4.policy-gradient-agent

September 29, 2021

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

```python
[2]: df = pd.read_csv('../dataset/GOOG-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
```

```python
[3]: class Agent:

         LEARNING_RATE = 1e-4
         LAYER_SIZE = 256
         GAMMA = 0.9
         OUTPUT_SIZE = 3

         def __init__(self, state_size, window_size, trend, skip):
             self.state_size = state_size
             self.window_size = window_size
             self.half_window = window_size // 2
             self.trend = trend
             self.skip = skip
```

```python
        self.X = tf.placeholder(tf.float32, (None, self.state_size))
        self.REWARDS = tf.placeholder(tf.float32, (None))
        self.ACTIONS = tf.placeholder(tf.int32, (None))
        feed_forward = tf.layers.dense(self.X, self.LAYER_SIZE, activation = tf.
↪nn.relu)
        self.logits = tf.layers.dense(feed_forward, self.OUTPUT_SIZE,␣
↪activation = tf.nn.softmax)
        input_y = tf.one_hot(self.ACTIONS, self.OUTPUT_SIZE)
        loglike = tf.log((input_y * (input_y - self.logits) + (1 - input_y) *␣
↪(input_y + self.logits)) + 1)
        rewards = tf.tile(tf.reshape(self.REWARDS, (-1,1)), [1, self.
↪OUTPUT_SIZE])
        self.cost = -tf.reduce_mean(loglike * (rewards + 1))
        self.optimizer = tf.train.AdamOptimizer(learning_rate = self.
↪LEARNING_RATE).minimize(self.cost)
        self.sess = tf.InteractiveSession()
        self.sess.run(tf.global_variables_initializer())

    def predict(self, inputs):
        return self.sess.run(self.logits, feed_dict={self.X:inputs})

    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 discount_rewards(self, r):
        discounted_r = np.zeros_like(r)
        running_add = 0
        for t in reversed(range(0, r.size)):
            running_add = running_add * self.GAMMA + r[t]
            discounted_r[t] = running_add
        return discounted_r

    def get_predicted_action(self, sequence):
        prediction = self.predict(np.array(sequence))[0]
        return np.argmax(prediction)

    def buy(self, initial_money):
        starting_money = initial_money
        states_sell = []
```

```python
        states_buy = []
        inventory = []
        state = self.get_state(0)
        for t in range(0, len(self.trend) - 1, self.skip):
            action = self.get_predicted_action(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):
            ep_history = []
            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.get_predicted_action(state)
                next_state = self.get_state(t + 1)
```

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

                elif action == 2 and len(inventory):
                    bought_price = inventory.pop(0)
                    total_profit += self.trend[t] - bought_price
                    starting_money += self.trend[t]
                ep_history.append([state,action,starting_money,next_state])
                state = next_state
            ep_history = np.array(ep_history)
            ep_history[:,2] = self.discount_rewards(ep_history[:,2])
            cost, _ = self.sess.run([self.cost, self.optimizer], feed_dict={self.X:np.vstack(ep_history[:,0]),
                                                                            self.REWARDS:ep_history[:,2],
                                                                            self.ACTIONS:ep_history[:,1]})
            if (i+1) % checkpoint == 0:
                print('epoch: %d, total rewards: %f.3, cost: %f, total money: %f'%(i + 1, total_profit, cost,
                    starting_money))
```

```python
[4]: close = df.Close.values.tolist()
    initial_money = 10000
    window_size = 30
    skip = 1
    agent = Agent(state_size = window_size,
              window_size = window_size,
              trend = close,
              skip = skip)
    agent.train(iterations = 200, checkpoint = 10, initial_money = initial_money)
```

```
epoch: 10, total rewards: 1781.590144.3, cost: -3782.833740, total money:
7062.900203
epoch: 20, total rewards: 1808.720396.3, cost: -6238.727539, total money:
10819.470396
epoch: 30, total rewards: 644.675288.3, cost: -10399.220703, total money:
10644.675288
epoch: 40, total rewards: 1696.944943.3, cost: -9798.079102, total money:
11696.944943
epoch: 50, total rewards: 593.719845.3, cost: -13938.982422, total money:
10593.719845
epoch: 60, total rewards: 634.539913.3, cost: -14890.398438, total money:
9645.289913
```

```
epoch: 70, total rewards: 1586.160156.3, cost: -10411.115234, total money:
11586.160156
epoch: 80, total rewards: 368.749937.3, cost: -15986.910156, total money:
10368.749937
epoch: 90, total rewards: 733.844603.3, cost: -15352.789062, total money:
8857.304625
epoch: 100, total rewards: 645.715148.3, cost: -15976.339844, total money:
10645.715148
epoch: 110, total rewards: 994.814937.3, cost: -11198.958984, total money:
4471.054988
epoch: 120, total rewards: 1771.289852.3, cost: -6539.313477, total money:
5164.829891
epoch: 130, total rewards: 1643.744995.3, cost: -11630.438477, total money:
11643.744995
epoch: 140, total rewards: 1877.095029.3, cost: -7103.230957, total money:
9104.255063
epoch: 150, total rewards: 481.749932.3, cost: -18531.593750, total money:
10481.749932
epoch: 160, total rewards: 638.035152.3, cost: -16995.314453, total money:
10638.035152
epoch: 170, total rewards: 1188.049925.3, cost: -13348.065430, total money:
10263.189940
epoch: 180, total rewards: 633.885008.3, cost: -14666.952148, total money:
10633.885008
epoch: 190, total rewards: 1675.079952.3, cost: -9106.298828, total money:
5977.189998
epoch: 200, total rewards: 567.955136.3, cost: -17828.587891, total money:
10567.955136
```

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

```
day 12: buy 1 unit at price 760.539978, total balance 9239.460022
day 19, sell 1 unit at price 758.039978, investment -0.328714 %, total balance
9997.500000,
day 23: buy 1 unit at price 759.109985, total balance 9238.390015
day 26, sell 1 unit at price 789.289978, investment 3.975708 %, total balance
10027.679993,
day 27: buy 1 unit at price 789.270020, total balance 9238.409973
day 28, sell 1 unit at price 796.099976, investment 0.865351 %, total balance
10034.509949,
day 31: buy 1 unit at price 790.799988, total balance 9243.709961
day 33, sell 1 unit at price 796.419983, investment 0.710672 %, total balance
10040.129944,
day 35: buy 1 unit at price 791.260010, total balance 9248.869934
day 36: buy 1 unit at price 789.909973, total balance 8458.959961
day 37, sell 1 unit at price 791.549988, investment 0.036648 %, total balance
9250.509949,
```

```
day 38: buy 1 unit at price 785.049988, total balance 8465.459961
day 39: buy 1 unit at price 782.789978, total balance 7682.669983
day 40: buy 1 unit at price 771.820007, total balance 6910.849976
day 41, sell 1 unit at price 786.140015, investment -0.477264 %, total balance
7696.989991,
day 42, sell 1 unit at price 786.900024, investment 0.235658 %, total balance
8483.890015,
day 44: buy 1 unit at price 806.150024, total balance 7677.739991
day 46: buy 1 unit at price 804.789978, total balance 6872.950013
day 50: buy 1 unit at price 804.609985, total balance 6068.340028
day 51: buy 1 unit at price 806.070007, total balance 5262.270021
day 52, sell 1 unit at price 802.174988, investment 2.476400 %, total balance
6064.445009,
day 53, sell 1 unit at price 805.020020, investment 4.301523 %, total balance
6869.465029,
day 54, sell 1 unit at price 819.309998, investment 1.632447 %, total balance
7688.775027,
day 56: buy 1 unit at price 835.669983, total balance 6853.105044
day 57: buy 1 unit at price 832.150024, total balance 6020.955020
day 59, sell 1 unit at price 802.320007, investment -0.306909 %, total balance
6823.275027,
day 61: buy 1 unit at price 795.695007, total balance 6027.580020
day 63: buy 1 unit at price 801.489990, total balance 5226.090030
day 64, sell 1 unit at price 801.340027, investment -0.406403 %, total balance
6027.430057,
day 65: buy 1 unit at price 806.969971, total balance 5220.460086
day 71: buy 1 unit at price 818.979980, total balance 4401.480106
day 72: buy 1 unit at price 824.159973, total balance 3577.320133
day 73, sell 1 unit at price 828.070007, investment 2.729291 %, total balance
4405.390140,
day 74, sell 1 unit at price 831.659973, investment -0.479856 %, total balance
5237.050113,
day 78, sell 1 unit at price 829.280029, investment -0.344889 %, total balance
6066.330142,
day 79: buy 1 unit at price 823.210022, total balance 5243.120120
day 80: buy 1 unit at price 835.239990, total balance 4407.880130
day 81, sell 1 unit at price 830.630005, investment 4.390501 %, total balance
5238.510135,
day 83: buy 1 unit at price 827.780029, total balance 4410.730106
day 86, sell 1 unit at price 838.679993, investment 4.640108 %, total balance
5249.410099,
day 88, sell 1 unit at price 845.539978, investment 4.779609 %, total balance
6094.950077,
day 89, sell 1 unit at price 845.619995, investment 3.252829 %, total balance
6940.570072,
day 90: buy 1 unit at price 847.200012, total balance 6093.370060
day 92, sell 1 unit at price 852.119995, investment 3.392548 %, total balance
6945.490055,
```

day 93, sell 1 unit at price 848.400024, investment 3.059973 %, total balance 7793.890079,
day 94: buy 1 unit at price 830.460022, total balance 6963.430057
day 95, sell 1 unit at price 829.590027, investment -0.676448 %, total balance 7793.020084,
day 98: buy 1 unit at price 819.510010, total balance 6973.510074
day 101: buy 1 unit at price 831.500000, total balance 6142.010074
day 102: buy 1 unit at price 829.559998, total balance 5312.450076
day 103, sell 1 unit at price 838.549988, investment 1.301065 %, total balance 6151.000064,
day 105, sell 1 unit at price 831.409973, investment -1.863791 %, total balance 6982.410037,
day 106: buy 1 unit at price 827.880005, total balance 6154.530032
day 108, sell 1 unit at price 824.729980, investment -0.689984 %, total balance 6979.260012,
day 109: buy 1 unit at price 823.349976, total balance 6155.910036
day 110, sell 1 unit at price 824.320007, investment 0.586936 %, total balance 6980.230043,
day 111, sell 1 unit at price 823.559998, investment -0.954901 %, total balance 7803.790041,
day 113, sell 1 unit at price 836.820007, investment 0.875164 %, total balance 8640.610048,
day 114: buy 1 unit at price 838.210022, total balance 7802.400026
day 115: buy 1 unit at price 841.650024, total balance 6960.750002
day 116, sell 1 unit at price 843.190002, investment 1.849301 %, total balance 7803.940004,
day 117, sell 1 unit at price 862.760010, investment 4.786547 %, total balance 8666.700014,
day 118, sell 1 unit at price 872.299988, investment 4.066996 %, total balance 9539.000002,
day 119: buy 1 unit at price 871.729980, total balance 8667.270022
day 120, sell 1 unit at price 874.250000, investment 3.873341 %, total balance 9541.520022,
day 121: buy 1 unit at price 905.960022, total balance 8635.560000
day 123, sell 1 unit at price 916.440002, investment 5.128884 %, total balance 9552.000002,
day 124: buy 1 unit at price 927.039978, total balance 8624.960024
day 127: buy 1 unit at price 934.299988, total balance 7690.660036
day 128: buy 1 unit at price 932.169983, total balance 6758.490053
day 129, sell 1 unit at price 928.780029, investment 2.518876 %, total balance 7687.270082,
day 130: buy 1 unit at price 930.599976, total balance 6756.670106
day 131: buy 1 unit at price 932.219971, total balance 5824.450135
day 132: buy 1 unit at price 937.080017, total balance 4887.370118
day 133: buy 1 unit at price 943.000000, total balance 3944.370118
day 134, sell 1 unit at price 919.619995, investment -0.800395 %, total balance 4863.990113,
day 135: buy 1 unit at price 930.239990, total balance 3933.750123

day 136: buy 1 unit at price 934.010010, total balance 2999.740113
day 137, sell 1 unit at price 941.859985, investment 0.809162 %, total balance 3941.600098,
day 138, sell 1 unit at price 948.820007, investment 1.786157 %, total balance 4890.420105,
day 139: buy 1 unit at price 954.960022, total balance 3935.460083
day 140, sell 1 unit at price 969.539978, investment 4.184397 %, total balance 4905.000061,
day 141: buy 1 unit at price 971.469971, total balance 3933.530090
day 142: buy 1 unit at price 975.880005, total balance 2957.650085
day 143, sell 1 unit at price 964.859985, investment 3.501321 %, total balance 3922.510070,
day 145: buy 1 unit at price 975.599976, total balance 2946.910094
day 146, sell 1 unit at price 983.679993, investment 4.972892 %, total balance 3930.590087,
day 147: buy 1 unit at price 976.570007, total balance 2954.020080
day 148, sell 1 unit at price 980.940002, investment 4.023330 %, total balance 3934.960082,
day 149, sell 1 unit at price 983.409973, investment 5.715728 %, total balance 4918.370055,
day 150: buy 1 unit at price 949.830017, total balance 3968.540038
day 151: buy 1 unit at price 942.900024, total balance 3025.640014
day 153, sell 1 unit at price 950.760010, investment 1.793343 %, total balance 3976.400024,
day 154: buy 1 unit at price 942.309998, total balance 3034.090026
day 155: buy 1 unit at price 939.780029, total balance 2094.309997
day 156: buy 1 unit at price 957.369995, total balance 1136.940002
day 157, sell 1 unit at price 950.630005, investment -0.453424 %, total balance 2087.570007,
day 158: buy 1 unit at price 959.450012, total balance 1128.119995
day 159: buy 1 unit at price 957.090027, total balance 171.029968
day 161, sell 1 unit at price 952.270020, investment -1.976381 %, total balance 1123.299988,
day 164, sell 1 unit at price 917.789978, investment -5.952579 %, total balance 2041.089966,
day 165, sell 1 unit at price 908.729980, investment -6.854243 %, total balance 2949.819946,
day 166: buy 1 unit at price 898.700012, total balance 2051.119934
day 168, sell 1 unit at price 906.690002, investment -7.155658 %, total balance 2957.809936,
day 170: buy 1 unit at price 928.799988, total balance 2029.009948
day 172, sell 1 unit at price 943.830017, investment -0.631692 %, total balance 2972.839965,
day 173, sell 1 unit at price 947.159973, investment 0.451792 %, total balance 3919.999938,
day 175: buy 1 unit at price 953.419983, total balance 2966.579955
day 176, sell 1 unit at price 965.400024, investment 2.450364 %, total balance 3931.979979,
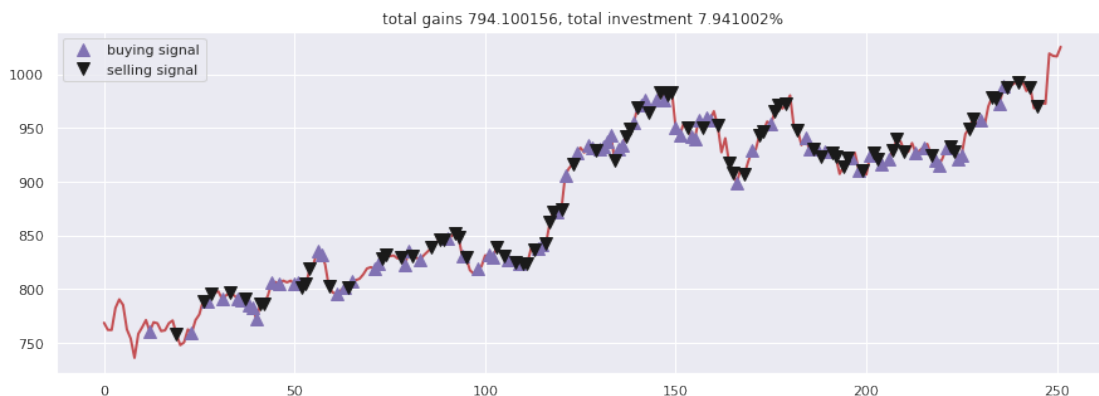
day 177, sell 1 unit at price 970.890015, investment 3.310348 %, total balance 4902.869994,
day 179, sell 1 unit at price 972.919983, investment 1.624240 %, total balance 5875.789977,
day 182, sell 1 unit at price 947.799988, investment -1.214240 %, total balance 6823.589965,
day 184: buy 1 unit at price 941.530029, total balance 5882.059936
day 185: buy 1 unit at price 930.500000, total balance 4951.559936
day 186, sell 1 unit at price 930.830017, investment -2.743735 %, total balance 5882.389953,
day 187: buy 1 unit at price 930.390015, total balance 4951.999938
day 188, sell 1 unit at price 923.650024, investment 2.776234 %, total balance 5875.649962,
day 189: buy 1 unit at price 927.960022, total balance 4947.689940
day 191, sell 1 unit at price 926.789978, investment -0.216409 %, total balance 5874.479918,
day 192, sell 1 unit at price 922.900024, investment -3.201103 %, total balance 6797.379942,
day 194, sell 1 unit at price 914.390015, investment -2.882544 %, total balance 7711.769957,
day 195, sell 1 unit at price 922.669983, investment -0.841485 %, total balance 8634.439940,
day 196: buy 1 unit at price 922.219971, total balance 7712.219969
day 198: buy 1 unit at price 910.979980, total balance 6801.239989
day 199, sell 1 unit at price 910.669983, investment -2.119545 %, total balance 7711.909972,
day 201: buy 1 unit at price 924.690002, total balance 6787.219970
day 202, sell 1 unit at price 927.000000, investment -0.103455 %, total balance 7714.219970,
day 203, sell 1 unit at price 921.280029, investment -0.101922 %, total balance 8635.499999,
day 204: buy 1 unit at price 915.890015, total balance 7719.609984
day 206: buy 1 unit at price 921.289978, total balance 6798.320006
day 207, sell 1 unit at price 929.570007, investment 2.040663 %, total balance 7727.890013,
day 208, sell 1 unit at price 939.330017, investment 1.583235 %, total balance 8667.220030,
day 210, sell 1 unit at price 928.450012, investment 1.371343 %, total balance 9595.670042,
day 213: buy 1 unit at price 926.500000, total balance 8669.170042
day 215: buy 1 unit at price 932.070007, total balance 7737.100035
day 217, sell 1 unit at price 925.109985, investment 0.414637 %, total balance 8662.210020,
day 218: buy 1 unit at price 920.289978, total balance 7741.920042
day 219: buy 1 unit at price 915.000000, total balance 6826.920042
day 221: buy 1 unit at price 931.580017, total balance 5895.340025
day 222, sell 1 unit at price 932.450012, investment 0.642203 %, total balance 6827.790037,

day 223, sell 1 unit at price 928.530029, investment -0.379797 %, total balance 7756.320066,
day 224: buy 1 unit at price 920.969971, total balance 6835.350095
day 225: buy 1 unit at price 924.859985, total balance 5910.490110
day 227, sell 1 unit at price 949.500000, investment 3.174002 %, total balance 6859.990110,
day 228, sell 1 unit at price 959.109985, investment 4.820763 %, total balance 7819.100095,
day 230: buy 1 unit at price 957.789978, total balance 6861.310117
day 233, sell 1 unit at price 978.890015, investment 5.078469 %, total balance 7840.200132,
day 234, sell 1 unit at price 977.000000, investment 6.083806 %, total balance 8817.200132,
day 235: buy 1 unit at price 972.599976, total balance 7844.600156
day 236: buy 1 unit at price 989.250000, total balance 6855.350156
day 237, sell 1 unit at price 987.830017, investment 6.808602 %, total balance 7843.180173,
day 240, sell 1 unit at price 992.179993, investment 3.590559 %, total balance 8835.360166,
day 243, sell 1 unit at price 988.200012, investment 1.603952 %, total balance 9823.560178,
day 245, sell 1 unit at price 970.539978, investment -1.891334 %, total balance 10794.100156,

```
[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()
```


total gains 794.100156, total investment 7.941002%

[ ]: