

# 機器學習第一次報告

陳慧如

Department of Applied Mathematics  
National Chung Hsing University  
South District, Taichung City 40227, Taiwan  
luluchenyo@smail.nchu.edu.tw

## I. 研究主題與目標

我想要研究的主題是跟財金領域相關的，主要目標是想要在金融市場上，應用機器學習的方法進行交易，盡可能的獲取最大的利潤。機器學習在各個領域的應用上已經逐漸成熟，舉凡在戰略遊戲、手寫語音辨識、醫學診斷還有自然語言辨識等等都做出了一定成績。而在金融領域上，拿股票市場來看，機器可以透過歷史股價自動提取特徵藉以預測未來的股價走勢，比起傳統的技術分析、基本面分析還有籌碼面分析等等方式，機器也許可以挖掘出傳統上人類從未考慮過的因素，以及各個因素之間是如何相互影響，相較於人類更全面的觀測整個市場的動向；再者，透過強化學習的概念，機器可以在股票市場中進行買進、賣出以及持平等動作，並在得到損益時更新獎勵，在與股票市場互動的過程中，自動化的學習到最適交易決策。要如何預測金融市場的動向是多年來備受關注的議題，隨著軟硬體技術不斷的進步，期許在未來人類可以靠著近年來蓬勃發展的機器學習方法，將人類對金融市場的理解推向一個新的高度。

## II. 文獻回顧

A. Deng, Y., Bao, F., Kong, Y., Ren, Z., & Dai, Q. (2017). Deep Direct Reinforcement Learning for Financial Signal Representation and Trading. *IEEE Transactions on Neural Networks and Learning Systems*, 28(3), 653–664. doi: 10.1109/tnnls.2016.2522401

首先，該文將歷史股價資料前後期取差，得到每一期相較於前一期的漲跌幅，接著用 K-means 方法將其做 fuzzy 的前處理，以模糊財務資料的隨機性，再來選擇用深度的 RNN 架構 (Fig. 1) 以進行特徵學習，體現了財務資料具有的時間序列特性，最後輸出為買、賣及持平三種決策。另外，由於 DRNN 的架構較為複雜，該文採用了 Task-Aware BPTT 方法來解決梯度消失的問題。

B. Jeong, G., & Kim, H. Y. (2019). Improving financial trading decisions using deep Q-learning: Predicting the number of shares, action strategies, and transfer learning. *Expert Systems with Applications*, 117, 125 – 138. doi: 10.1016/j.eswa.2018.09.036

相較於前一篇，決定買、賣及持平之外，更進一步決定要交易的股票數量。為了解決數據不足造成的 over-fitting，該文用了四個國家的股價加權指數 (S&P500、KOSPI、HSI 及 EuroStoxx50) 來進行 transfer learning。

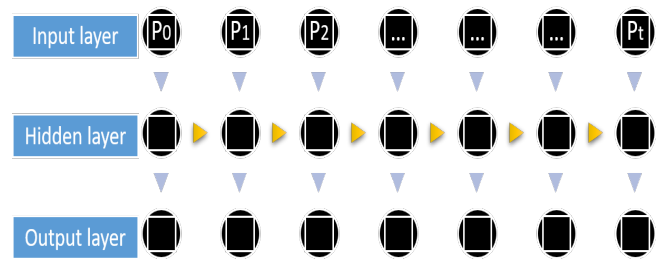


Fig. 1: RNN

在決策方面，主要用到的方法為 Q-learning，首先將股價資料投入兩層的 DNN 架構，接著分成兩支，其中一支以 Q function 的值選取在某個市場情況下，買、賣或是持平中最具有價值的決策，另外一支用 Num Q 函數決定要交易的股票數量。該文另外定義了市場混亂的情況，即在某個時間點，進行買進或是賣出的行動價值相差不遠時，將進行預設的動作（平倉或是不進行動作），而非訓練出來的決策。

C. Carapuço, J., Neves, R., & Horta, N. (2018). Reinforcement learning applied to Forex trading. *Applied Soft Computing*, 73, 783–794. doi: 10.1016/j.asoc.2018.09.017

該文將 Q 網路系統應用在外匯市場上，輸入為歐元兌美元的匯率資料，輸出為買、賣及持平，投資數量固定為一單位，考慮 Sharpe 比率及 Sortino 比率當作風險。在 reward 的設計上，新增未實現損益的增減，避免在標的價格持續下跌的時候，agent 因為只考慮到平倉後會得到的負 reward 而選擇繼續持倉。另外，該文運用三個歐元兌美元的數據集，進行 12 個月的訓練以及 6 個月的測試來選擇 hyper-parameter，試圖取得穩定的且較一般化的學習曲線。

D. Ryan, L., Yi, W., Aviv, T., Jean, H., Pieter, A., & Igor, M. (2017, June 7). Multi-Agent Actor-Critic for Mixed Cooperative-Competitive Environments. Retrieved from <https://arxiv.org/abs/1706.02275>.

該文討論了多個玩家運用 Actor-Critic 方法一同進行遊戲的情形，主要分為三類：競爭、合作還有競爭兼合作的情況。在競爭遊戲下，critic 會藉由預測其他 agents 的動作，然後決定傳給自己 agent 的 actor 的訊息。如果可以將該文的模型運用在股票市場上的話，機器藉由

預測其他股民的動作來決定交易決策，就假設市場上的 agents 都預測某一支股票會上漲來說，隨著進行買進動作的 agents 越來越多，成交價越來越貴。當運用此模型在股票市場上的時候，agent 可能會試圖比其他 agents 先進行買進的動作（買低），接著在其他 agents 還在買進或是還未賣出的時候進行賣出（賣高）。

*E. Feuerriegel, S., & Helmut P. (2016). News-based trading strategies. Decision Support Systems 90, 65-74. doi: 10.1016/j.dss.2016.06.020*

想要一個競爭環境下的股票市場獲得最大利益，比別的競爭者先一步進行動作是必要的。我們通常說股價資訊反映了市場上所有舊資訊以及新資訊，在我前幾篇參考的文獻中，都只考慮了舊資訊（歷史股價），而新資訊（金融事件）也是會影響股價的一部份因素，我們只能利用歷史股價來得到金融事件一開始發生時，市場上進行內線交易的資訊，一旦金融事件的新聞播報出來，假設市場上有一群人參考財經新聞在進行投資，這些人的交易動作只會反映在下一期的股價資料上。假設有一則新聞，具有足夠的價值，可以使得某一支股票上漲的機率提升到 agent 可以進行買進的狀態，有多參考新聞資訊的 agent 就可以比只參考歷史股價的 agent 早一步進場。所以參考該文，利用新聞情緒以及歷史股價進行交易決策，當新聞情緒的指向與歷史股價預測的股價指向為同一方向時，agent 才會將該狀態視為可交易訊號，最後利用 Q-learning 方法訓練交易決策。