



機器學習第二組會議紀錄

大家好，我們是機器學習第二組組員，以下是我們專案的會議紀錄

指導老師：

巨資：蔡芸琤

經濟：林巧涵

第二組組員

# 學號	Aa 姓名	≡ 系級
9370013	黃碩彥	巨資碩一
9351081	劉昊東	經碩一
9370006	魏嗣宸	巨資碩一
9351016	李冠瑩	經碩一
	Untitled	

第一次會議記錄

會議記錄：

問題定義與說明：探討影響公司成長率的指標-以台灣股市資料為例

資料來源：TEJ資料(現金流量表、損益表、股票年報)

前處理步驟：去空值、去除離異值、敘述統計、看是否要正則化（可以再討論）

資料EDA與視覺化解讀：相關係數圖、盒須圖、分佈圖（可以再討論）

如何分工：

資料蒐集：李冠瑩、劉昊東

預處理：黃碩彥、魏嗣宸

資料視覺化：黃碩彥、魏嗣宸

第二次會議記錄

會議記錄：

問題定義與說明：探討影響公司成長率的指標-以台灣股市資料為例

資料來源：TEJ資料(現金流量表、損益表、股票年報)

前處理步驟：去空值、去除離異值、敘述統計（可以再討論）

資料EDA與視覺化解讀：相關係數圖、盒須圖、資料分佈圖（可以再討論）

如何分工：

報告人：劉昊東

簡報：李冠瑩

資料蒐集：李冠瑩、劉昊東

預處理：黃碩彥、魏嗣宸

資料視覺化：黃碩彥、魏嗣宸

第三次會議記錄

會議記錄：

問題定義與說明：探討影響公司成長率的指標-以台灣股市資料為例

資料來源：TEJ資料(現金流量表、損益表、股票年報)

前處理步驟：去空值、去除離異值、敘述統計（可以再討論）

資料EDA與視覺化解讀：相關係數圖、盒須圖、資料分佈圖（可以再討論）

如何分工：

報告人：劉昊東

簡報：李冠瑩

資料蒐集：李冠瑩、劉昊東

預處理：黃碩彥、魏嗣宸

資料視覺化：黃碩彥、魏嗣宸

如何解決問題：

透過四個指標(Cash flow,Debt ratio,ROE,Total asset turnover)預測公司是否成長(0,1)

第四次會議記錄

會議記錄：

問題定義與說明：探討影響公司成長率的指標-以台灣股市資料為例

資料來源：TEJ資料(現金流量表、損益表、股票年報)

前處理步驟：去空值、去除離異值、敘述統計（可以再討論）

資料EDA與視覺化解讀：相關係數圖、盒須圖、資料分佈圖（可以再討論）

劉昊東：

資料更換：改成大盤指數資料

黃碩彥：

演算法討論：運用隨機森林與Kmeans、視覺化、資料前處理

魏嗣宸：

運用共線圖：畫出公司間的指標相關關係

黃碩彥：

特徵工程：將Kmeans分群視為一個特徵，當作隨機森林預測的自變數

討論拍攝影片：

15分鐘:劉昊東

3分鐘:魏嗣宸

影片內容討論：

黃碩彥、魏嗣宸、李冠瑩

Github 網址

<https://github.com/MachineLearningGroup2/Project>
