CS677\_A3\_Fall2023\_WillMohr

**Question 1**

Part 1)

Note: add\_true\_label defined in code.

gme = add\_true\_label(gme)

spy = add\_true\_label(spy)

Part 2) Default probabilities are (0.507, 0.569) for gme, spy respectively

Part 3)

gme: [(1, 0.499), (2, 0.473), (3, 0.5)]  
spy: [(1, 0.563), (2, 0.592), (3, 0.586)]  
where i is number of prior consecutive negative days and j is the probability of a positive return on the following day in the above list of (i, j) tuples.

Part 4)

gme: [(1, 0.513), (2, 0.51), (3, 0.51)]\

spy: [(1, 0.572), (2, 0.567), (3, 0.518)]\

Where i is number of prior consecutive positive days and j is the probability of a positive return on the following day in the above list of (i, j) tuples.

**Question 2**

Part 1)

w\_labs\_gme = predict\_45(gme,[2,3,4])

w\_labs\_spy = predict\_45(spy,[2,3,4])

Part 2)

SPY accuracies:



GME accuracies:



Part 3)

GME: a W of 3 gave the highest accuracy of 0.502

SPY: W’s of 2 and 3 each gave accuracy of 0.579

**Question 3**

Part 1)

w\_e\_labs\_gme = majority\_predict(w\_labs\_gme.copy(deep=True))

w\_e\_labs\_spy = majority\_predict(w\_labs\_spy.copy(deep=True))

Part 2)

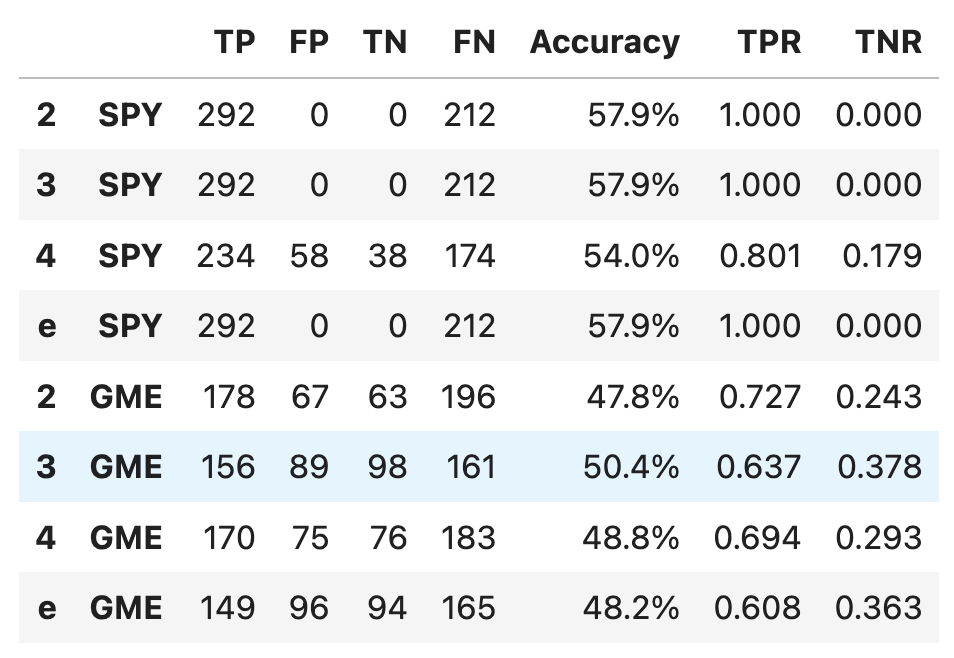
GME: 48.2% accuracy | SPY: 57.9% accuracy

Part 3) For SPY, ensemble accuracy was 0.00% along with w\_2 and w\_3, while w\_4 scored 17.9% accuracy | For GME, ensemble accuracy (36.3%) was better than w\_3 (37.8) and w\_4 (29.3%) but worse than w\_2.

Part 4)

GME: ensemble accuracy (60.8%) was worse than each model alone. | SPY: ensemble accuracy (100%) was better than w\_4 (80.1%) but same as w\_2 (100%) & w\_4 (100%)

**Question 4**

Part 1)

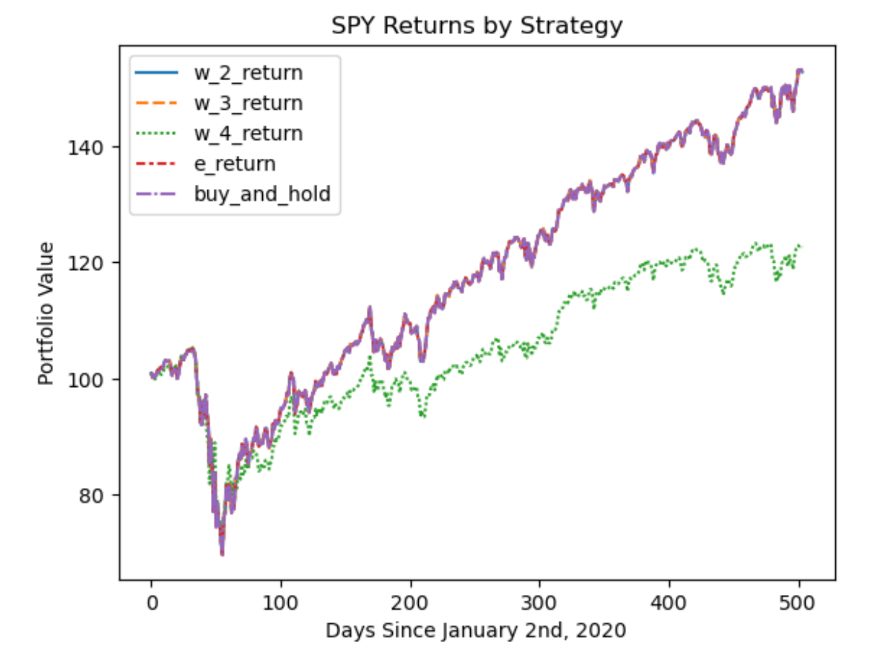
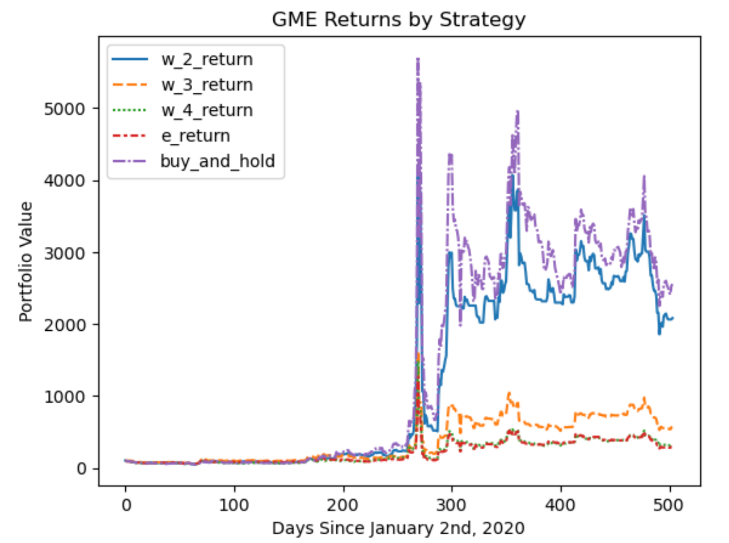
From the statistics above, we can see that for the s&p 500 there are only positive predictions for hyperparams of 2,3, and ensemble while a W of 4 at least makes an attempt at negative predictions, so that some of the length 4 permutations of '+' and '-' are associated with negative returns more than 50% of the time. This seems to match the generally consistent upward trend (and more often than not positive returns) over the five years of the S&P 500 data examined. For GME, the higher number of negative predictions, and hence slightly more acceptable \*\*TNR\*\* indicates that returns following a pattern were less universally positive on average. The surprising fact that the accuracy was under 50% for GME for hyperparameters of W= [2,4, and ensemble] indicate that the overall trend in the returns shifted from years 1,2 & 3 to years 4 & 5. In other words, either a stable trend became more definitively positive or negative or there was a directional switch.

Most accurate, arguably "best" values of hyperparameters for SPY & GME:

SPY: 2,3 and Ensemble are the most accurate

GME: a W of 3 leads to the greatest accuracy.

**Question 5**

Part 1) 

Most accurate, arguably "best" values of hyperparams for SPY & GME:

SPY: 2,3 and Ensemble are the most accurate

GME: a W of 3 leads to the greatest accuracy.

From the charts above, trading based on the most accurate hyperparameters does not necessarily lead to the best cumulative returns over years 4 & 5.

For SPY, the 'time in the market' strategy does in fact generate better returns than the other strategies tested.

For GME, staying out of the market on days predicted to be negative by the W=2 strategy does worse than the buy & hold strategy, while following a trading strategy based on the most accurate hyperparameter does quite horribly in comparison to both the 'time in the market' and W\_2 approaches, while returning slightly better than W\_4. Interestingly, the ensemble strategy does by far the worst! This could be because w\_2, w\_3, and w\_4 tend to predict the same days as negative, so that if one predicts positive and the other two negative the ensemble will chose to stay out of the market, while one of the individual strategies stays in the market and earns a positive return. The high variability in GME's returns could be due to the randomness inherent in extreme volatility, where a small number of key trades or non-trades can make all the difference.

After examining the charts, the analysis resulting from looking at the tables largely hold. The immense returns and losses around the 280 day mark for GME as well as high subsequent volatility supports the notion that missing out on a few trading days can have a massively negative impact on one's portfolio. Similarly, for SPY, if one missed out on a few trading days because one's model decided to stay out of the market, the overall return was significantly worse. Notably, there appeared to be little protection from the pandemic-induced bear market of early 2020 by following the W=4 strategy and choosing not to trade some days.