

專題期末報告

411412292 吳品儒

在這學期我專題負責的是深度學習中的圖形辨識，主要工作內容是如何利用深度強化學習 DRL 與卷積神經網路 CNN 分析股價時間序列資料上，能否透過 CNN 辨識股票 K 線圖提供 buy/sell/hold 的訊號。因此在學期期初我先是與許祐凱整理常見的技術分析圖形像是頭肩頂(底)、上升(下降)三角等等大約 20 個型態，並探討其背後支撐壓力、成交量等故事背景。在本學期中我嘗試做了一次 Kaggle 上 CNN 的實作，並在期中後閱讀多篇學術期刊論文。本學期最主要在於專業知識與程式技術的建構，並閱讀一些相關文獻做成文獻回顧，

一、 CNN Cats&Dogs 實作

我在 kaggle 上找了一個貓狗辨識的議題做實作，幫助我更了解 CNN 的運作原理跟流程，並熟悉 CNN 在 python 上如何撰寫以及有哪些參數可以做調整等等。整套流程為安裝套件>標記 dog&cat>將圖片調整為套件需要的格式>將圖片、標籤向量化轉換為 numpy 數據組>將數據標準化並打亂>建構 CNN 模型，包括 Dense、Activation、Dropout、Flattem、Conv2D、MaxPooling2D、SGD&ADAM>設立 batchsize、epochs 作為訓練>結果顯示為圖表>加入 test>Confusion Martirx。

資料網址:<https://www.kaggle.com/datasets/tongpython/cat-and-dog/code>

二、 文獻回顧

a. Interpretable image-based deep learning for price trend prediction in ETF markets

探討將金融時間序列轉換為圖像的可行性，作者自行建立 CNN+通道注意機制的 CS-ACNN 模型，利圖像化的深度學習模型來提 ETF 價格趨勢的預測準確性。最終目標是預測上漲 or 下跌，而非商品的價格。文獻中標的使用 S&P500、Hang Seng Index ETF、SSE 50 ETF，本文提出兩種方法，第一是 Augmented candlestick charts，強化部分黑白灰燭圖的資訊讓機器更能捕捉到相關變動。第二是較為常見的 GAF，將蠟燭圖正規化並轉換為矩陣。本文中故事性較為理工科，作者想開發出 Sharpe Ratio 更高、準確率更高的模型應用在生活中。

b. Deep reinforcement learning stock market trading, utilizing a CNN with candlestick images

這篇文章是用 OpenAI Gym 模擬市場，再將蠟燭圖作為 CNN 輸入，透過 DDQN 算法，根據 Qvalue 產出 long、short、no position，標的是 S&P 前 30 大個股 OHLC，本文目標希望解釋 DRL 判斷的決策，因此有加入 Heat Map 來視覺化圖表，並探討 CNN+DDQN 能否打敗大盤 S&P500。DQN 為 Q-learning 的延伸，我認為這是我們可以參考的模型，因為他是 off-policy 且他透過 neural 預測下一步 action，可以解省龐大 Q-table。本文故事也屬於偏技術面的東西，較少關於經濟直覺的結果，大多在比較演算法之間的優劣。

三、景氣循環投資組合-Benchmark

在寒假期間我負責做 Benchmark 的研究，我們希望以景氣循環四種類別：衰退、復甦、擴張、趨緩四種階段決定股票、債券、商品的權重，並將此投組作為 Benchmark 提供後來 DQN 的對照組。我有選擇兩種適合的總經指標作為景氣循環的判斷方式。

1. 美林時鐘:由美林證券根據美國過去 30 年數據編制的方法，主要由 CPI+GDP 作為判斷，我個人認為可以再加入失業率或領先指標 PMI 作為判斷。

低增長 + 低通脹 + 高失業：適合債券	>衰退
高增長 + 低通脹 + 失業率下降：適合股票	>復甦
高增長 + 高通脹 + 失業率低：適合原物料	>擴張
低增長 + 高通脹 + 失業率上升：適合債券	>趨緩

2. 長短天期殖利率利差

通常為 10 年期美國公債利率-2 年期美國公債利率，藉由判斷利差為縮小、放大、負數來判斷目前景氣處於哪個階段。

利差放大:主要為景氣衰退時，央行降息刺激經濟，利差放大。>衰退期

利差縮小:主要為經濟良好且通膨發生時，央行升息導致短天期利率上升，進而使利差縮小。>復甦期+擴張期

利率倒掛:市場過熱處於景氣循環高點，經濟開始緊縮，資金開始進駐債。>擴張期末期+趨緩期

四、標的

我們組員內部討論後主要想以美國作為主要投資國家，以及想以大盤作為標的。我認為做模型比對可以以指數為主，股:S&P500、債: 美國十年期公債殖利率、商品: 以黃金為主可使用 LBMA Gold Price，以全部大眾商品可使用 S&P GSCI。但若是想做投資績效的評比則可以選擇 ETF，股:SPY、債 IEF、商品:依據選擇黃金或全部商品決定 GLD 或 GSG。我認為 ETF 的優勢在於它相較指數為可直接投資的標的，在績效評估上更貼近現實層面，但最大缺點在於 ETF 為近 20 年發明的商品，而景氣循環一輪參考 kitchen、juglar 循環及 NBER 認定的循環多為 5~10 年，不確定是否資料時間太短會影響到模型結果，可能要在與組員跟老師做討論。

五、 總結

這學期主要在建立對 DRL 與 CNN 領域的知識，我對未來研究方向有進一步的了解與認識。接下來的工作及目標為建立 Benchmark，內容包含如何用總經指標判斷景氣循環處在哪個階段以及如何根據馬可維茲建立股、債、商品的投資組合。