

5.q-learning-agent

September 29, 2021

```
[1]: import numpy as np
import pandas as pd
import tensorflow as tf
import matplotlib.pyplot as plt
import seaborn as sns
sns.set()
```

```
[2]: df = pd.read_csv('../dataset/G00G-year.csv')
df.head()
```

```
[2]:
```

| | Date | Open | High | Low | Close | Adj Close | \ |
|---|------------|------------|------------|------------|------------|------------|---|
| 0 | 2016-11-02 | 778.200012 | 781.650024 | 763.450012 | 768.700012 | 768.700012 | |
| 1 | 2016-11-03 | 767.250000 | 769.950012 | 759.030029 | 762.130005 | 762.130005 | |
| 2 | 2016-11-04 | 750.659973 | 770.359985 | 750.560974 | 762.020020 | 762.020020 | |
| 3 | 2016-11-07 | 774.500000 | 785.190002 | 772.549988 | 782.520020 | 782.520020 | |
| 4 | 2016-11-08 | 783.400024 | 795.632996 | 780.190002 | 790.510010 | 790.510010 | |

| | Volume |
|---|---------|
| 0 | 1872400 |
| 1 | 1943200 |
| 2 | 2134800 |
| 3 | 1585100 |
| 4 | 1350800 |

```
[3]: from collections import deque
import random

class Agent:
    def __init__(self, state_size, window_size, trend, skip, batch_size):
        self.state_size = state_size
        self.window_size = window_size
        self.half_window = window_size // 2
        self.trend = trend
        self.skip = skip
        self.action_size = 3
        self.batch_size = batch_size
```

```

self.memory = deque(maxlen = 1000)
self.inventory = []

self.gamma = 0.95
self.epsilon = 0.5
self.epsilon_min = 0.01
self.epsilon_decay = 0.999

tf.reset_default_graph()
self.sess = tf.InteractiveSession()
self.X = tf.placeholder(tf.float32, [None, self.state_size])
self.Y = tf.placeholder(tf.float32, [None, self.action_size])
feed = tf.layers.dense(self.X, 256, activation = tf.nn.relu)
self.logits = tf.layers.dense(feed, self.action_size)
self.cost = tf.reduce_mean(tf.square(self.Y - self.logits))
self.optimizer = tf.train.GradientDescentOptimizer(1e-5).minimize(
    self.cost
)
self.sess.run(tf.global_variables_initializer())

def act(self, state):
    if random.random() <= self.epsilon:
        return random.randrange(self.action_size)
    return np.argmax(
        self.sess.run(self.logits, feed_dict = {self.X: state})[0]
    )

def get_state(self, t):
    window_size = self.window_size + 1
    d = t - window_size + 1
    block = self.trend[d : t + 1] if d >= 0 else -d * [self.trend[0]] +
↪self.trend[0 : t + 1]
    res = []
    for i in range(window_size - 1):
        res.append(block[i + 1] - block[i])
    return np.array([res])

def replay(self, batch_size):
    mini_batch = []
    l = len(self.memory)
    for i in range(l - batch_size, l):
        mini_batch.append(self.memory[i])
    replay_size = len(mini_batch)
    X = np.empty((replay_size, self.state_size))
    Y = np.empty((replay_size, self.action_size))
    states = np.array([a[0][0] for a in mini_batch])
    new_states = np.array([a[3][0] for a in mini_batch])

```

```

Q = self.sess.run(self.logits, feed_dict = {self.X: states})
Q_new = self.sess.run(self.logits, feed_dict = {self.X: new_states})
for i in range(len(mini_batch)):
    state, action, reward, next_state, done = mini_batch[i]
    target = Q[i]
    target[action] = reward
    if not done:
        target[action] += self.gamma * np.amax(Q_new[i])
    X[i] = state
    Y[i] = target
cost, _ = self.sess.run(
    [self.cost, self.optimizer], feed_dict = {self.X: X, self.Y: Y}
)
if self.epsilon > self.epsilon_min:
    self.epsilon *= self.epsilon_decay
return cost

def buy(self, initial_money):
    starting_money = initial_money
    states_sell = []
    states_buy = []
    inventory = []
    state = self.get_state(0)
    for t in range(0, len(self.trend) - 1, self.skip):
        action = self.act(state)
        next_state = self.get_state(t + 1)

        if action == 1 and initial_money >= self.trend[t] and t < (len(self.
→trend) - self.half_window):
            inventory.append(self.trend[t])
            initial_money -= self.trend[t]
            states_buy.append(t)
            print('day %d: buy 1 unit at price %f, total balance %f' % (t,
→self.trend[t], initial_money))

        elif action == 2 and len(inventory):
            bought_price = inventory.pop(0)
            initial_money += self.trend[t]
            states_sell.append(t)
            try:
                invest = ((close[t] - bought_price) / bought_price) * 100
            except:
                invest = 0
            print(
                'day %d, sell 1 unit at price %f, investment %f %%, total_
→balance %f,'

```

```

        % (t, close[t], invest, initial_money)
    )

    state = next_state
    invest = ((initial_money - starting_money) / starting_money) * 100
    total_gains = initial_money - starting_money
    return states_buy, states_sell, total_gains, invest

def train(self, iterations, checkpoint, initial_money):
    for i in range(iterations):
        total_profit = 0
        inventory = []
        state = self.get_state(0)
        starting_money = initial_money
        for t in range(0, len(self.trend) - 1, self.skip):
            action = self.act(state)
            next_state = self.get_state(t + 1)

            if action == 1 and starting_money >= self.trend[t] and t <
→ (len(self.trend) - self.half_window):
                inventory.append(self.trend[t])
                starting_money -= self.trend[t]

            elif action == 2 and len(inventory) > 0:
                bought_price = inventory.pop(0)
                total_profit += self.trend[t] - bought_price
                starting_money += self.trend[t]

            invest = ((starting_money - initial_money) / initial_money)
            self.memory.append((state, action, invest,
                                next_state, starting_money < initial_money))
            state = next_state
            batch_size = min(self.batch_size, len(self.memory))
            cost = self.replay(batch_size)
            if (i+1) % checkpoint == 0:
                print('epoch: %d, total rewards: %f.3, cost: %f, total money:
→ %f'%(i + 1, total_profit, cost,
→ starting_money))

```

```

[4]: close = df.Close.values.tolist()
initial_money = 10000
window_size = 30
skip = 1
batch_size = 32
agent = Agent(state_size = window_size,
              window_size = window_size,

```

```

        trend = close,
        skip = skip,
        batch_size = batch_size)
agent.train(iterations = 200, checkpoint = 10, initial_money = initial_money)

```

```

epoch: 10, total rewards: 274.710201.3, cost: 0.810730, total money:
10274.710201
epoch: 20, total rewards: 161.429929.3, cost: 0.406487, total money:
10161.429929
epoch: 30, total rewards: 89.659849.3, cost: 0.225568, total money: 10089.659849
epoch: 40, total rewards: 121.209836.3, cost: 0.152499, total money:
10121.209836
epoch: 50, total rewards: 94.869810.3, cost: 0.120762, total money: 10094.869810
epoch: 60, total rewards: 123.609922.3, cost: 0.097353, total money:
10123.609922
epoch: 70, total rewards: 130.149901.3, cost: 0.131718, total money:
10130.149901
epoch: 80, total rewards: 55.369871.3, cost: 0.072531, total money: 10055.369871
epoch: 90, total rewards: 177.780026.3, cost: 0.062346, total money:
10177.780026
epoch: 100, total rewards: 151.249997.3, cost: 0.056566, total money:
10151.249997
epoch: 110, total rewards: 101.629942.3, cost: 0.050717, total money:
10101.629942
epoch: 120, total rewards: 138.329892.3, cost: 0.075178, total money:
10138.329892
epoch: 130, total rewards: 187.559812.3, cost: 0.039170, total money:
10187.559812
epoch: 140, total rewards: 125.699889.3, cost: 0.035156, total money:
10125.699889
epoch: 150, total rewards: 138.249876.3, cost: 0.403965, total money:
10138.249876
epoch: 160, total rewards: 141.329832.3, cost: 0.029966, total money:
10141.329832
epoch: 170, total rewards: 179.989928.3, cost: 0.027219, total money:
10179.989928
epoch: 180, total rewards: 191.619871.3, cost: 0.025002, total money:
10191.619871
epoch: 190, total rewards: 191.929868.3, cost: 0.149151, total money:
10191.929868
epoch: 200, total rewards: 113.759886.3, cost: 0.021398, total money:
10113.759886

```

```

[5]: states_buy, states_sell, total_gains, invest = agent.buy(initial_money =
    ↪initial_money)

```

```

day 4: buy 1 unit at price 790.510010, total balance 9209.489990
day 5, sell 1 unit at price 785.309998, investment -0.657805 %, total balance

```

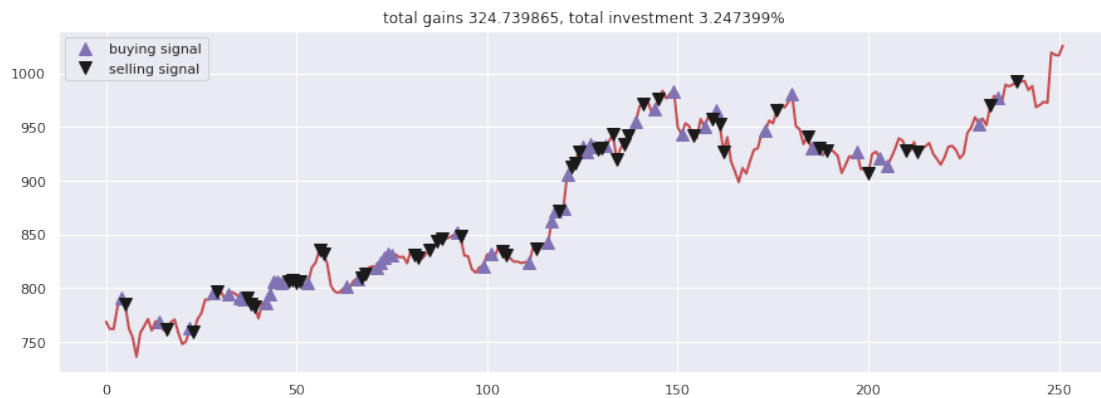
9994.799988,
 day 14: buy 1 unit at price 768.270020, total balance 9226.529968
 day 16, sell 1 unit at price 761.679993, investment -0.857775 %, total balance
 9988.209961,
 day 22: buy 1 unit at price 762.520020, total balance 9225.689941
 day 23, sell 1 unit at price 759.109985, investment -0.447206 %, total balance
 9984.799926,
 day 28: buy 1 unit at price 796.099976, total balance 9188.699950
 day 29, sell 1 unit at price 797.070007, investment 0.121848 %, total balance
 9985.769957,
 day 32: buy 1 unit at price 794.200012, total balance 9191.569945
 day 35: buy 1 unit at price 791.260010, total balance 8400.309935
 day 36: buy 1 unit at price 789.909973, total balance 7610.399962
 day 37, sell 1 unit at price 791.549988, investment -0.333672 %, total balance
 8401.949950,
 day 38, sell 1 unit at price 785.049988, investment -0.784827 %, total balance
 9186.999938,
 day 39, sell 1 unit at price 782.789978, investment -0.901368 %, total balance
 9969.789916,
 day 42: buy 1 unit at price 786.900024, total balance 9182.889892
 day 43: buy 1 unit at price 794.020020, total balance 8388.869872
 day 44: buy 1 unit at price 806.150024, total balance 7582.719848
 day 45: buy 1 unit at price 806.650024, total balance 6776.069824
 day 46: buy 1 unit at price 804.789978, total balance 5971.279846
 day 48, sell 1 unit at price 806.359985, investment 2.472990 %, total balance
 6777.639831,
 day 49, sell 1 unit at price 807.880005, investment 1.745546 %, total balance
 7585.519836,
 day 50, sell 1 unit at price 804.609985, investment -0.191036 %, total balance
 8390.129821,
 day 51, sell 1 unit at price 806.070007, investment -0.071904 %, total balance
 9196.199828,
 day 53: buy 1 unit at price 805.020020, total balance 8391.179808
 day 56, sell 1 unit at price 835.669983, investment 3.837027 %, total balance
 9226.849791,
 day 57, sell 1 unit at price 832.150024, investment 3.370103 %, total balance
 10058.999815,
 day 63: buy 1 unit at price 801.489990, total balance 9257.509825
 day 66: buy 1 unit at price 808.380005, total balance 8449.129820
 day 67, sell 1 unit at price 809.559998, investment 1.006876 %, total balance
 9258.689818,
 day 68, sell 1 unit at price 813.669983, investment 0.654392 %, total balance
 10072.359801,
 day 71: buy 1 unit at price 818.979980, total balance 9253.379821
 day 72: buy 1 unit at price 824.159973, total balance 8429.219848
 day 73: buy 1 unit at price 828.070007, total balance 7601.149841
 day 74: buy 1 unit at price 831.659973, total balance 6769.489868
 day 75: buy 1 unit at price 830.760010, total balance 5938.729858

day 81, sell 1 unit at price 830.630005, investment 1.422504 %, total balance 6769.359863,
 day 82, sell 1 unit at price 829.080017, investment 0.596977 %, total balance 7598.439880,
 day 85, sell 1 unit at price 835.369995, investment 0.881567 %, total balance 8433.809875,
 day 87, sell 1 unit at price 843.250000, investment 1.393602 %, total balance 9277.059875,
 day 88, sell 1 unit at price 845.539978, investment 1.779090 %, total balance 10122.599853,
 day 92: buy 1 unit at price 852.119995, total balance 9270.479858
 day 93, sell 1 unit at price 848.400024, investment -0.436555 %, total balance 10118.879882,
 day 99: buy 1 unit at price 820.919983, total balance 9297.959899
 day 101: buy 1 unit at price 831.500000, total balance 8466.459899
 day 104, sell 1 unit at price 834.570007, investment 1.662772 %, total balance 9301.029906,
 day 105, sell 1 unit at price 831.409973, investment -0.010827 %, total balance 10132.439879,
 day 111: buy 1 unit at price 823.559998, total balance 9308.879881
 day 113, sell 1 unit at price 836.820007, investment 1.610084 %, total balance 10145.699888,
 day 116: buy 1 unit at price 843.190002, total balance 9302.509886
 day 117: buy 1 unit at price 862.760010, total balance 8439.749876
 day 118: buy 1 unit at price 872.299988, total balance 7567.449888
 day 119, sell 1 unit at price 871.729980, investment 3.384762 %, total balance 8439.179868,
 day 120: buy 1 unit at price 874.250000, total balance 7564.929868
 day 121: buy 1 unit at price 905.960022, total balance 6658.969846
 day 122, sell 1 unit at price 912.570007, investment 5.773332 %, total balance 7571.539853,
 day 123, sell 1 unit at price 916.440002, investment 5.060187 %, total balance 8487.979855,
 day 124, sell 1 unit at price 927.039978, investment 6.038316 %, total balance 9415.019833,
 day 125: buy 1 unit at price 931.659973, total balance 8483.359860
 day 126: buy 1 unit at price 927.130005, total balance 7556.229855
 day 127: buy 1 unit at price 934.299988, total balance 6621.929867
 day 128: buy 1 unit at price 932.169983, total balance 5689.759884
 day 129, sell 1 unit at price 928.780029, investment 2.518876 %, total balance 6618.539913,
 day 130, sell 1 unit at price 930.599976, investment -0.113775 %, total balance 7549.139889,
 day 131: buy 1 unit at price 932.219971, total balance 6616.919918
 day 133, sell 1 unit at price 943.000000, investment 1.711734 %, total balance 7559.919918,
 day 134, sell 1 unit at price 919.619995, investment -1.571229 %, total balance 8479.539913,

day 136, sell 1 unit at price 934.010010, investment 0.197392 %, total balance 9413.549923,
 day 137, sell 1 unit at price 941.859985, investment 1.034092 %, total balance 10355.409908,
 day 139: buy 1 unit at price 954.960022, total balance 9400.449886
 day 141, sell 1 unit at price 971.469971, investment 1.728863 %, total balance 10371.919857,
 day 144: buy 1 unit at price 966.950012, total balance 9404.969845
 day 145, sell 1 unit at price 975.599976, investment 0.894562 %, total balance 10380.569821,
 day 149: buy 1 unit at price 983.409973, total balance 9397.159848
 day 151: buy 1 unit at price 942.900024, total balance 8454.259824
 day 154, sell 1 unit at price 942.309998, investment -4.179333 %, total balance 9396.569822,
 day 157: buy 1 unit at price 950.630005, total balance 8445.939817
 day 159, sell 1 unit at price 957.090027, investment 1.504932 %, total balance 9403.029844,
 day 160: buy 1 unit at price 965.590027, total balance 8437.439817
 day 161, sell 1 unit at price 952.270020, investment 0.172519 %, total balance 9389.709837,
 day 162, sell 1 unit at price 927.330017, investment -3.962345 %, total balance 10317.039854,
 day 173: buy 1 unit at price 947.159973, total balance 9369.879881
 day 176, sell 1 unit at price 965.400024, investment 1.925762 %, total balance 10335.279905,
 day 180: buy 1 unit at price 980.340027, total balance 9354.939878
 day 184, sell 1 unit at price 941.530029, investment -3.958830 %, total balance 10296.469907,
 day 185: buy 1 unit at price 930.500000, total balance 9365.969907
 day 186: buy 1 unit at price 930.830017, total balance 8435.139890
 day 187, sell 1 unit at price 930.390015, investment -0.011820 %, total balance 9365.529905,
 day 189, sell 1 unit at price 927.960022, investment -0.308326 %, total balance 10293.489927,
 day 197: buy 1 unit at price 926.960022, total balance 9366.529905
 day 200, sell 1 unit at price 906.659973, investment -2.189959 %, total balance 10273.189878,
 day 203: buy 1 unit at price 921.280029, total balance 9351.909849
 day 205: buy 1 unit at price 913.809998, total balance 8438.099851
 day 210, sell 1 unit at price 928.450012, investment 0.778263 %, total balance 9366.549863,
 day 213, sell 1 unit at price 926.500000, investment 1.388692 %, total balance 10293.049863,
 day 229: buy 1 unit at price 953.270020, total balance 9339.779843
 day 232, sell 1 unit at price 969.960022, investment 1.750816 %, total balance 10309.739865,
 day 234: buy 1 unit at price 977.000000, total balance 9332.739865
 day 239, sell 1 unit at price 992.000000, investment 1.535312 %, total balance

10324.739865,

```
[6]: fig = plt.figure(figsize = (15,5))
plt.plot(close, color='r', lw=2.)
plt.plot(close, '^', markersize=10, color='m', label = 'buying signal',
↪markevery = states_buy)
plt.plot(close, 'v', markersize=10, color='k', label = 'selling signal',
↪markevery = states_sell)
plt.title('total gains %f, total investment %f%%'%(total_gains, invest))
plt.legend()
plt.show()
```



[]: