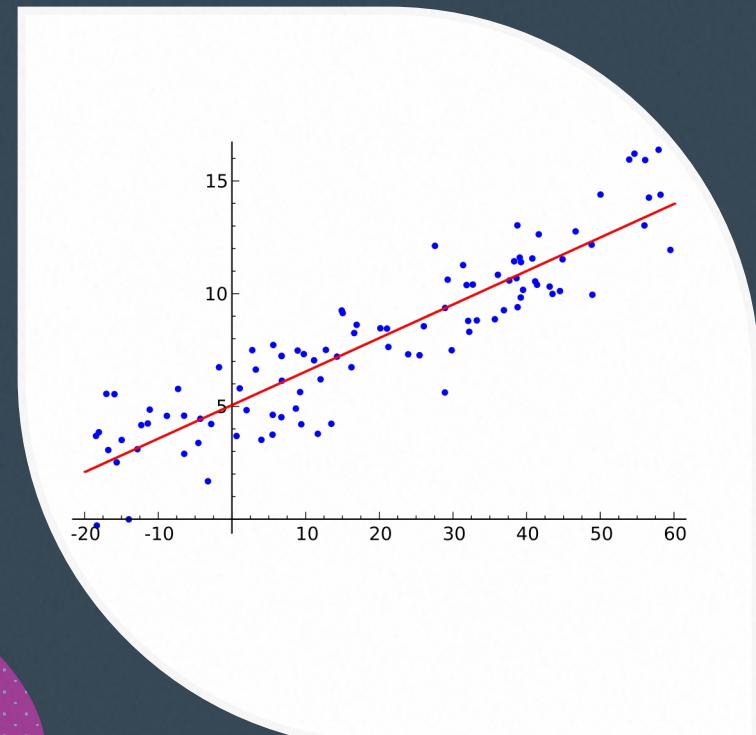
Linear Regression

線性迴歸模型

- 概念：利用稱為線性迴歸方程式的最小平方函數對一個或多個自變數和應變數之間關係進行建模的一種迴歸分析。

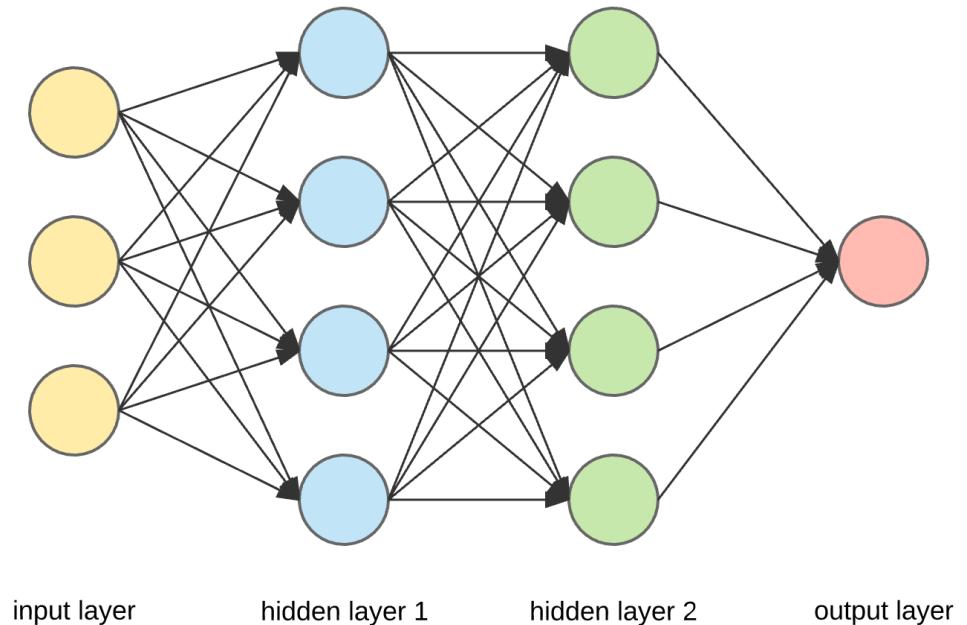
```
lr_model = LinearRegression()  
lr_model.fit(X_train, y_train)  
y_pred = lr_model.predict(X_test)
```

- 平均絕對值誤差(Mean Absolute Error):3.54%
- 平均次方誤差(Mean Squared Error):0.21%
- 方均根誤差(Root Mean Squared Error):4.61%



Neural Network Regression 類神經網路迴歸模型

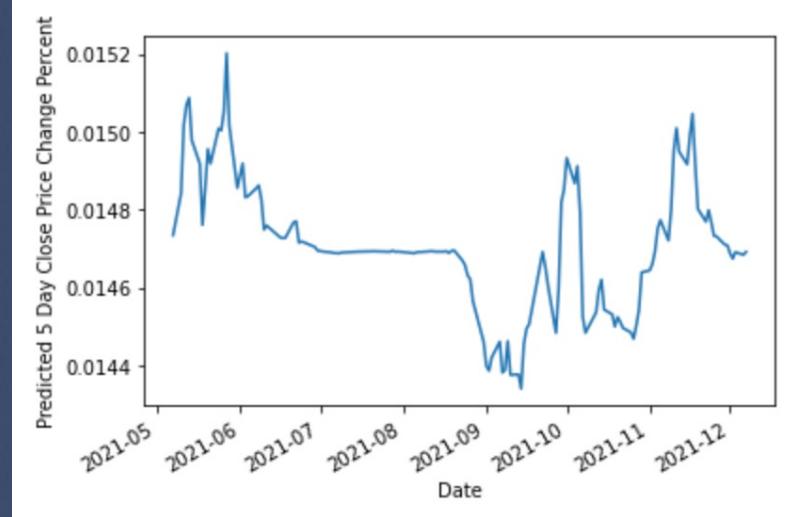
- 概念：模仿生物神經網路的結構和功能的數學模型，用於對函式進行估計和近似。
- `MLPRegressor (max_iter=500, activation = 'logistic', hidden_layer_sizes = (100, 100))`
- 平均絕對值誤差(Mean Absolute Error):3.17%
- 平均次方誤差(Mean Squared Error):0.17%
- 方均根誤差(Root Mean Squared Error):4.21%



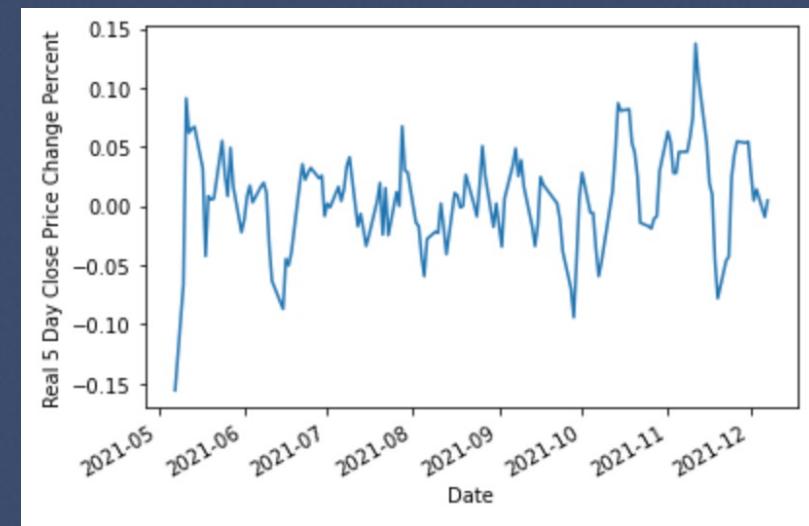
結論 Conclusion

用模型預測股市：

1. 收集資料
 2. 特徵工程
 3. 建構模型
 4. 檢視預測結果
-
- 機器學習可一次分析大量資料，對可量化資訊進行數學分析，但只看數字，無法觀測不能被量化的資訊（事件日、情報面），被歷史資料限制，對於預測沒辦法達到很高的精準度，然而，這些估計可以作為投資策略的參考，以降低風險。



預測資料



真實資料