Assignment 2

-Attack of the True Labels-

Question 1:

Part 1 - true_label.py reads in the csv files for both stocks and assigns a “True Label” determined by the day’s return. This can be seen in the python file by uncommenting lines 42 and 43, where the Dataframes are printed after the “True Label” column is generated.

Part 2 - Assuming that all days are independent, there is a 54.64% chance for COST and a 55.44% chance for SPY that the next day will be an “up” day. This output can be seen by uncommenting lines 96 and 97 in true_label.py.

Part 3 -
*note - all probabilities are shown in percentage format and rounded to the nearest 2 decimal places. Unrounded values are shown in the outputs when the python files are run.

Output for this table can be seen when lines 131 - 134 are uncommented in true_label.py

	COST	SPY
Value of K “Down” Days	Probability of Next “Up” Day	
1	50.88%	59.40%
2	55.95%	59.56%
3	60.81%	63.64%

Part 4 -
*note - all probabilities are shown in percentage format and rounded to the nearest 2 decimal places. Unrounded values are shown in the outputs when the python files are run.

Output for this table can be seen when lines 137 - 140 are uncommented in true_label.py

	COST	SPY
Value of K “Up” Days	Probability of Next “Up” Day	
1	57.56%	52.04%
2	55.32%	50.23%
3	59.69%	46.79%

Question 2:

Part 1 - predict.py computes the predicted labels for w = 2, 3, and 4 for each day in years 4 and 5 based on the true labels gathered from years 1, 2, and 3.

Part 2 - For the last two years of data, my accuracies for w = 2, 3 and 4:
*note - all probabilities are shown in percentage format and rounded to the nearest 2 decimal places. Unrounded values are shown in the outputs when the python files are run.

	COST	SPY
W	Accuracy	
2	54.76%	58.33%
3	54.76%	58.33%
4	54.96%	57.54%

Part 3 - w = 4 gave me the best accuracy for both my stock (COST), where the accuracy was 54.96%. For SPY w = 2 or w = 3 gave me the best accuracy, where they were both 58.33%.

Question 3:

Part 1 - The predictLabels function in predict.py generates the ensemble value from the three values w = 2, 3, and 4 for each day in years 4 and 5.

Part 2 - The ensemble prediction had a 54.76% accuracy for COST and a 58.73% accuracy for SPY.

Part 3 - No, overall the ensemble prediction for “-“ labels was around the midpoint of W values for both my stock and SPY.

Part 4 - No, overall the ensemble prediction for “+” labels was also around the midpoint of W values for both my stock and SPY.

Question 4:

Parts 1 - 6 - Each of the values named in parts 1 - 6 are calculated in the calcAccuracy function in predict.py and printed on lines 206 - 213.

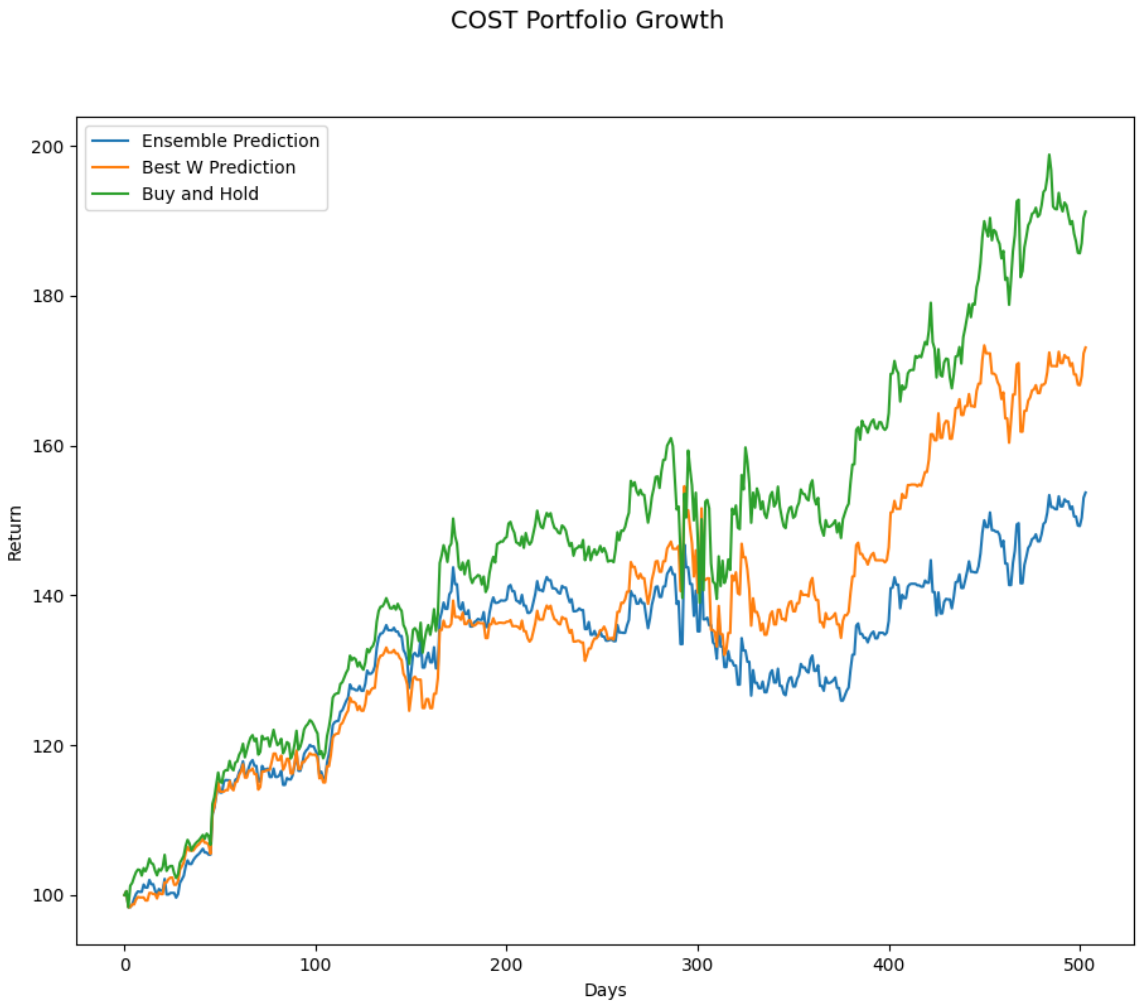
Part 7 -
*note - all probabilities are shown in percentage format and rounded to the nearest 2 decimal places. Unrounded values are shown in the outputs when the python files are run.

W	Ticker	TP	FP	TN	FN	Accuracy	TPR	TNR
2	SPY	294	210	0	0	58.33%	100%	0%
3	SPY	246	162	48	48	58.33%	83.67%	22.86%
4	SPY	236	156	54	58	57.54%	80.27%	25.71%
Ensemble	SPY	271	185	25	23	58.73%	92.18%	11.90%
2	COST	222	165	54	63	54.76%	77.89%	24.66%
3	COST	222	165	54	63	54.76%	77.89%	24.66%
4	COST	214	156	63	71	54.96%	75.09%	28.77%
Ensemble	COST	222	165	54	63	54.76%	77.89%	24.66%

Part 8 - For both my stock (COST) and SPY, there are two values for W (W = 2 and W = 3) that have similar or identical sets of stats. Another interesting note is for w = 2 in SPY, there was a 100% true positive rate and 0% true negative rate which means the training model only had groupings of three days that ended in positives. The true positive rates were all about 75% and the true negative rates were all below 29% which shows the model was much better at predicting positive days than it was negative days.

Question 5:

Part 1 -



Part 2 - The most noticeable pattern over the course of the two years of returns is the trend of Buy and Hold resulting in the highest return, Best W Prediction having the middle return, and Ensemble Prediction having the lowest return. This is especially interesting since the prediction method with the most data that went into it (ensemble) turned out to be the least effective. Another trend is that in the first year the values, despite changing from 100 to around 140, all stayed very similar in return un until around day 300. From then on in the second year, each of the three methods had much larger differences despite having similar overall motion.