Jefferson Parker

CS 677

March 29, 2022

Homework 2, Written Answers

For homework 2, my selected stock symbol was PFE (Pfizer, Inc).

**Question 1.**

* 1. – See code in japarker\_hw2\_1.py
  2. The default probability of an up day is 0.538 for PFE and 0.556 for SPY.
  3. The probability of seeing an up day following k consecutive down days is:

|  |  |  |
| --- | --- | --- |
|  | PFE | SPY |
| k = 2 | 0.624 | 0.678 |
| k = 3 | 0.625 | 0.690 |
| k = 4 | 0.636 | 0.714 |

* 1. The probability of seeing an up day after k consecutive up days is:

|  |  |  |
| --- | --- | --- |
|  | PFE | SPY |
| k = 2 | 0.445 | 0.419 |
| k = 3 | 0.386 | 0.381 |
| k = 4 | 0.370 | 0.310 |

**Question 2.**

2.1 – See code in japarker\_hw2\_2.py

2.2 – The accuracies for the hyperparameter W = 2,3,4 for PFE and SPY are:

|  |  |  |
| --- | --- | --- |
|  | PFE | SPY |
| W = 2 | 73.176 | 57.988 |
| W= 3 | 73.176 | 39.053 |
| W = 4 | 54.043 | 36.292 |

2.3 For PFE, both there was a tie for W = 2 and W = 3, for SPY W = 2 was clearly better.

**Question 3.**

3.1 – See code in japarker\_hw2\_3.py

3.2 – For the ensemble models, the accuracy for PFE was 73.176 and for SPY was 48.718.

3.3 - I did not calculate the accuracy for - predictions (sorry).

3.4 - I did not calculate the accuracy for + predictions (sorry).

**Question 4.**

4.1 – 6 – See code in japarker\_hw2\_4.py

4.7: The summary table of results is:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| W | Ticker | TP | FP | TN | FN | Accuracy | TPR | TNR |
| 2 | PFE | 254 | 3 | 117 | 133 | 73.176 | 0.656 | 0.975 |
| 3 | PFE | 254 | 3 | 117 | 133 | 73.176 | 0.656 | 0.975 |
| 4 | PFE | 215 | 42 | 59 | 191 | 54.043 | 0.53 | 0.584 |
| Ensemble | PFE | 254 | 3 | 117 | 133 | 73.176 | 0.656 | 0.975 |
| 2 | SPY | 295 | 2 | 0 | 210 | 58.185 | 0.584 | 0 |
| 3 | SPY | 199 | 98 | 0 | 210 | 39.25 | 0.487 | 0 |
| 4 | SPY | 159 | 138 | 25 | 185 | 36.292 | 0.462 | 0.153 |
| Ensemble | SPY | 247 | 50 | 0 | 210 | 48.718 | 0.54 | 0 |

4.8 –

I was initially concerned that the PFE predictions were identical across W=2, 3 and ensemble, but seeing that the SPY values were different (and the code was copied directly, only changing the symbols) I was less concerned that this was an error and accepted it was just a fluke of the data.

The ability to accurately predict down days (negative labels) is much higher for the PFE. All models for SPY were either unable to accurately predict downturns or with very poor accuracy (W = 4).

For the S&P 500, the only model that was able to predict a gain day better than random chance was W = 2 … looking at the prior day’s performance.

All in all, this limited sample reaffirms my suspicion that investing the stock market is not much better than games of chance.

**Question 5.**

I’m out of time to get this turned in for credit. I’ll figure out the plotting but accept the lost points.