

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

Andrew Brim,, Nicholas S. Flann

這篇文章是用OpenAI Gym模擬市場，再來將蠟燭圖作為CNN輸入，透過DDQN算法，根據Qvalue產出long、short、no position

目標:

- 1.探討黑箱，將Feature視覺化來解釋CNN的注意力集中點**
- 2.探討其在重大市場事件（ex:2020年新冠疫情股市崩盤）期間的表現**
- 3.測試CNN+DDQN與是否超越能S&P 500**

資料集:

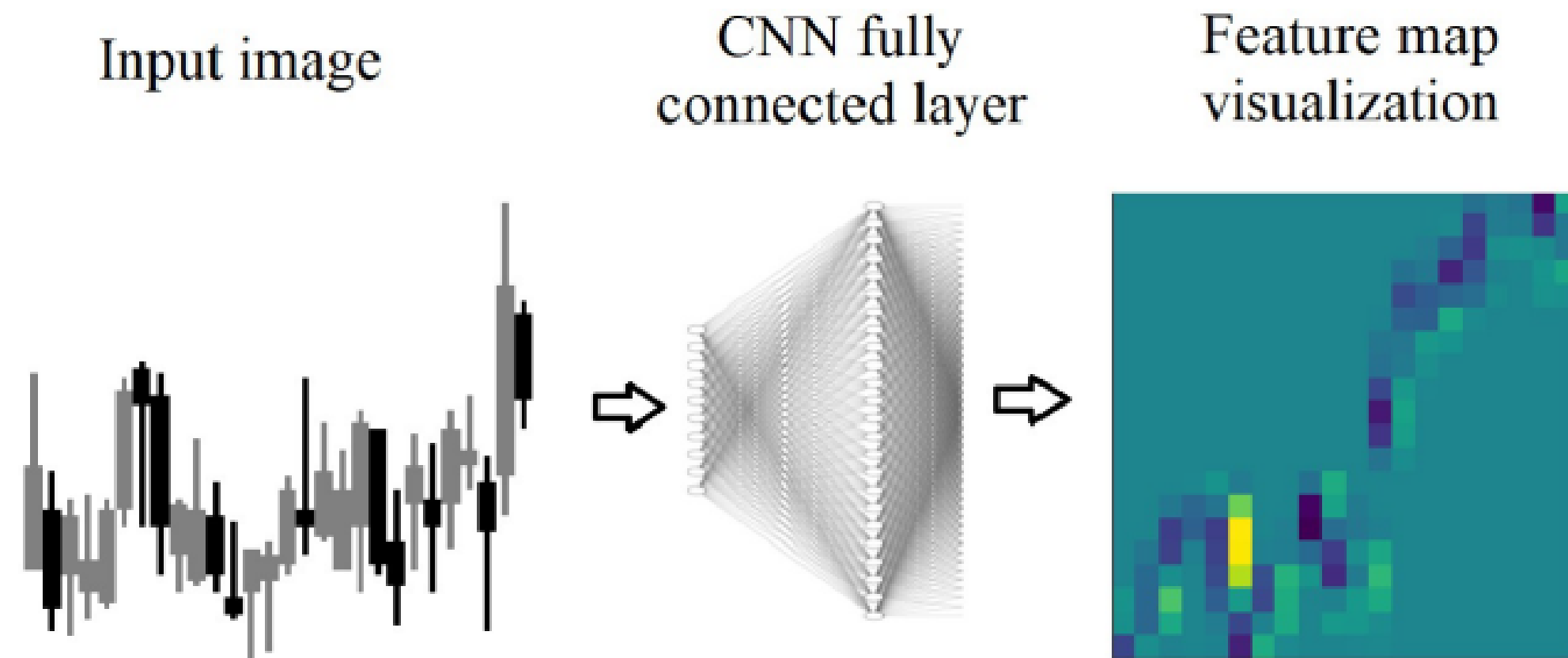
S&P 500中前30大個股OHLC

Training:January 2013 -December 2019. >52920 observations

Testing:January 2020 -June 2020. >3780 observations

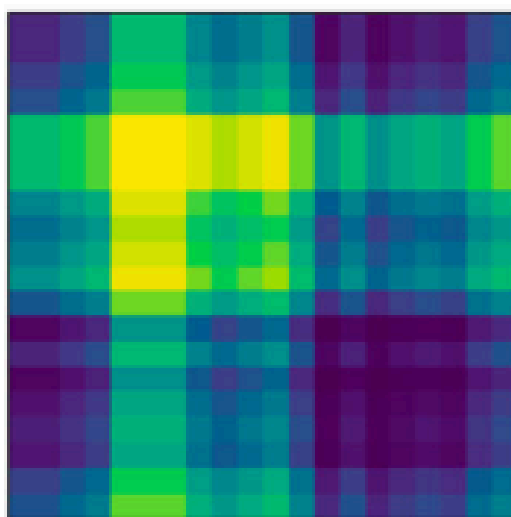
資料處理:

- 自訂蠟燭圖，灰色為上漲、黑色為下跌，每張為過去28天日K。
尺寸為84x84像素，O、C為兩像素寬，H、L為一像素寬。
- 利用OpenAI Gym模擬股票市場，透過輸入燭圖的CNN來逼近Q函數，輸出 long、short、no position。

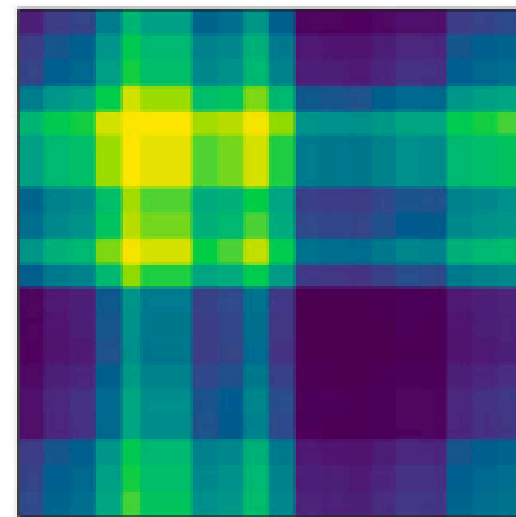


資料處理:

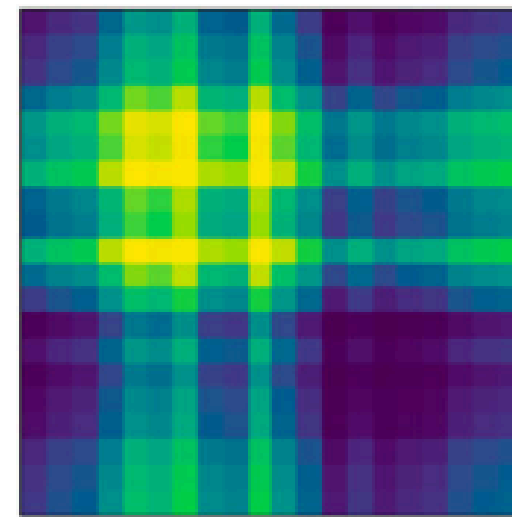
- 不考慮使用GASF原因，作者希望模擬交易員觀看蠟燭圖型態進行決策的過程，並觀察每根k線的關係。
- 若使用GASF，將會轉為一種對稱矩陣，儘管保留了時間序列資料，但無法對於K線型態(ex:十字星、吞沒、影像長短)進行有效分析



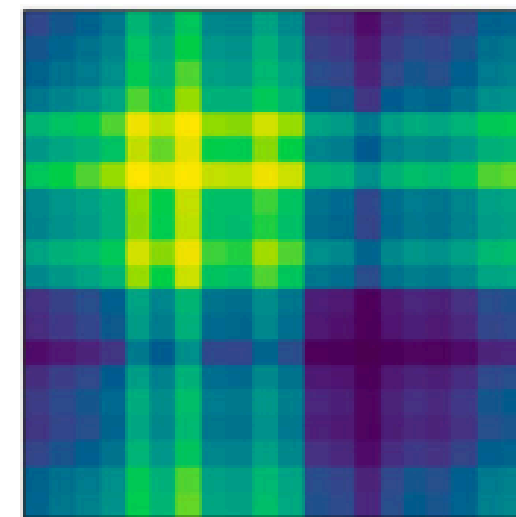
(a)



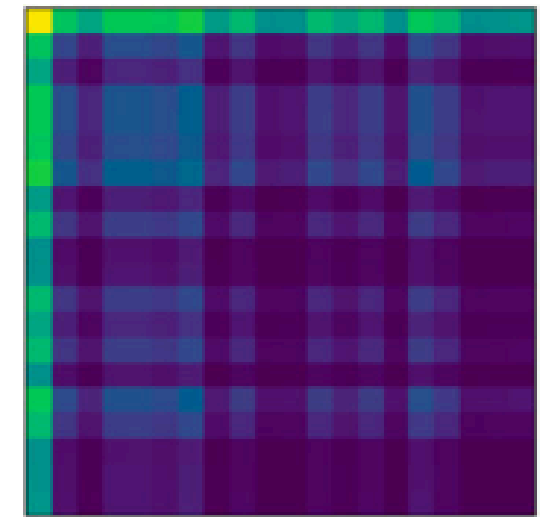
(b)



(c)



(d)



(e)

DQN&DDQN

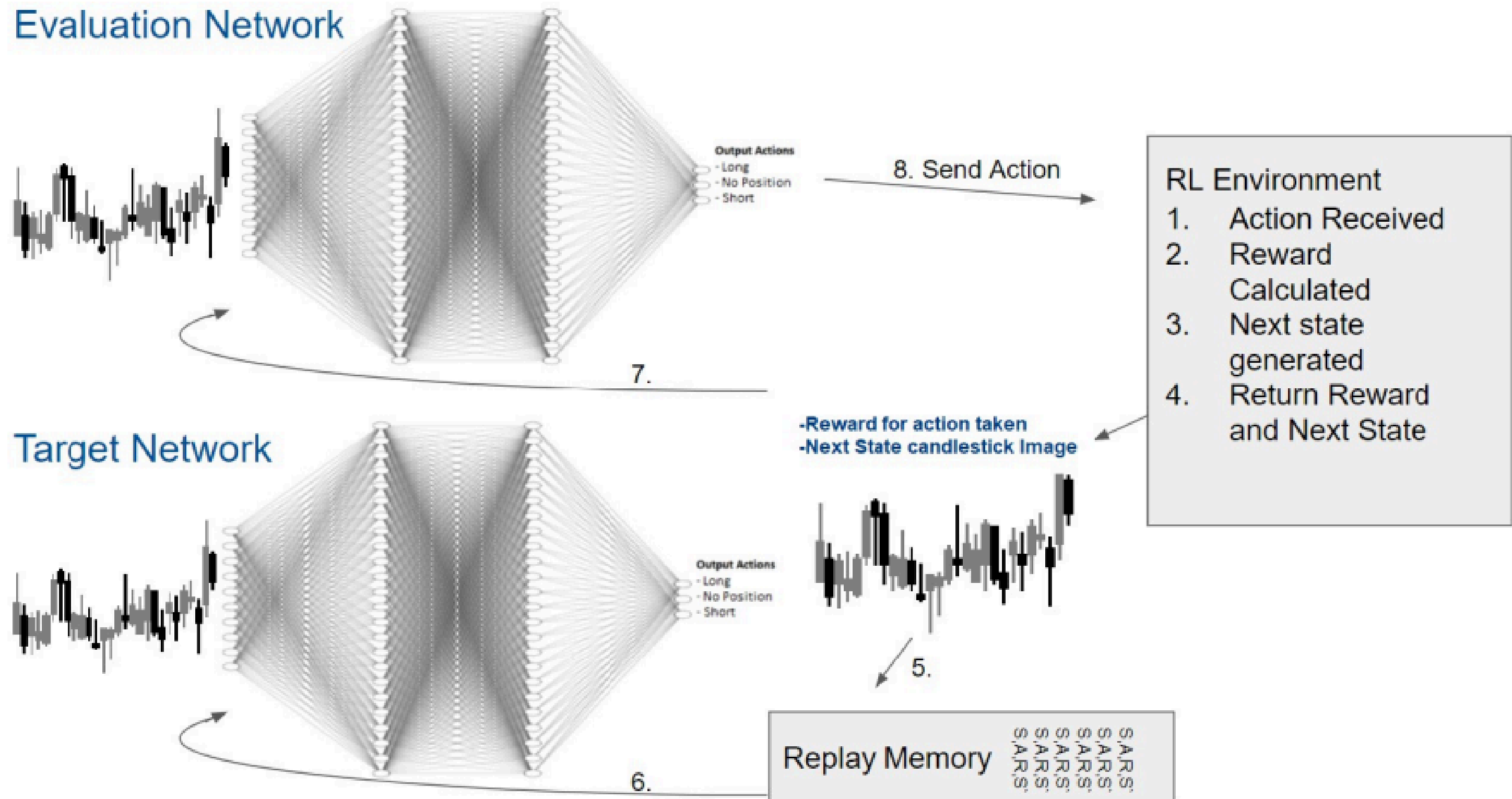
- Deep Learning + Reinforcement Learning
- 透過neural來預測接下來的action，可以省去龐大的Q-table，更可以計算更複雜的環境狀態來提升強化學習的效果
- 主要差異在於如何計算目標Q值 (Target Q Value)

DQN由同一網路同時負責選擇和估算，容易導致過估計偏差

DDQN將動作選擇與價值估算分開，有效減少過估計偏差，提升訓練穩定性和性能

1.DDQN 使用兩個網絡分離行動選擇和目標估算，減少了 Q 值的過估計問題

2.模型能調整對圖像特徵的關注，ex疫情間集中注意最近的蠟燭數據



OpenAI Gym中的DDQN訓練獎勵計算

$$T = a * r * N$$

T：訓練獎勵

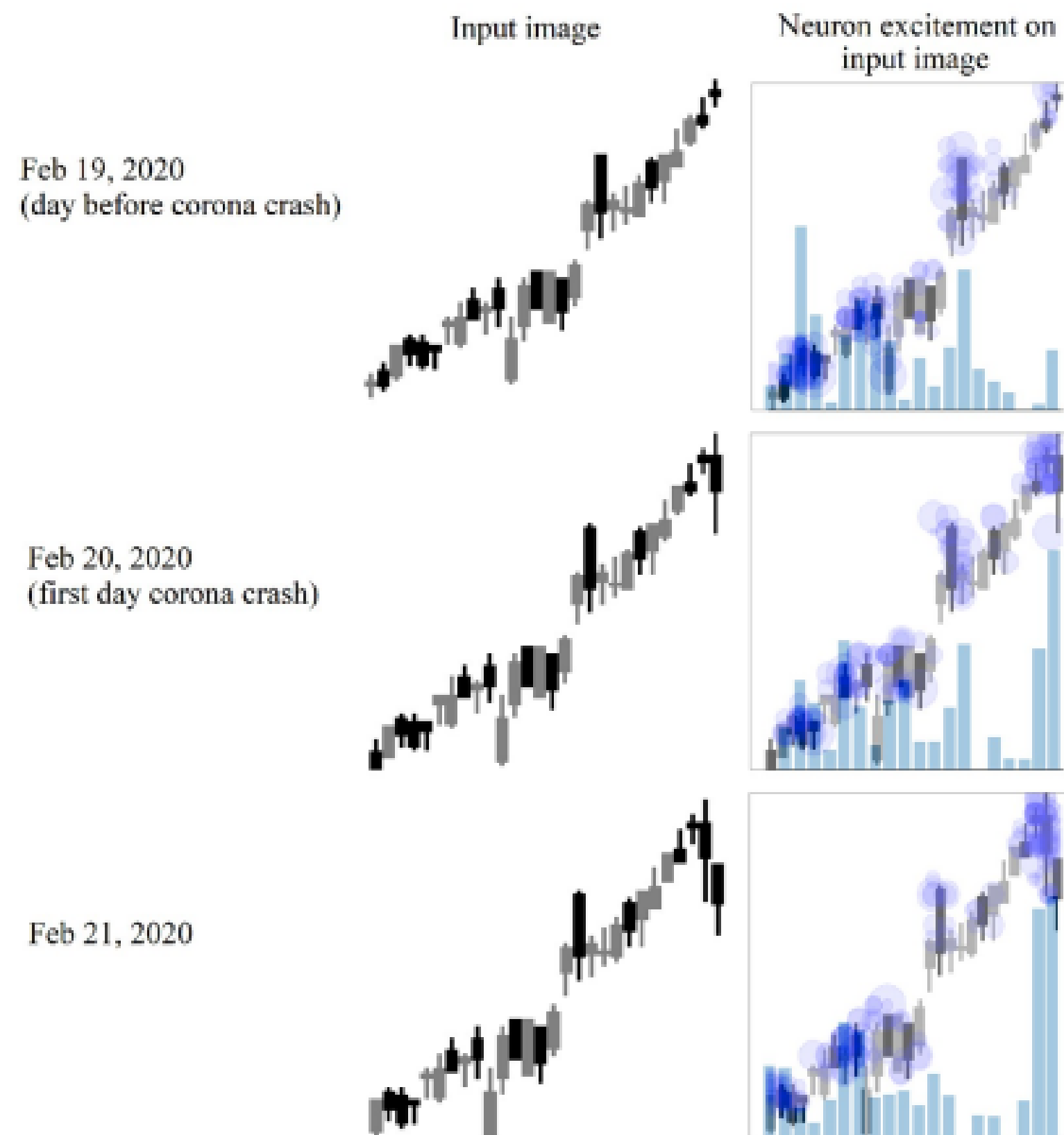
a：DDQN 輸出的動作， $\text{action} = \{1, -1, 0\}$ long、short、no position

r：每日收益

N：負獎勵

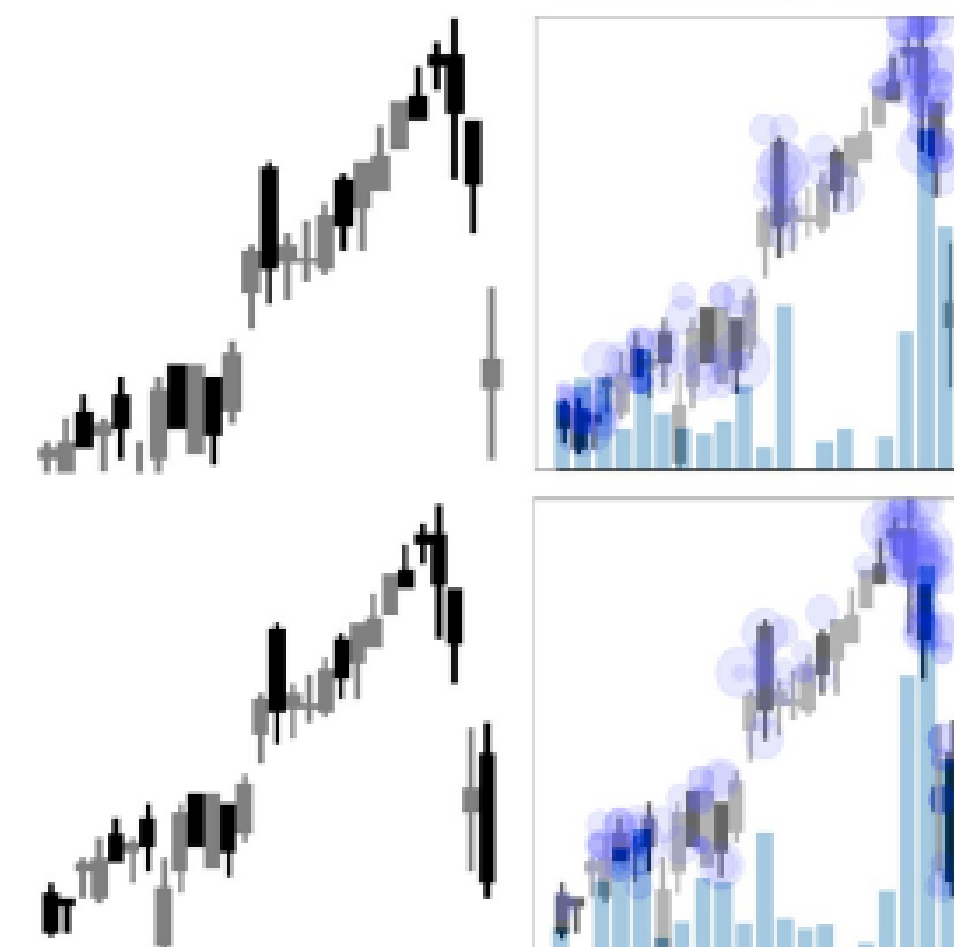
讓模型在多空不明的情況下，學習no position

資料視覺化:計算神經元權重

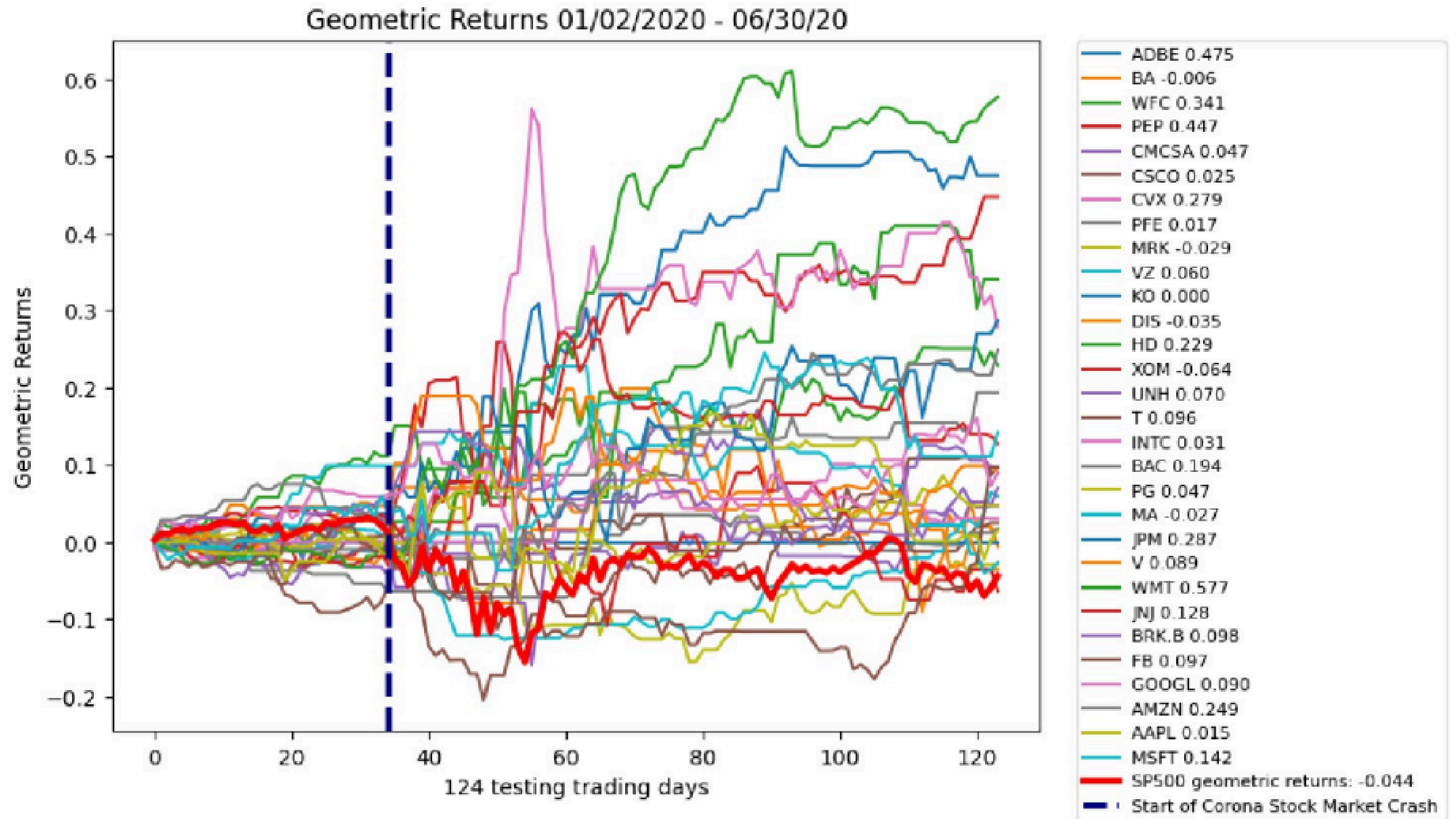


Feb 24, 2020

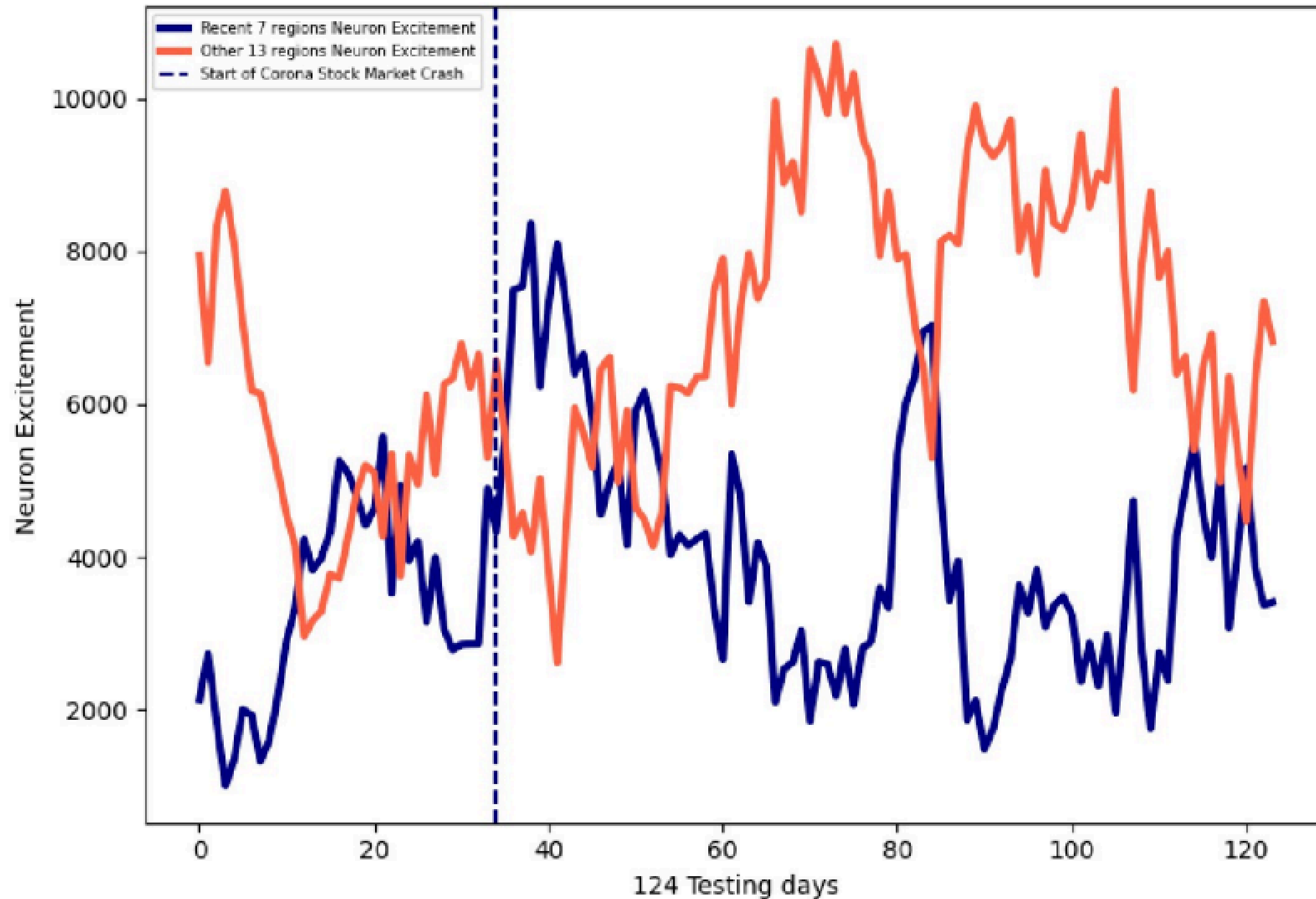
Feb 25, 2020



30個個股6個月Return



Adobe Neuron Excitement Shift to Recent Regions:
Sum of recent 7 regions neuron excitement and Sum of other 13 regions neuron excitement



A solid blue vertical bar is located on the far left side of the image, extending from the top to the bottom.

Thanks