



# 專題報告

題目：深度強化學習於股票之應用

**Deep Reinforcement for Stock Trading**


指導教授：陳煥

學生：陳俊元

學生：黃韻儒



# 目錄

- ◎ 專題目的與動機
  - ◎ 使用的技術
  - ◎ 實作方式
  - ◎ 實作結果
  - ◎ 實作困難
  - ◎ 總結
- 

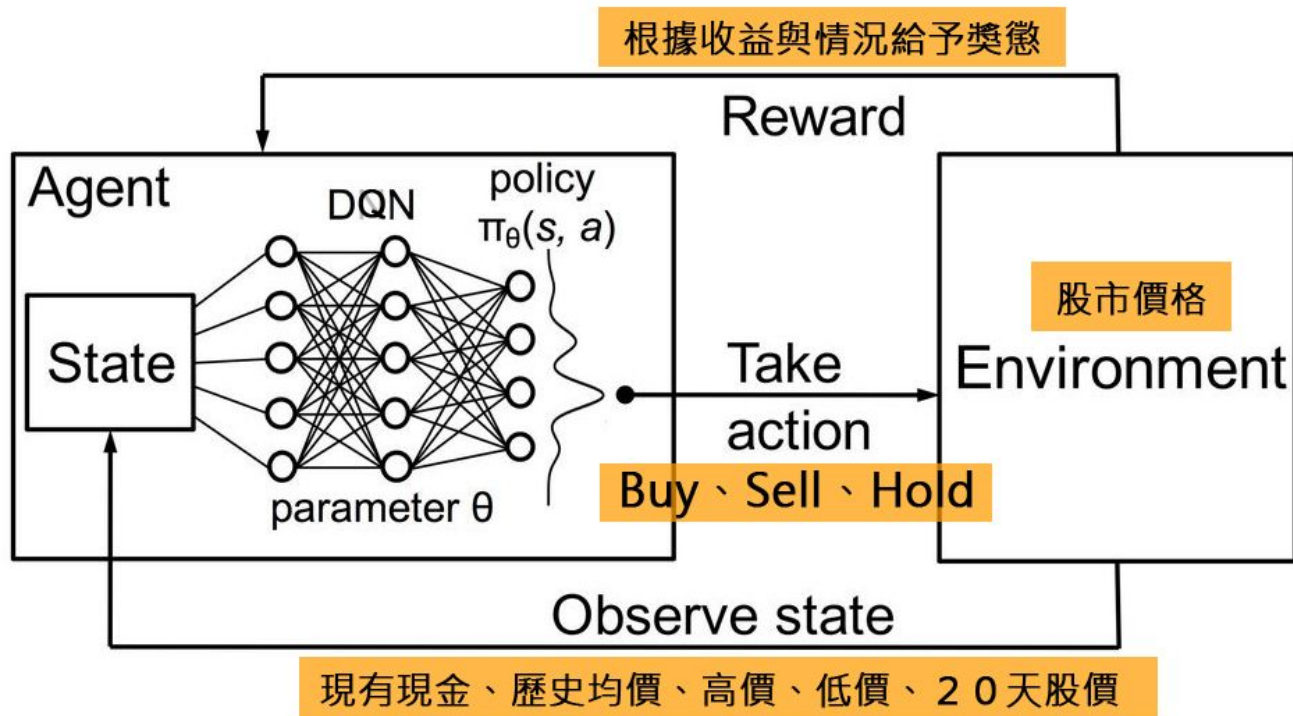
# 專題目的與動機

- ◎ 學習投資的基礎知識
- ◎ 熟悉RL的相關操作
- ◎ 尋找適當的股票策略
- ◎ 實驗各種基本股票指標

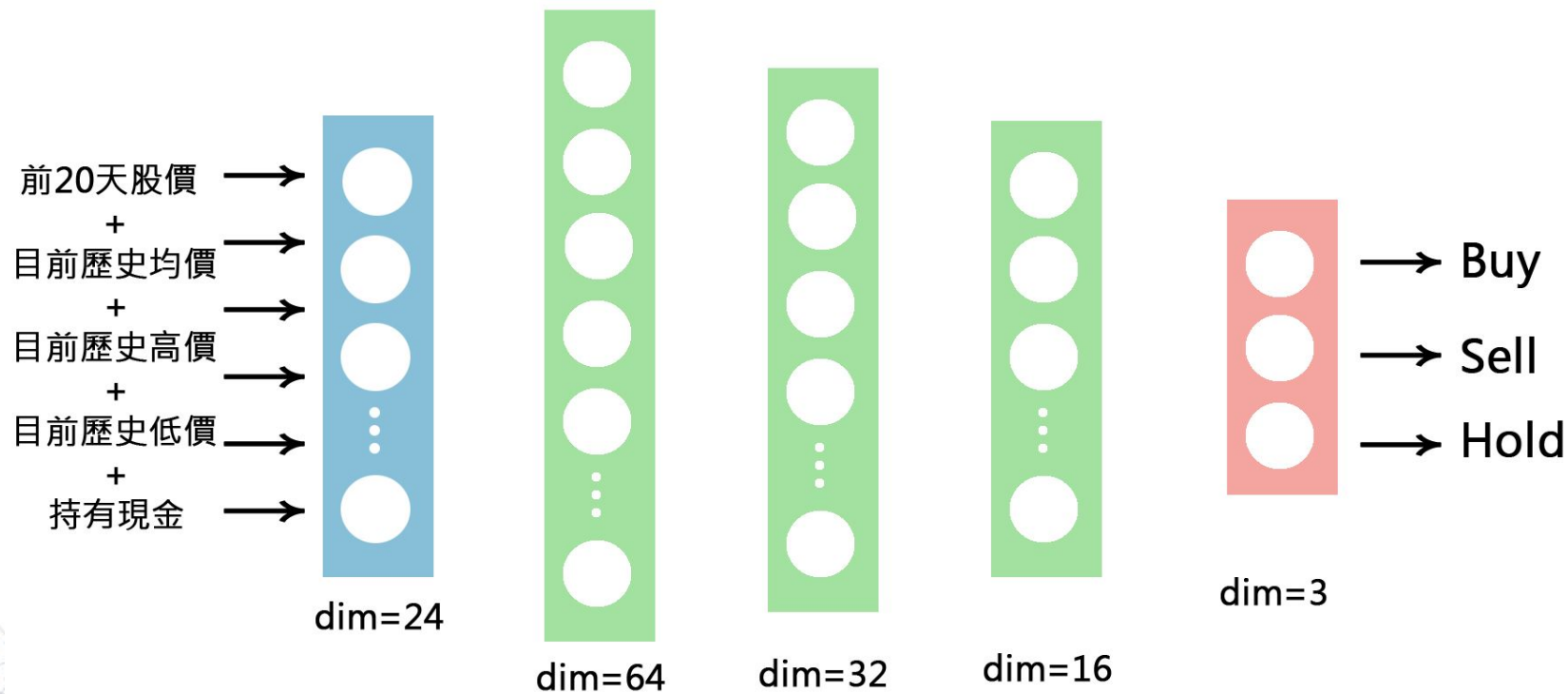


# 使用的技術

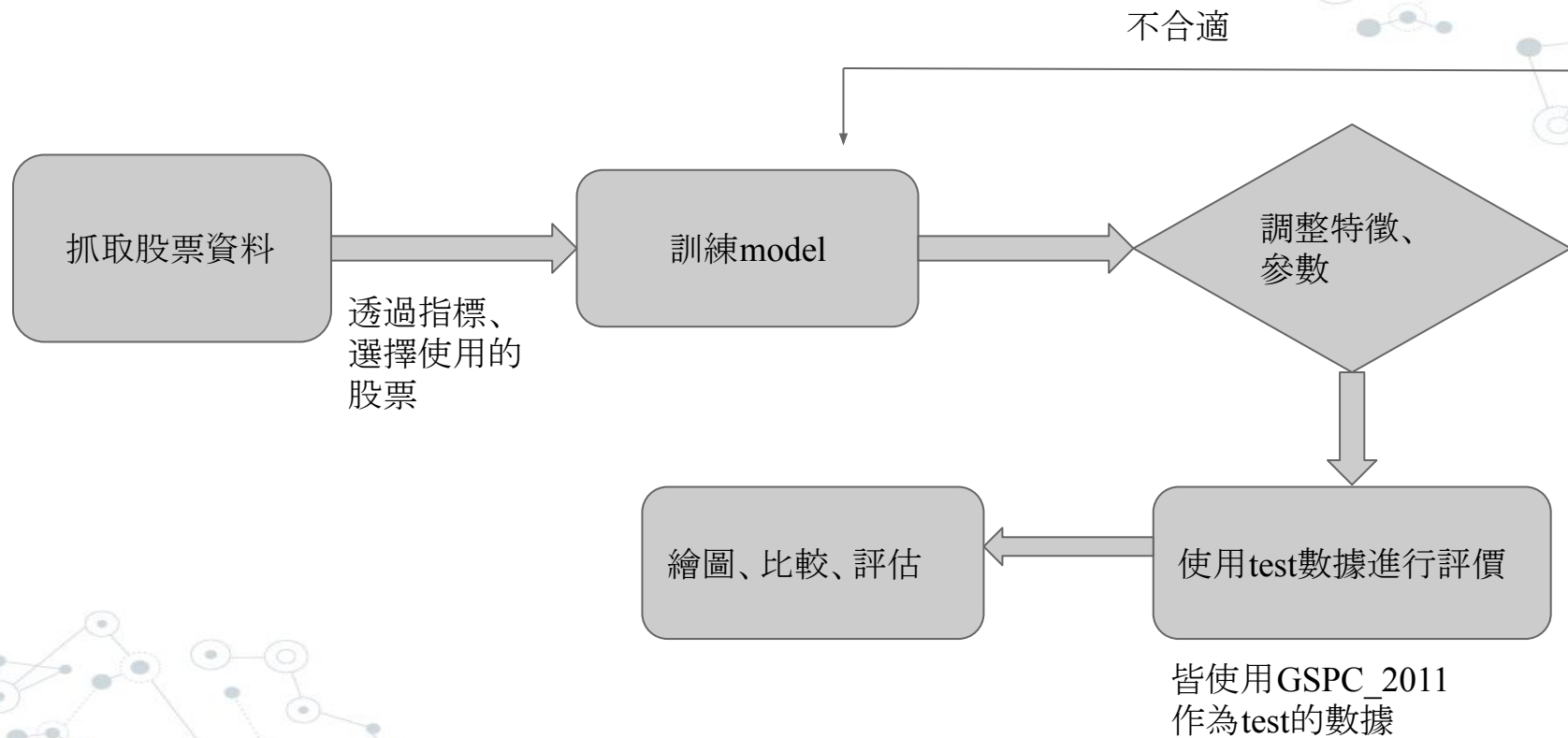
## Reinforcement Learning、Deep Q-learning



# model架構



# 實作方式



# 透過指標選擇股票



# 透過指標選擇訓練股票

找出時間序列較為穩定的股票-ADF test:

```
ADF Statistic: -0.560965  
p-value: 0.879561  
Critical Values:  
1%: -3.433  
5%: -2.863  
10%: -2.567
```

2001-2011(花仙子)

```
ADF Statistic: -1.959003  
p-value: 0.304833  
Critical Values:  
1%: -3.433  
5%: -2.863  
10%: -2.567
```

2001-2011(GSPC)

ADF-t 統計量越小越能提供證據證明序列穩定



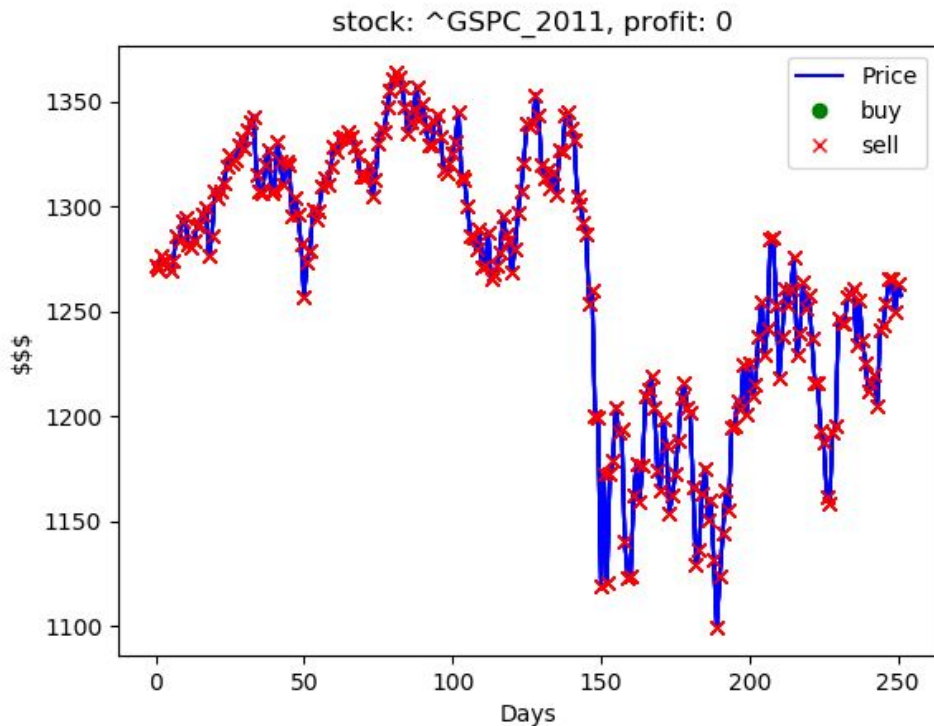
# 實作困難

## DQN中的Reward設置問題



# 實作困難

問題1. 僅以買賣盈虧值作為Reward, profit=0



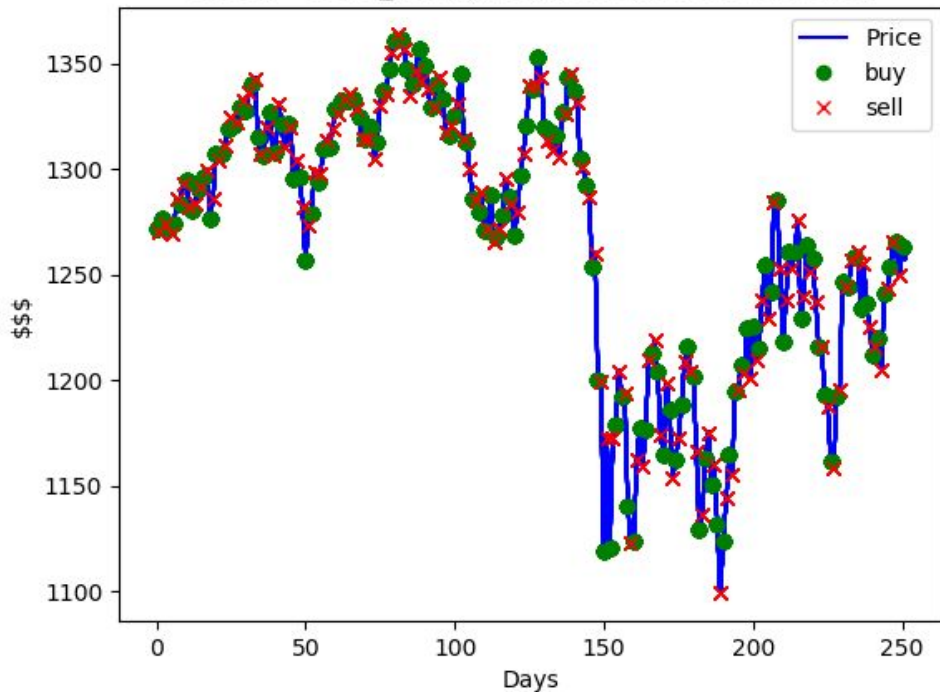
解法:

無持股卻要賣股時, 強迫Agent  
選擇其他行動 (Buy與hold擇其  
Action-Value 高者 )。

# 實作困難

問題2. 無持股時強迫Agent改變行動, 造成Agent為了賣而買

stock: ^GSPC\_2011, profit: 109.239618000000097

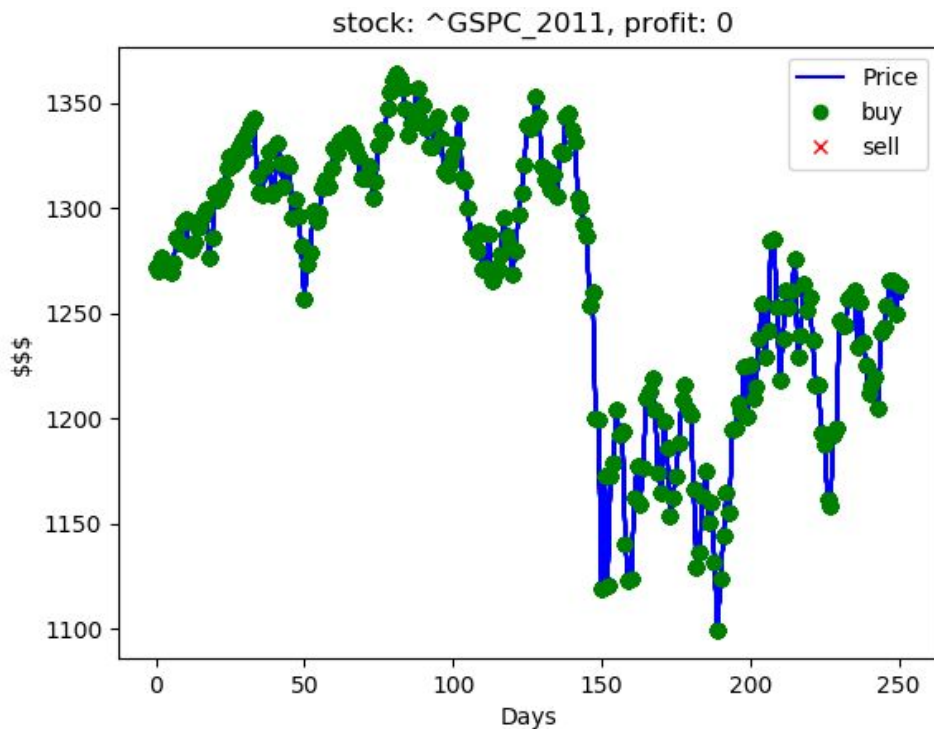


解法:

無持股卻要賣股時, 不更換其Sell  
行動, 改為給予負 reward 值。

# 實作困難

問題3. Sell有可能被懲罰, Buy卻沒有懲罰, 造成所有action都是buy

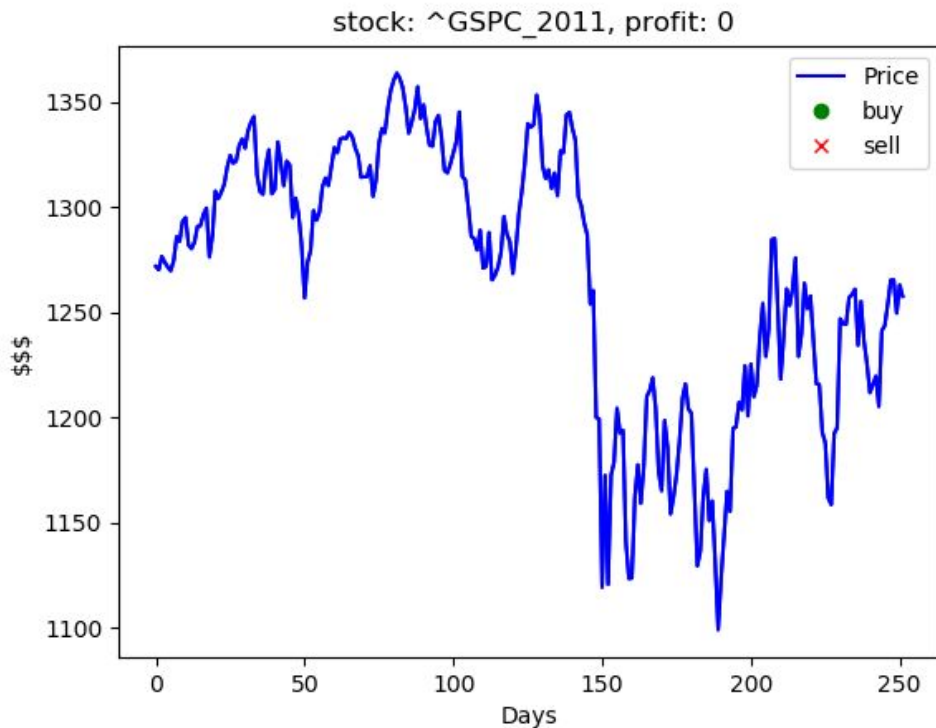


法:

限制資金, 沒有錢卻還是Buy時  
會給予懲罰

# 實作困難

問題4. 在sell與buy都有懲罰的情況下，機器選擇從頭到尾都hold(不買不賣)

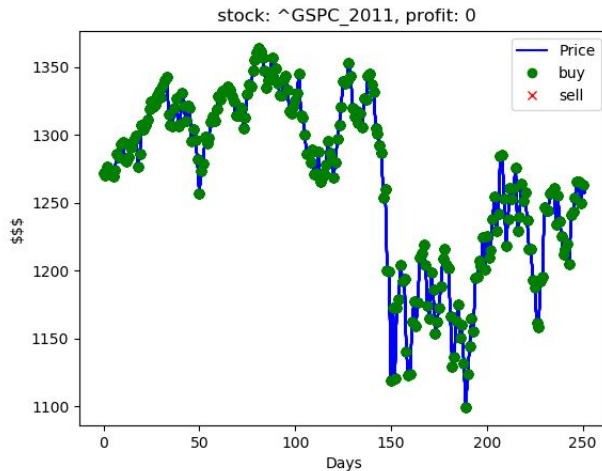
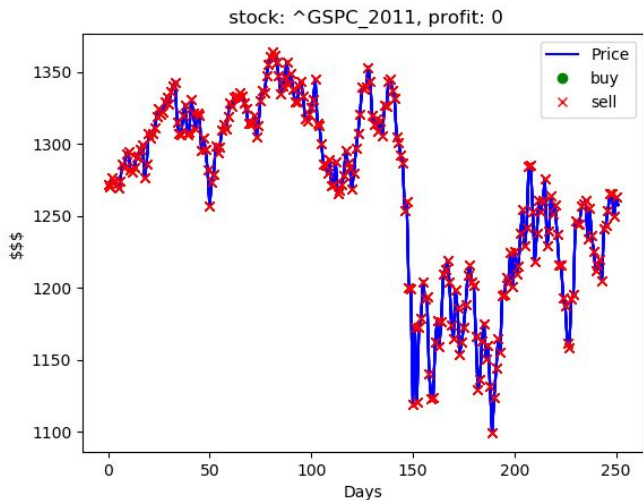


解法：

替hold加上懲罰，reward = -1

# 實作困難

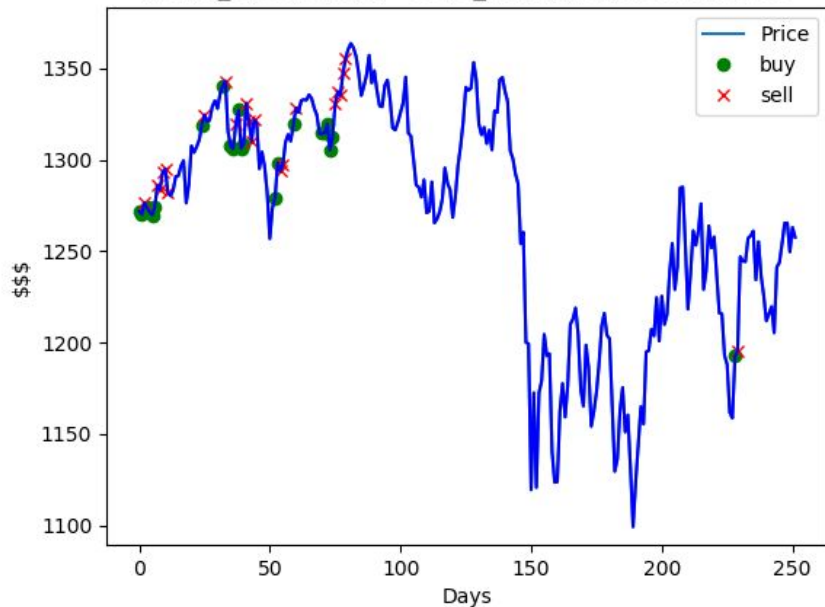
問題5. Reward的制定大幅影響決策，導致決策與股價趨勢無關



解法: 重新調整各結果的Reward比例

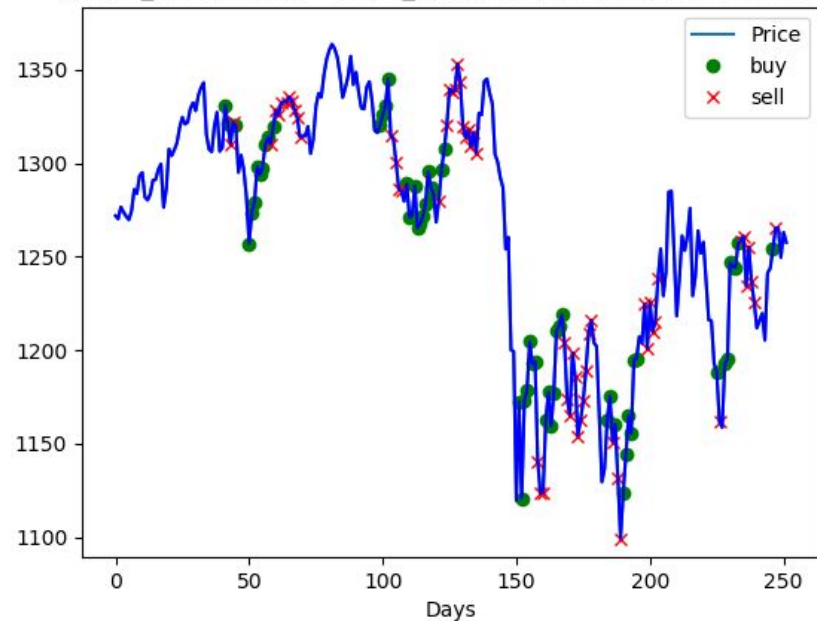
# 訓練結果

model\_ep60, stock: ^GSPC\_2011, profit: 306070.068



利用1730作為training  
data, PROFIT:0.3%

model\_ep60, stock: ^GSPC\_2011, profit: 828589.9630000005



利用GSPC作為training  
data, PROFIT:0.8%



## 總結

特徵的調整應該要用更客觀的方式來執行，例如使用逆向強化學習來調整特徵與reward。

買賣股票應依據比例調整股票的單位。



# 圖片來源

finlab部落

格:<https://www.finlab.tw/%E8%82%A1%E7%A5%A8%E5%85%A5%E9%96%80%E6%87%B6%E4%BA%BA%E5%8C%85/>

deep reinforcement learning:

<https://quantdare.com/deep-reinforcement-trading/>

Reinforcement Learning for the Enterprise:

<https://dzone.com/articles/reinforcement-learning-for-the-enterprise>

yahoo finance stock:

<https://finance.yahoo.com/quote/%5EGSPC/history?period1=1526491438&period2=1558027438&interval=1d&filter=history&frequency=1d>