Project 2

By: Richard Foggio, Alex Trainham, and Noah Weingand
CMSC 409 - Fall 2020 - Dr. Manic

Student certification:

Team member 1:
Print Name: ___Noah Weingand_____ Date: __10-1-2020_____
I have contributed by doing the following:

Driver for group programming testing for hard and soft activation, sampling training data from datasets, calculating false/true positives and negatives and associated accuracies, filling in results from program, finding best results for Graph B's hard activation, Alt Group C comparisons

Signed: 

Team member 2:
Print Name: _____Richard Foggio_____ Date: _____10-1-2020_____
I have contributed by doing the following:

Coding algorithm, sampling data, selecting train_in, train_out, test_in_test_out arrays. Graphing results. Answering questions. GroupA, GroupC, and Alt A comparisons.

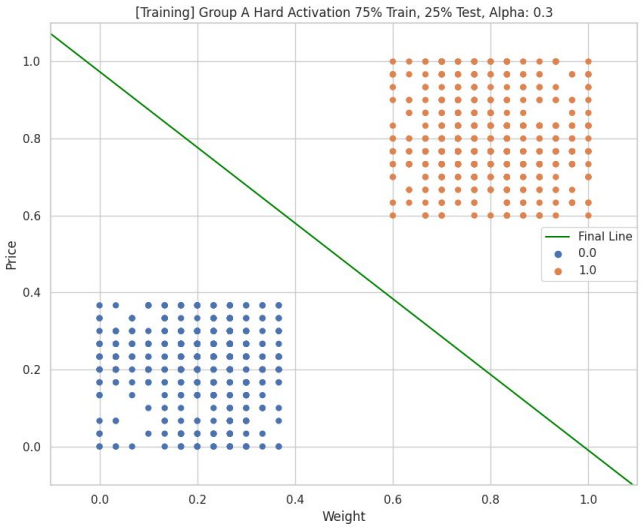
Signed: 

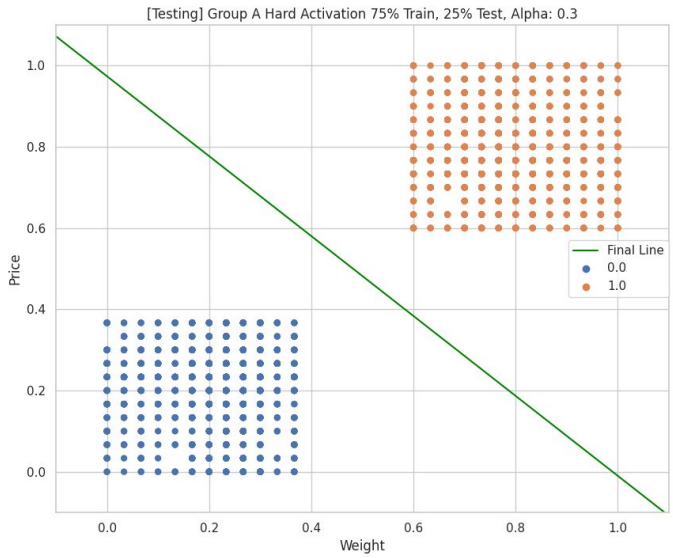
Team member 3:
Print Name: __Alex Trainham_____ Date: ___10-01-2020_____
I have contributed by doing the following: Filling out the response questions, Determining testing/training splits. Make an initial perceptron algorithm with Noah. Wrote the testing function and analysis on the trained perceptron.

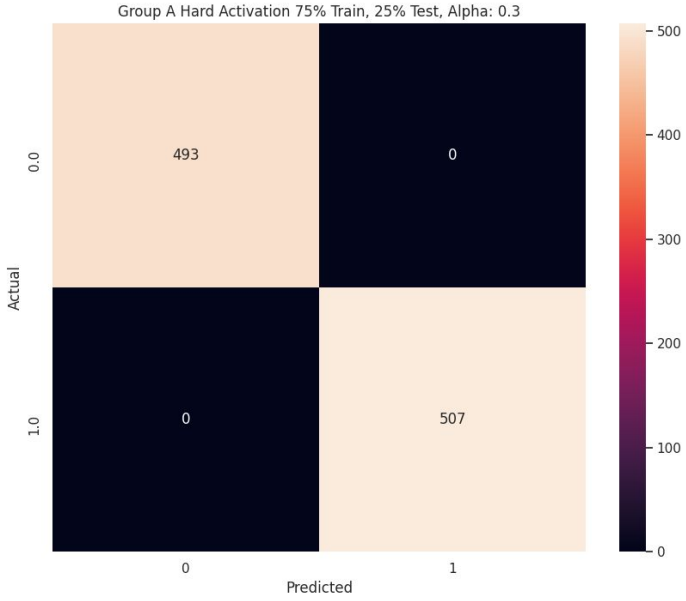
Signed:

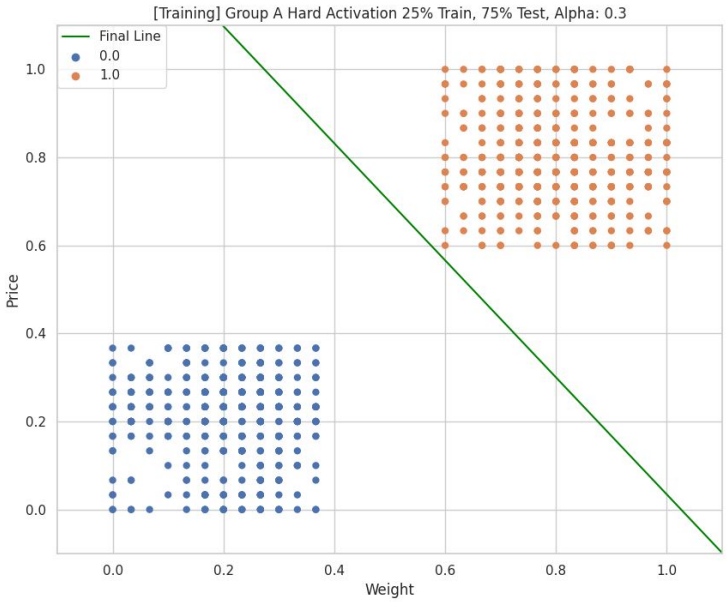
Hard Activation (Graphs)

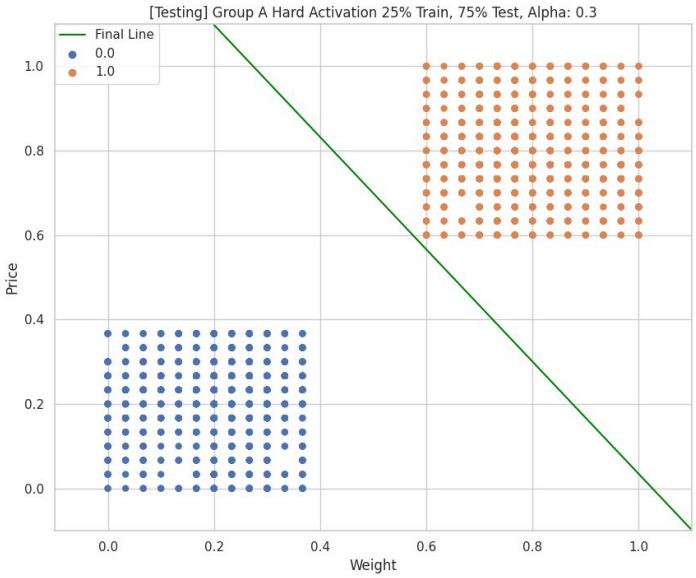
Group A

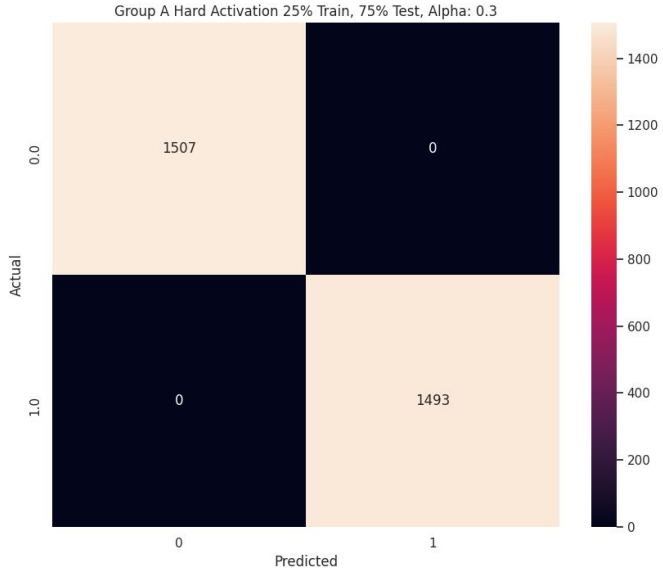



[Testing] Group A Hard Activation 75% Train, 25% Test, Alpha: 0.3
Price
Weight
Final Line
0.0
1.0
0.0
0.2
0.4
0.6
0.8
1.0

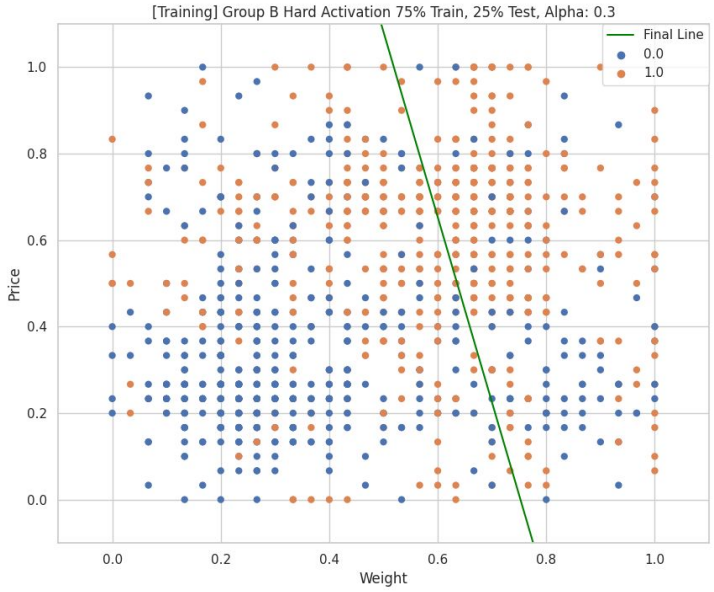

Group A Hard Activation 75% Train, 25% Test, Alpha: 0.3
Actual
Predicted
493
0
0
507
0.0
1.0
0
1
500
400
300
200
100
0

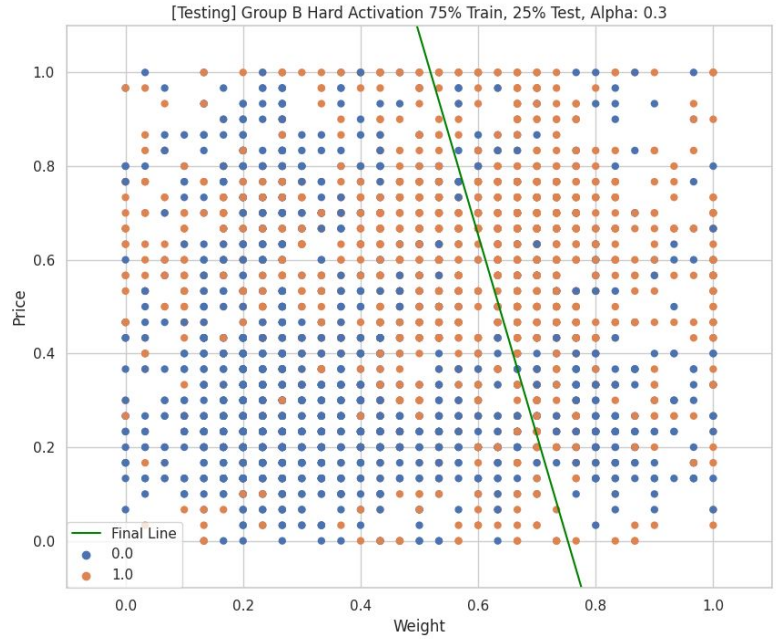

[Training] Group A Hard Activation 25% Train, 75% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

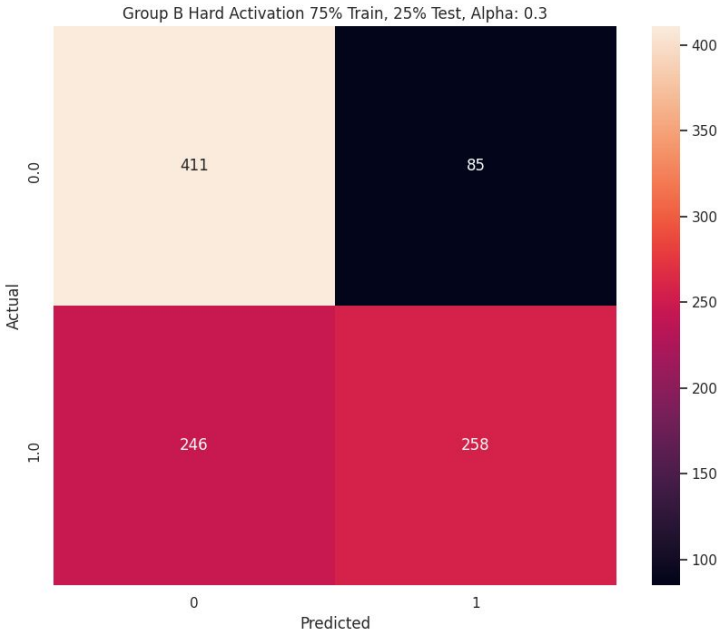

[Testing] Group A Hard Activation 25% Train, 75% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

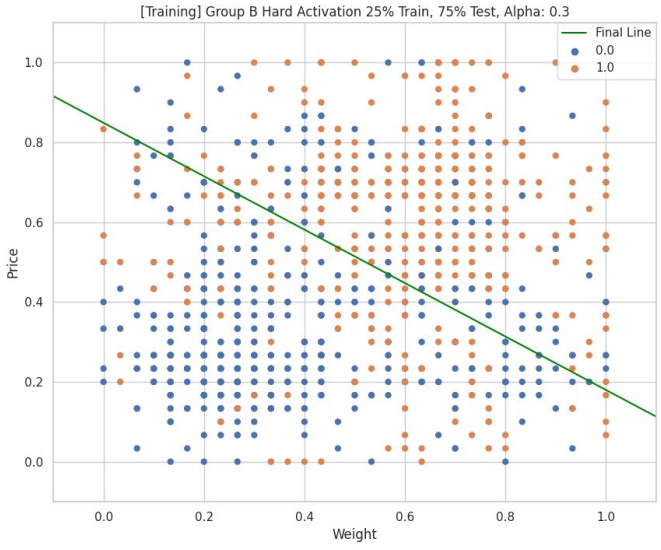

Group A Hard Activation 25% Train, 75% Test, Alpha: 0.3
1507
0
0
1493
Actual
Predicted
0.0
1.0
0
1
1400
1200
1000
800
600
400
200
0

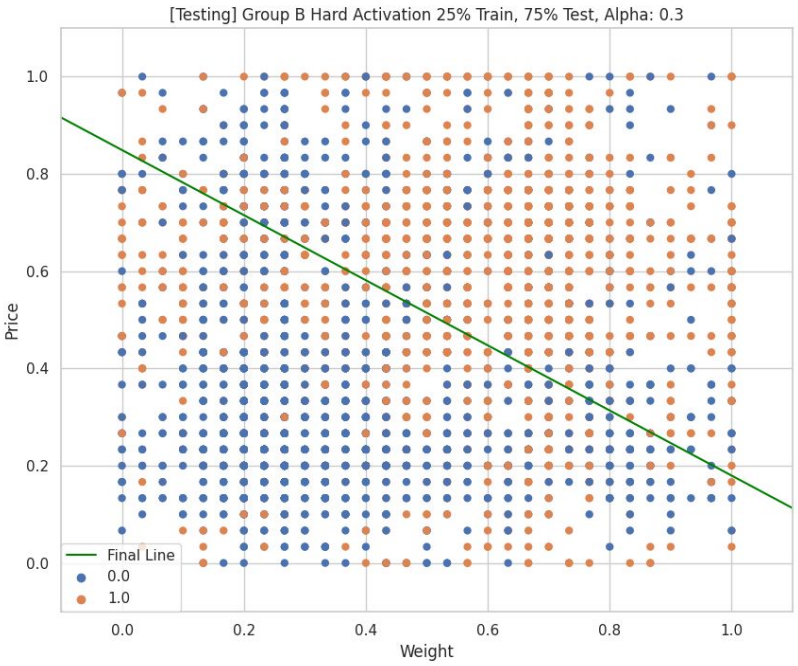
Group B

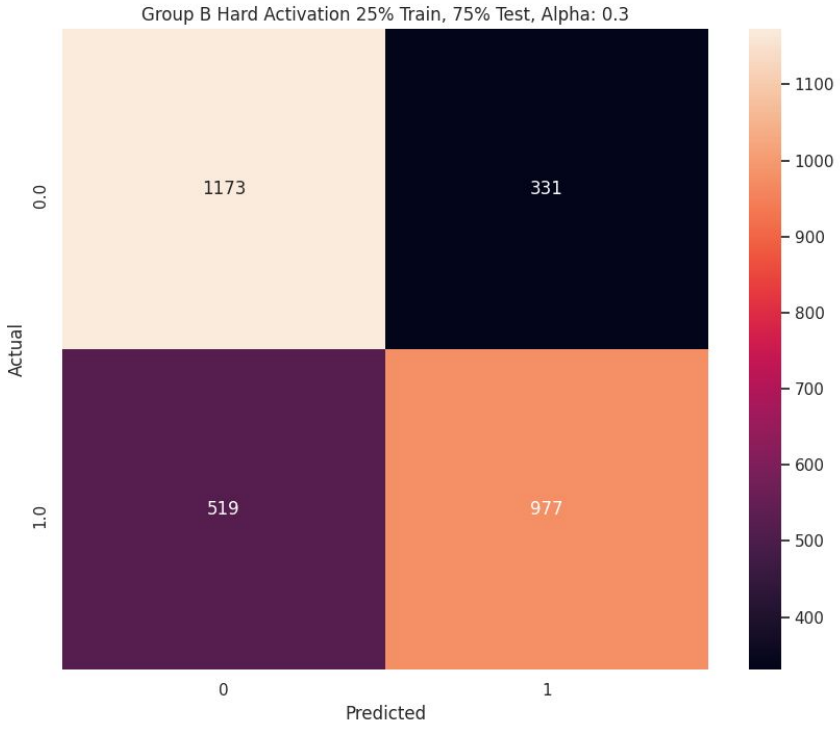

[Training] Group B Hard Activation 75% Train, 25% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0


[Testing] Group B Hard Activation 75% Train, 25% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

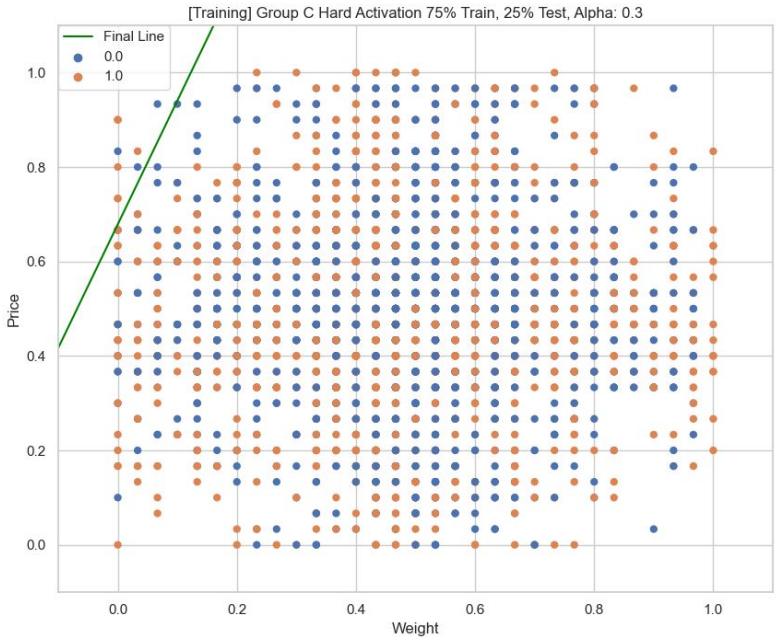

Group B Hard Activation 75% Train, 25% Test, Alpha: 0.3
411
85
246
258
0.0
1.0
Actual
0
1
Predicted
400
350
300
250
200
150
100

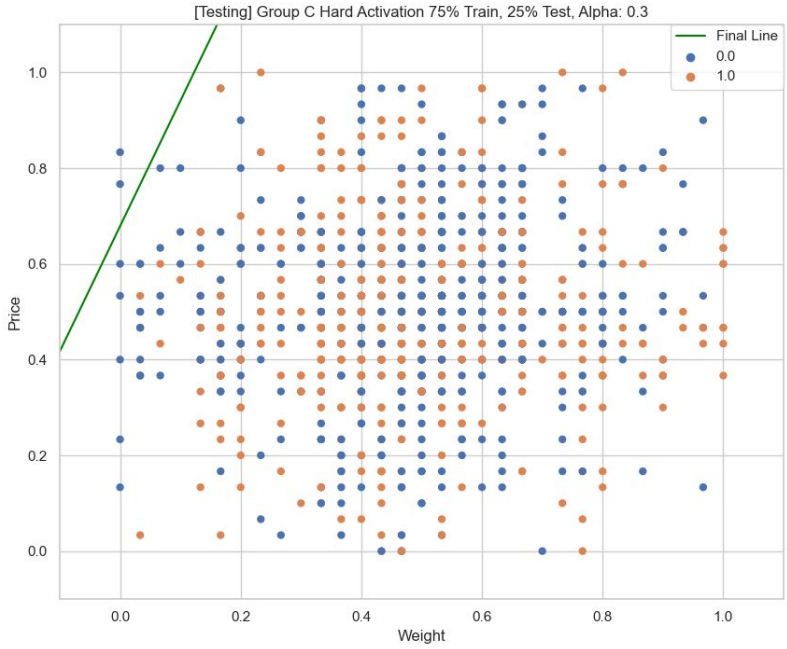

[Training] Group B Hard Activation 25% Train, 75% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

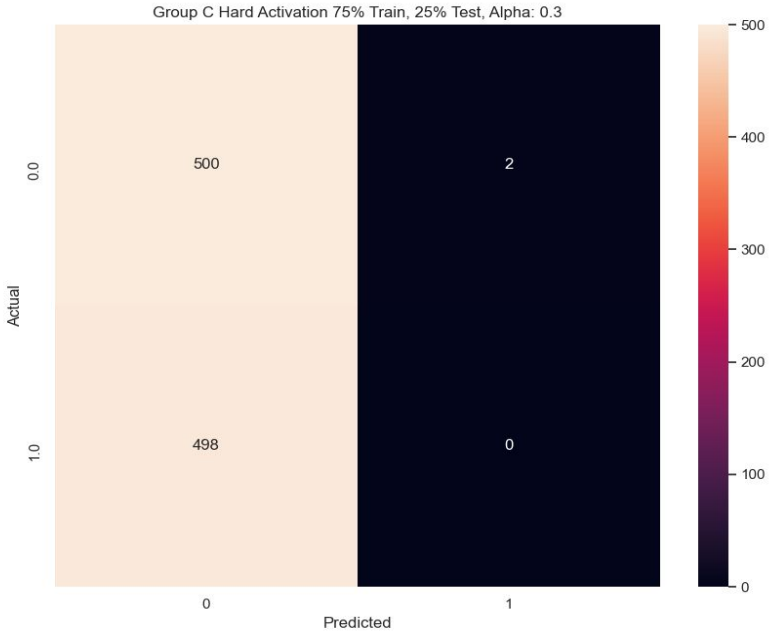

[Testing] Group B Hard Activation 25% Train, 75% Test, Alpha: 0.3
Price
Weight
Final Line
0.0
1.0
0.0
0.2
0.4
0.6
0.8
1.0

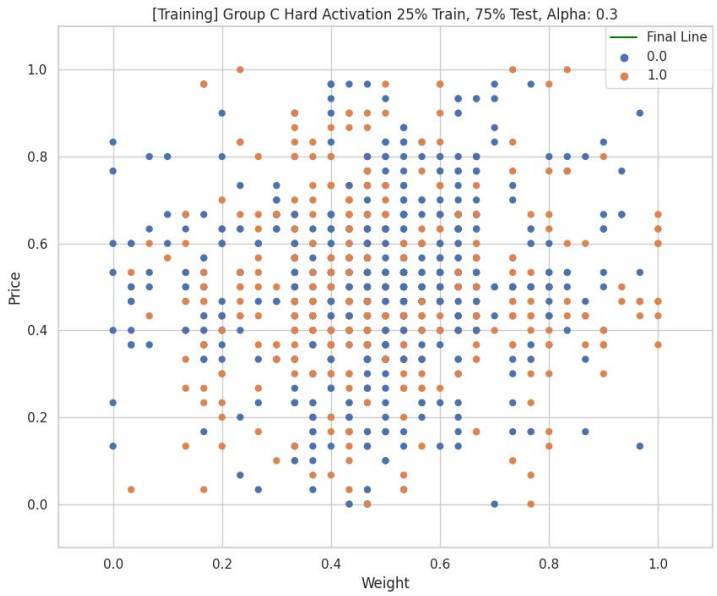

Group B Hard Activation 25% Train, 75% Test, Alpha: 0.3
1173
331
519
977
Actual
Predicted
0.0
1.0
0
1
1100
1000
900
800
700
600
500
400

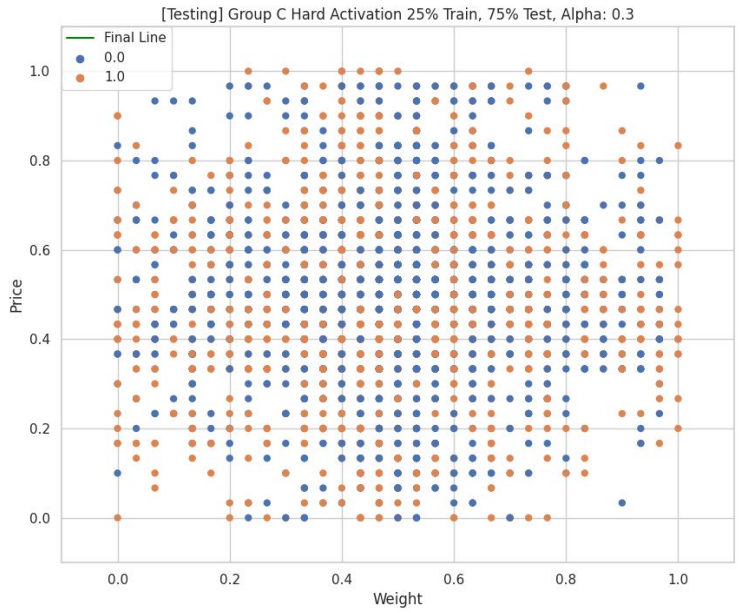
Group C

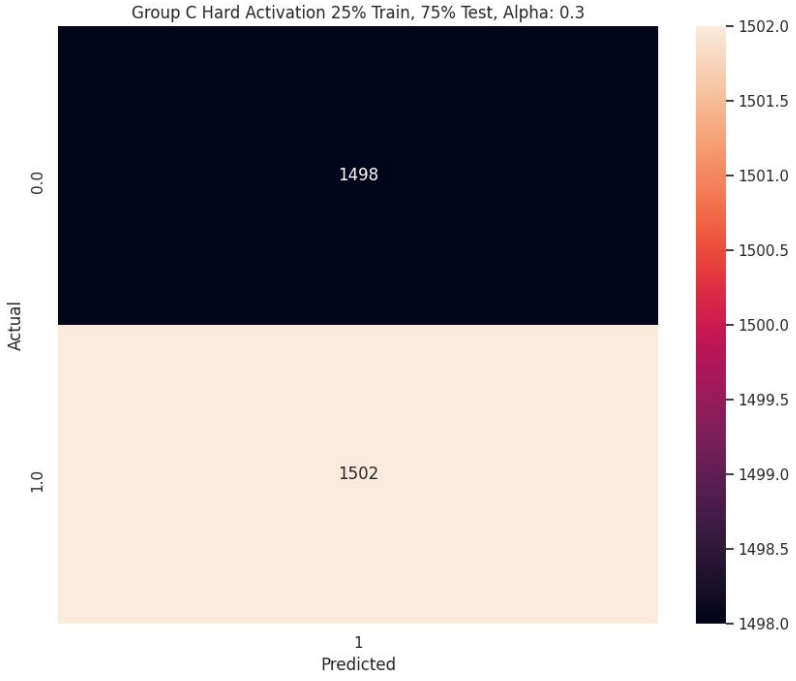

[Training] Group C Hard Activation 75% Train, 25% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0


[Testing] Group C Hard Activation 75% Train, 25% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

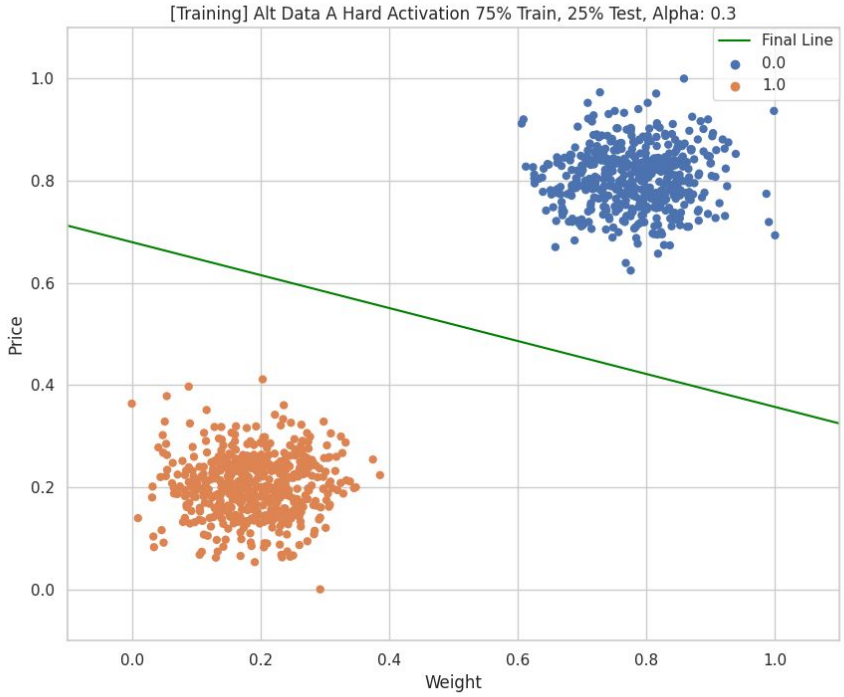

Group C Hard Activation 75% Train, 25% Test, Alpha: 0.3
500
2
498
0
0.0
1.0
Actual
0
1
Predicted
500
400
300
200
100
0

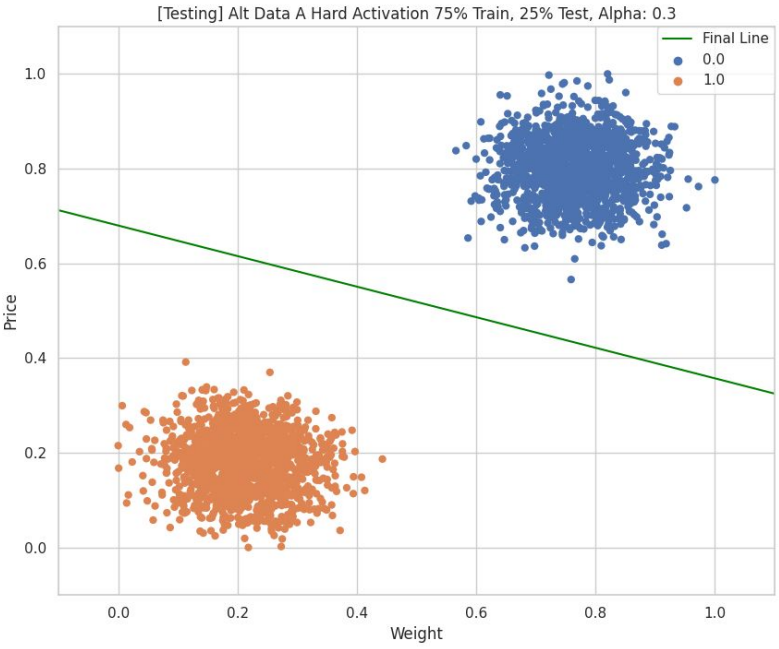

[Training] Group C Hard Activation 25% Train, 75% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

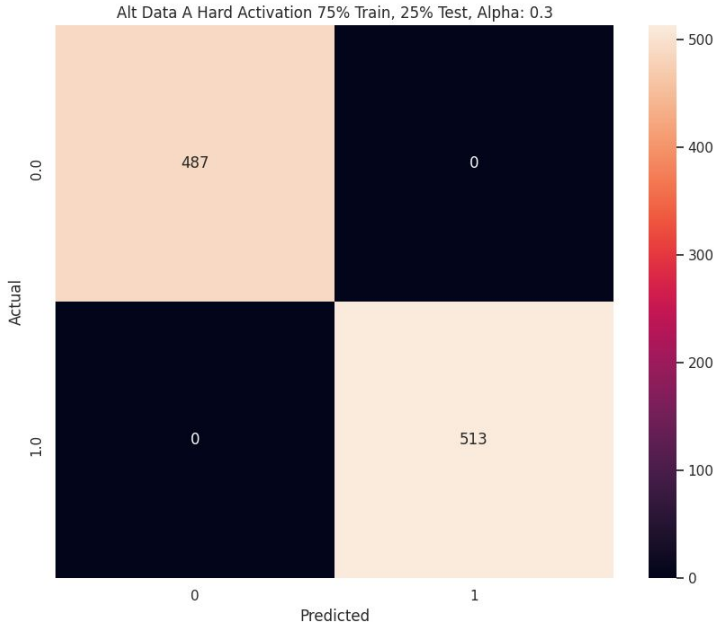

[Testing] Group C Hard Activation 25% Train, 75% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0


Group C Hard Activation 25% Train, 75% Test, Alpha: 0.3
1498
1502
Actual
Predicted
0.0
1.0
1
1502.0
1501.5
1501.0
1500.5
1500.0
1499.5
1499.0
1498.5
1498.0

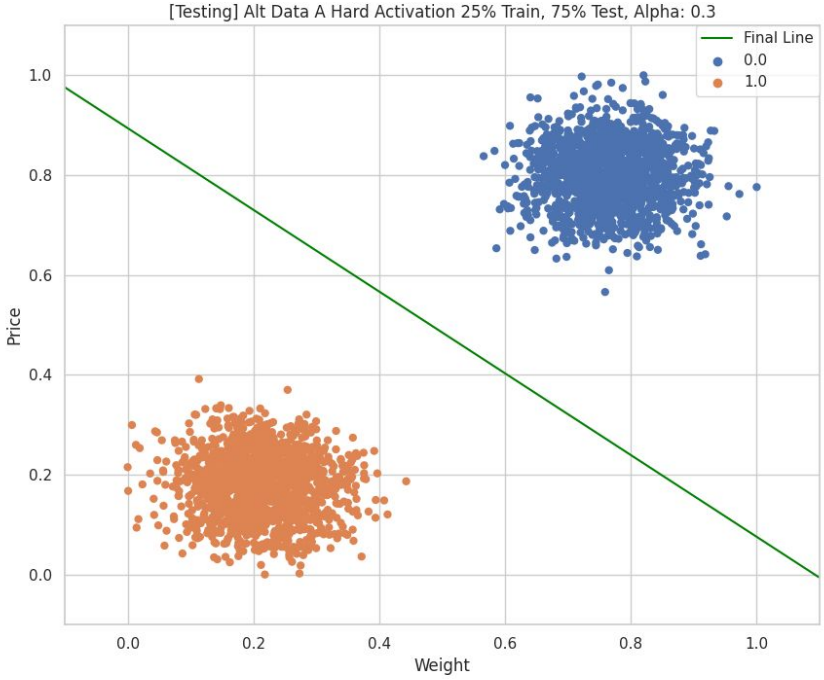
Group A Alt

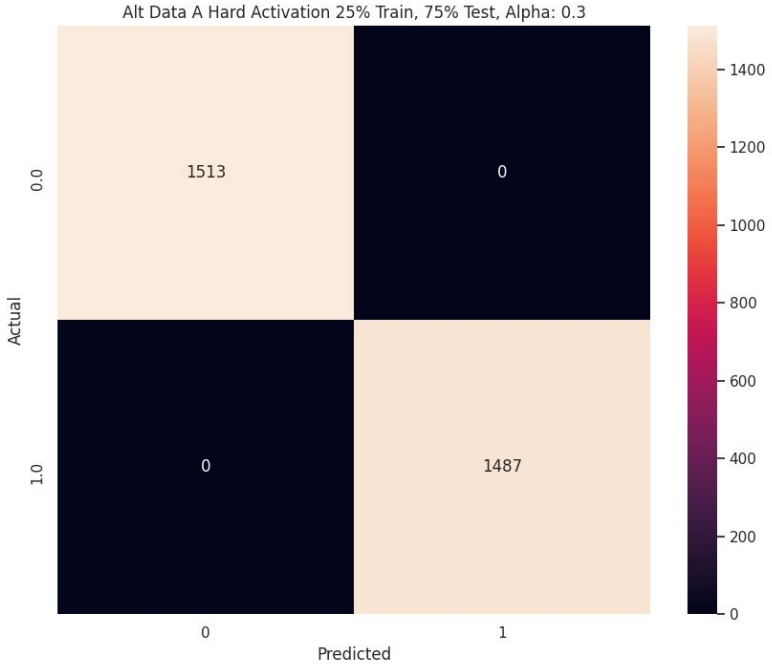

[Training] Alt Data A Hard Activation 75% Train, 25% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0


[Testing] Alt Data A Hard Activation 75% Train, 25% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

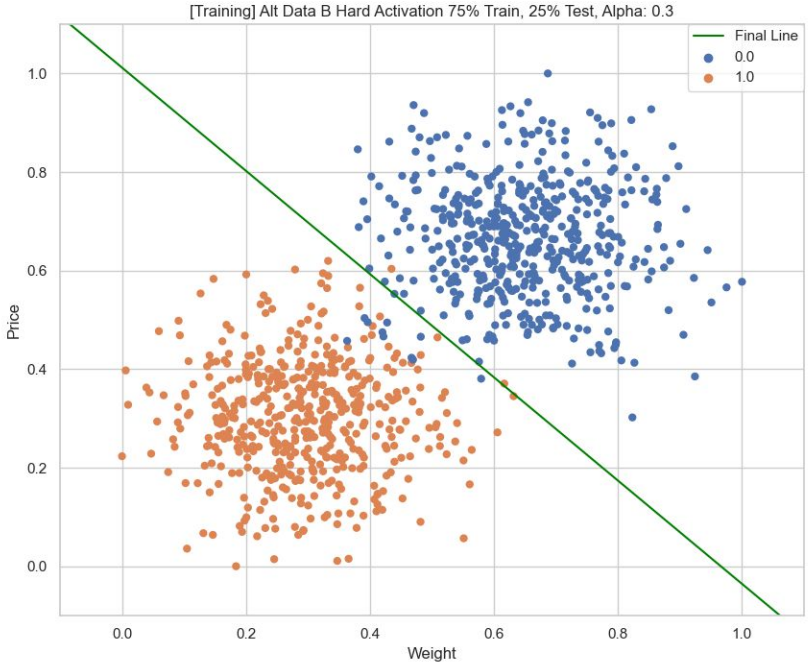

Alt Data A Hard Activation 75% Train, 25% Test, Alpha: 0.3
487
0
0
513
0.0
1.0
Actual
0
1
Predicted
500
400
300
200
100
0

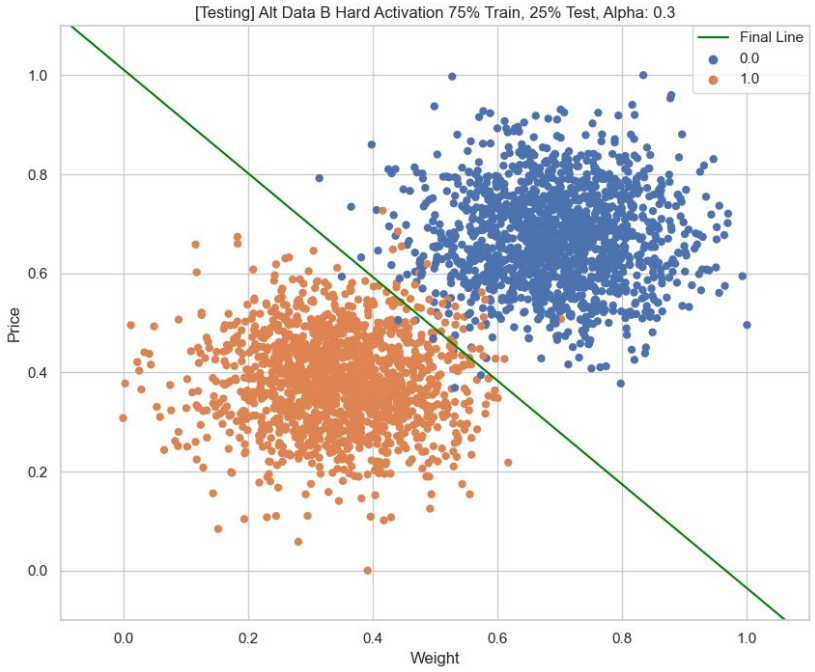

[Training] Alt Data A Hard Activation 25% Train, 75% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

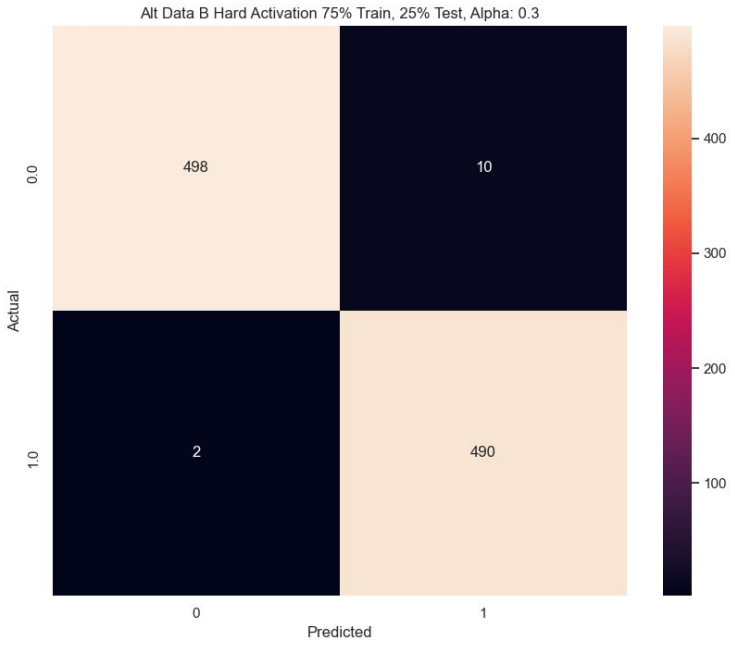

[Testing] Alt Data A Hard Activation 25% Train, 75% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

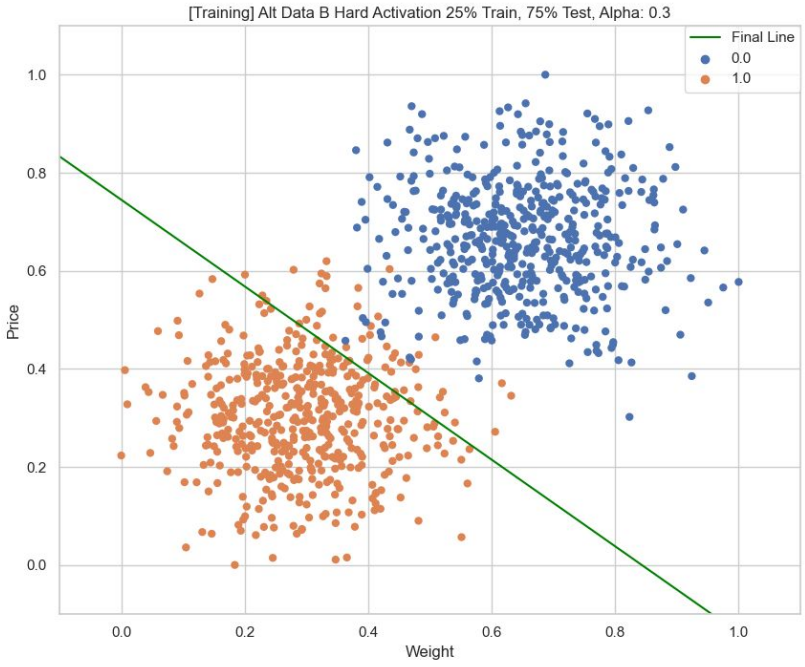

Alt Data A Hard Activation 25% Train, 75% Test, Alpha: 0.3
1513
0
0
1487
Actual
Predicted
0.0
1.0
0
1
1400
1200
1000
800
600
400
200
0

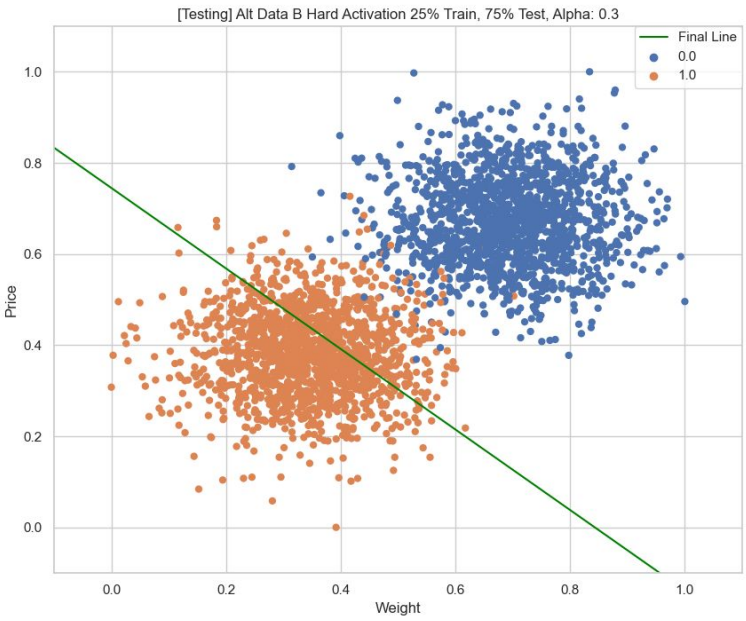
Group B Alt

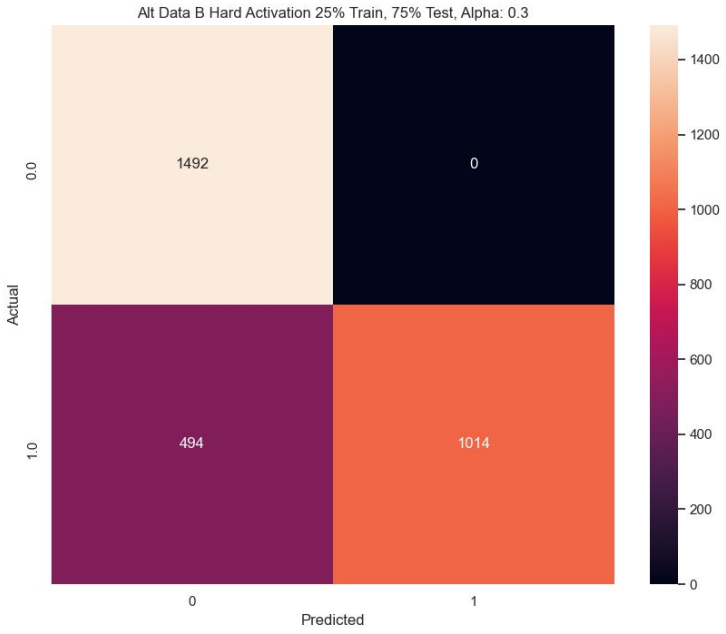

[Training] Alt Data B Hard Activation 75% Train, 25% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0


[Testing] Alt Data B Hard Activation 75% Train, 25% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

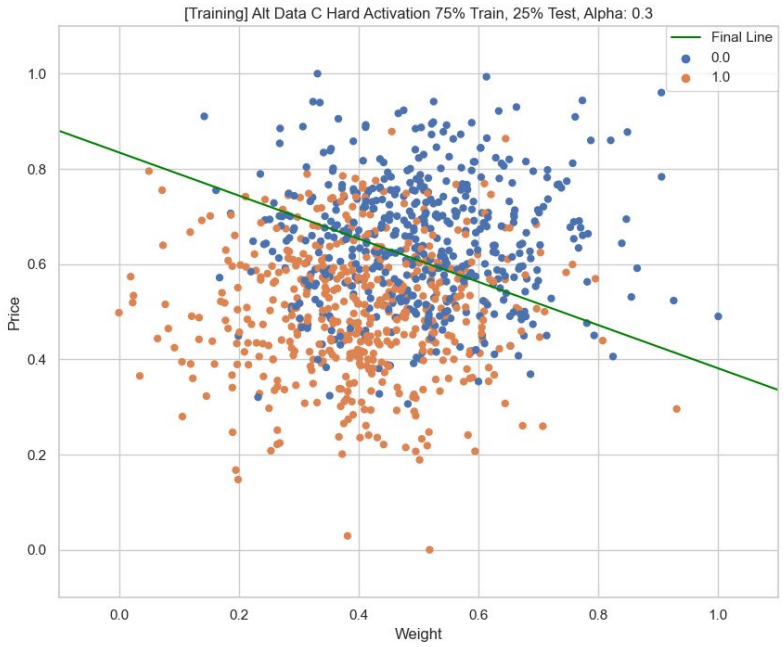

Alt Data B Hard Activation 75% Train, 25% Test, Alpha: 0.3
498
10
2
490
0.0
1.0
Actual
0
1
Predicted
400
300
200
100

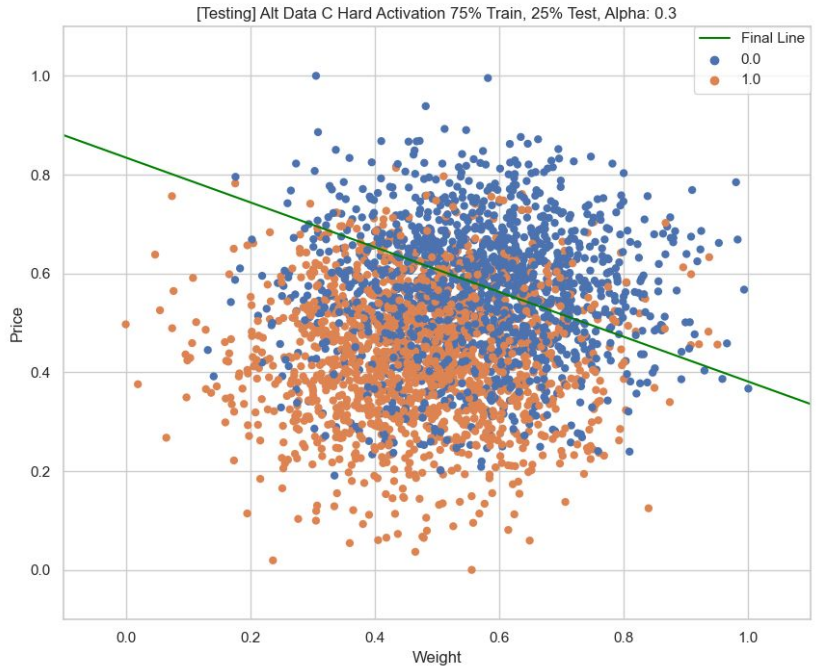

[Training] Alt Data B Hard Activation 25% Train, 75% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

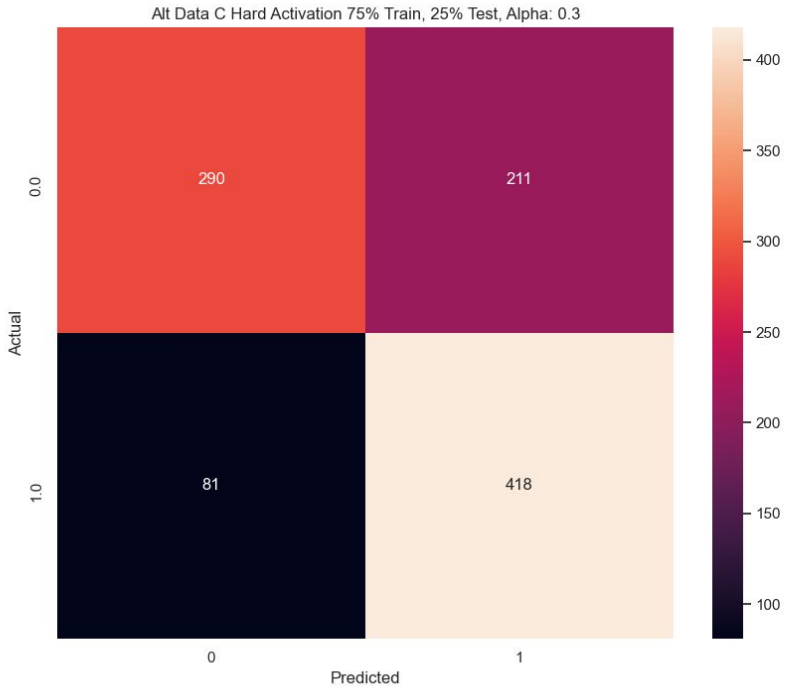

[Testing] Alt Data B Hard Activation 25% Train, 75% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

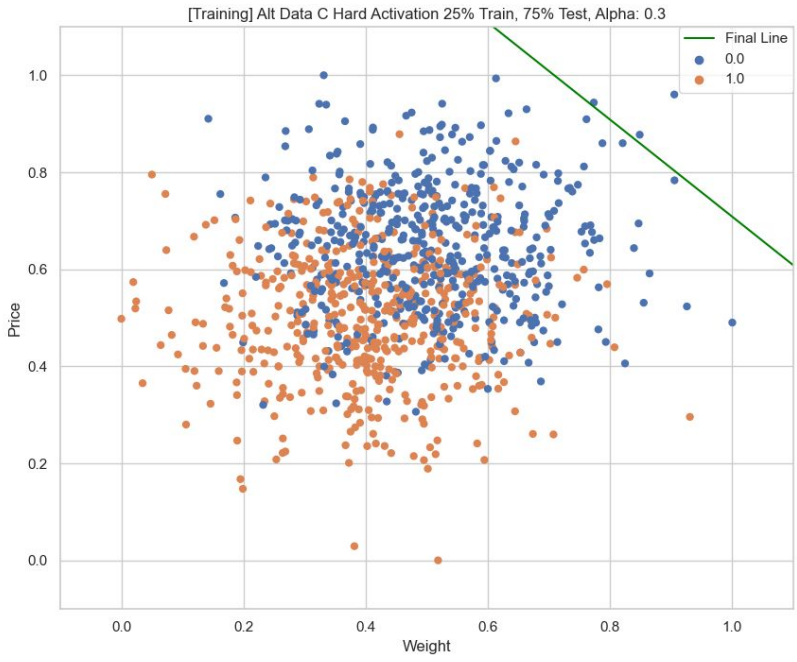

Alt Data B Hard Activation 25% Train, 75% Test, Alpha: 0.3
1492
0
494
1014
Actual
Predicted
0.0
1.0
0
1
1400
1200
1000
800
600
400
200
0

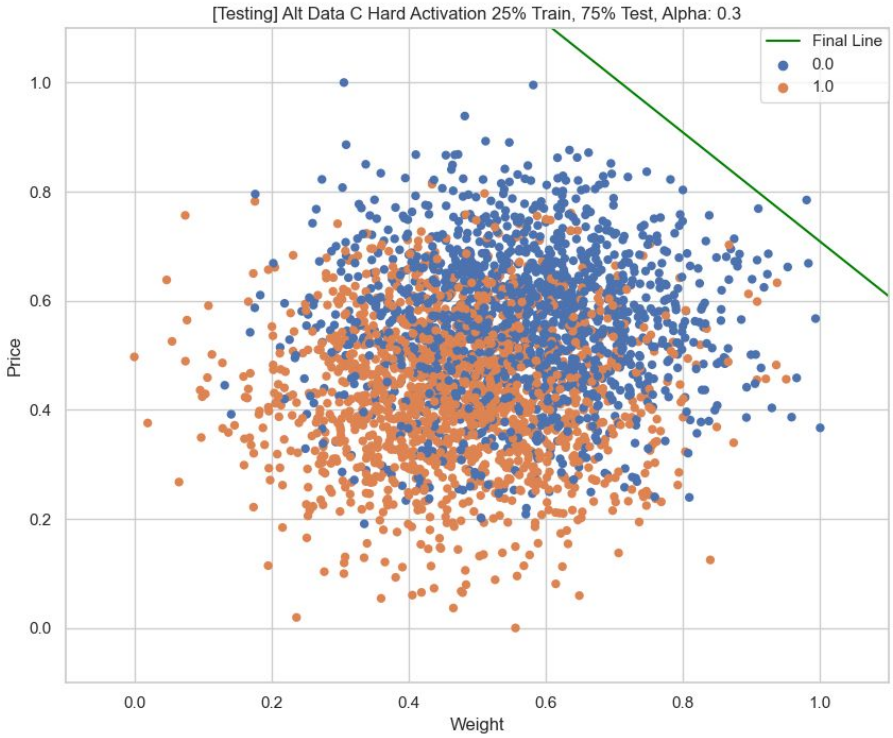
Group C Alt

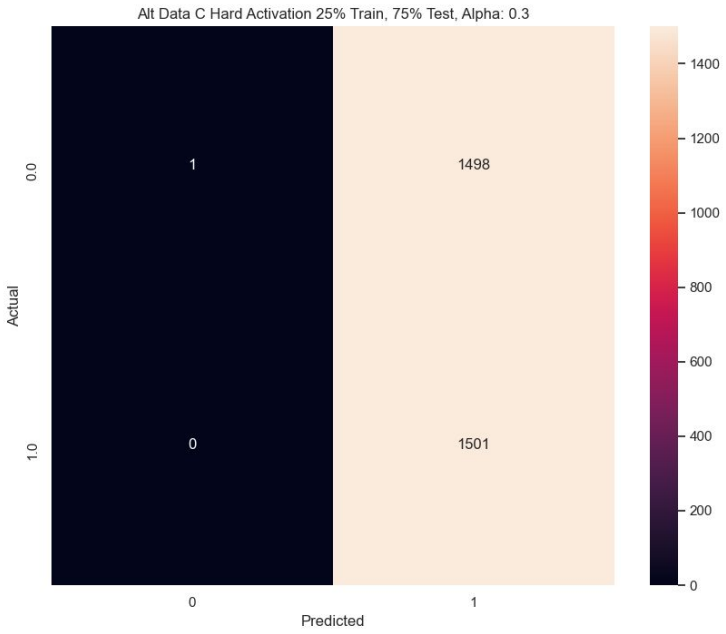

[Training] Alt Data C Hard Activation 75% Train, 25% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0


[Testing] Alt Data C Hard Activation 75% Train, 25% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0


Alt Data C Hard Activation 75% Train, 25% Test, Alpha: 0.3
290
211
81
418
0.0
1.0
Actual
0
1
Predicted
400
350
300
250
200
150
100

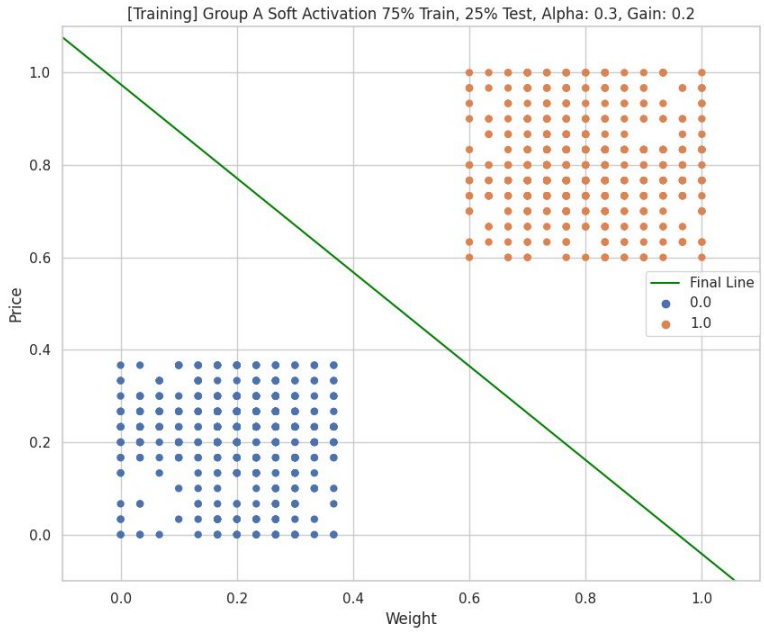

[Training] Alt Data C Hard Activation 25% Train, 75% Test, Alpha: 0.3
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

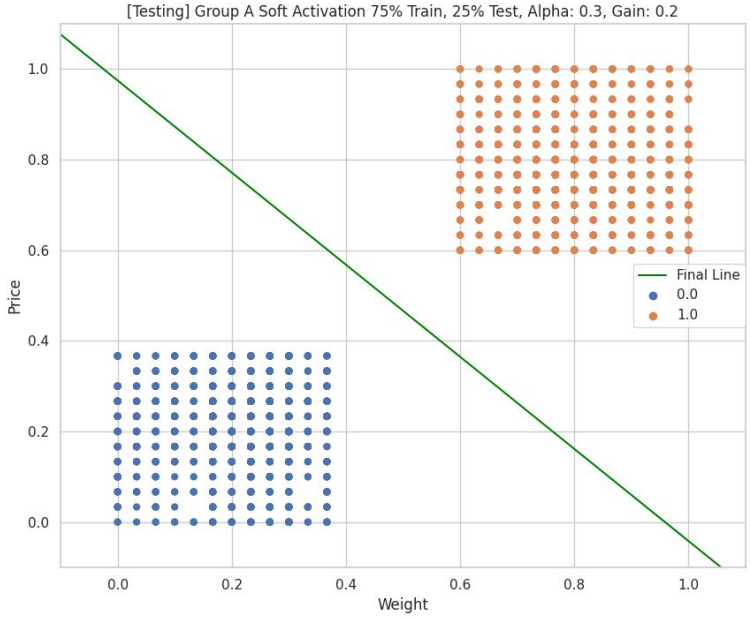


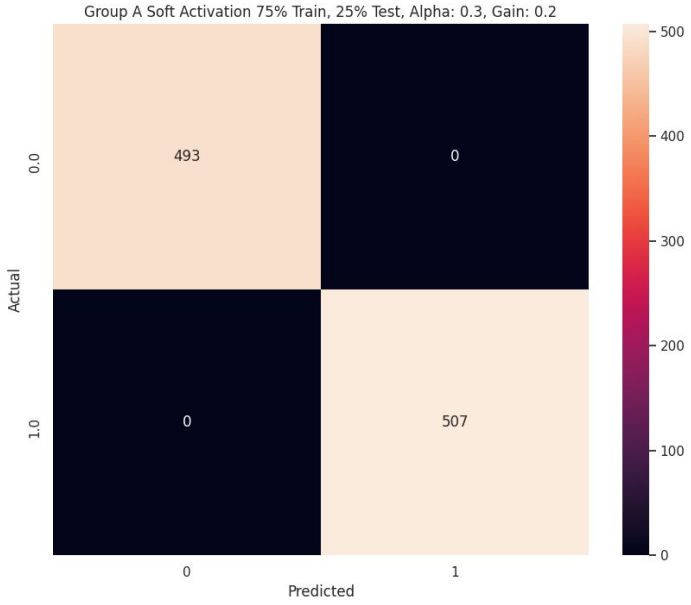


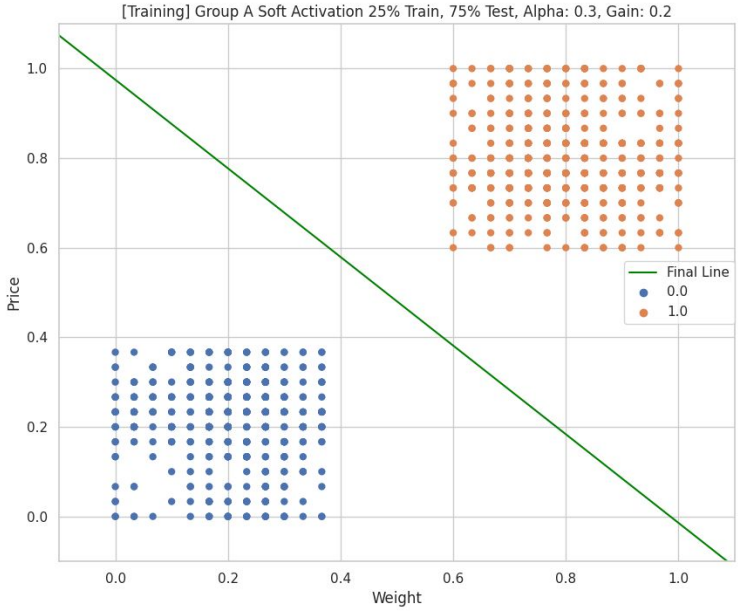
Soft Activation (Graphs)

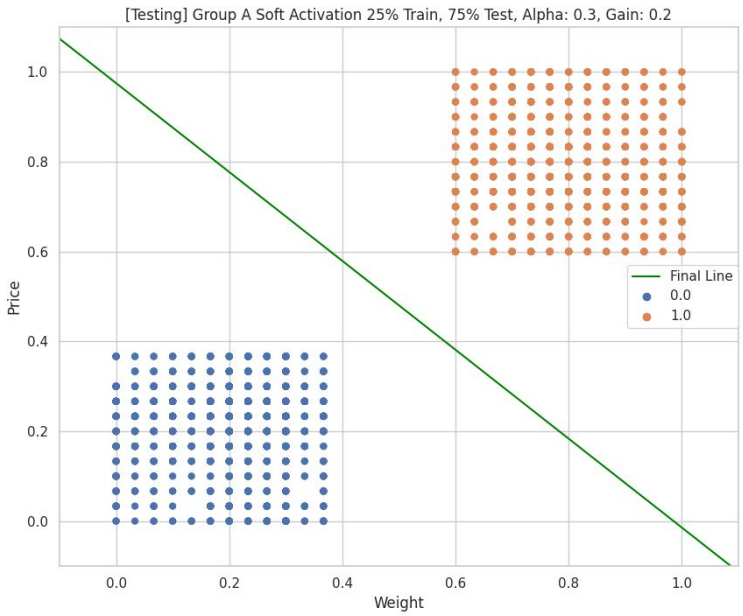
Group A

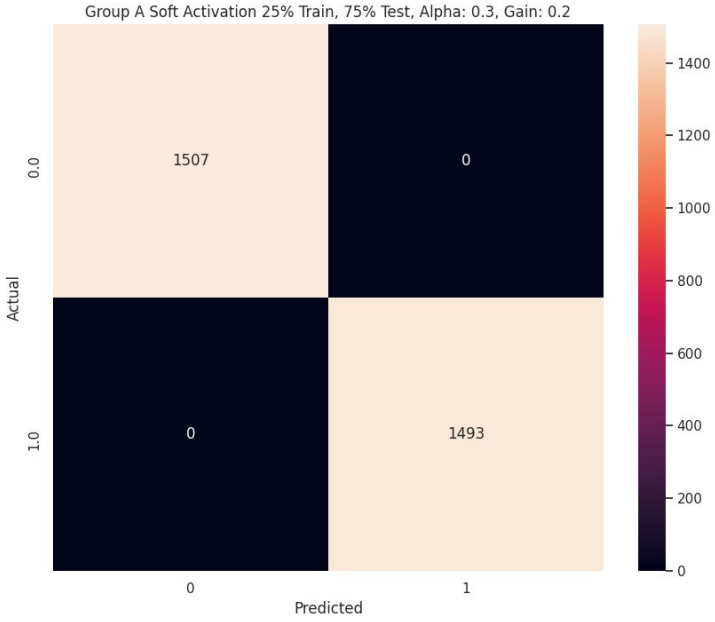

[Training] Group A Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
Price
Weight
Final Line
0.0
1.0


[Testing] Group A Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
Price
Weight
Final Line
0.0
1.0

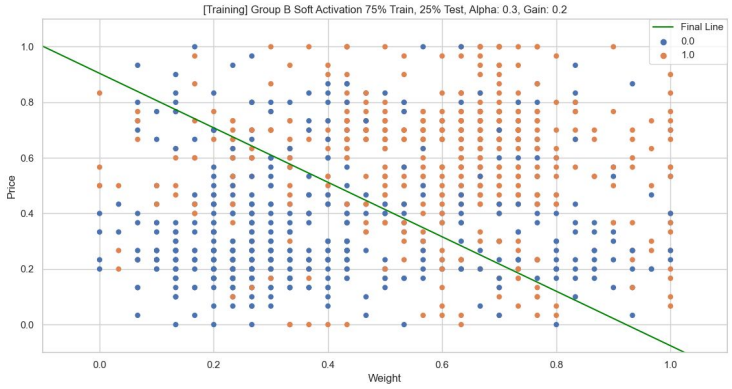

Group A Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
493
0
0
507
0.0
1.0
Actual
0
1
Predicted
500
400
300
200
100
0

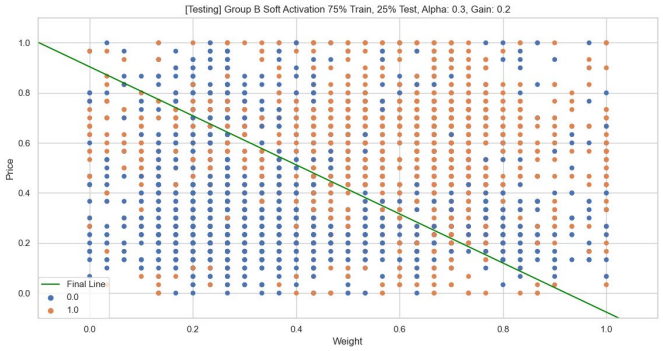

[Training] Group A Soft Activation 25% Train, 75% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

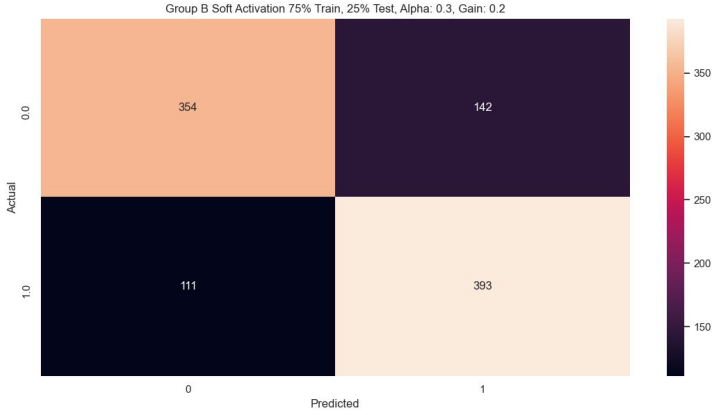

[Testing] Group A Soft Activation 25% Train, 75% Test, Alpha: 0.3, Gain: 0.2
Price
Weight
Final Line
0.0
1.0
0.0
0.2
0.4
0.6
0.8
1.0

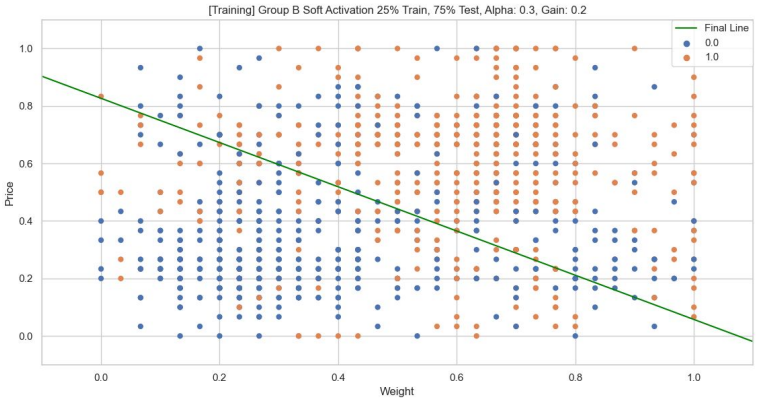

Group A Soft Activation 25% Train, 75% Test, Alpha: 0.3, Gain: 0.2
Actual
Predicted
1507
0
0
1493
0.0
1.0
0
1
1400
1200
1000
800
600
400
200
0

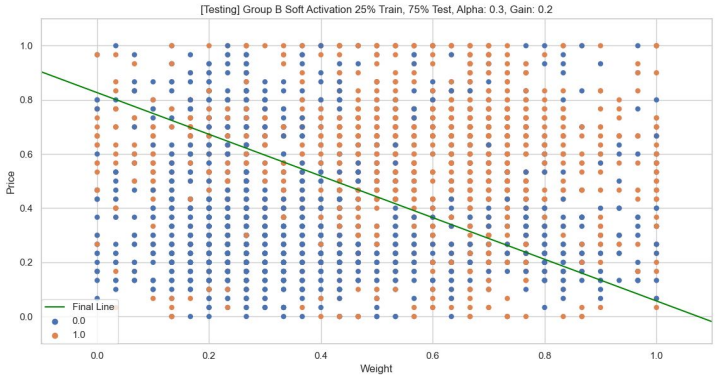
Group B

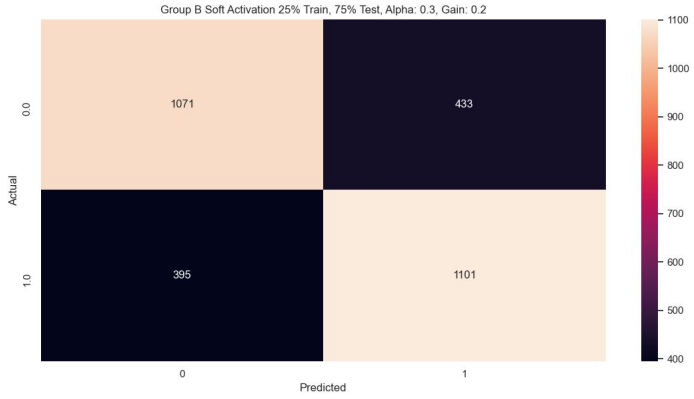

[Training] Group B Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0


[Testing] Group B Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

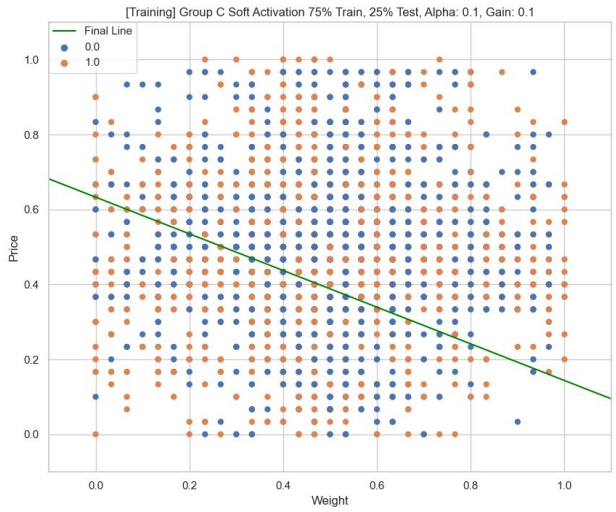

Group B Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
354
142
111
393
Actual
Predicted
0.0
1.0
0
1
350
300
250
200
150

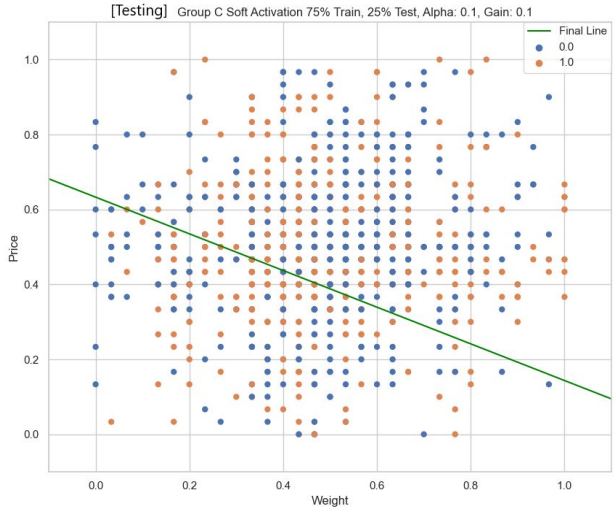

[Training] Group B Soft Activation 25% Train, 75% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

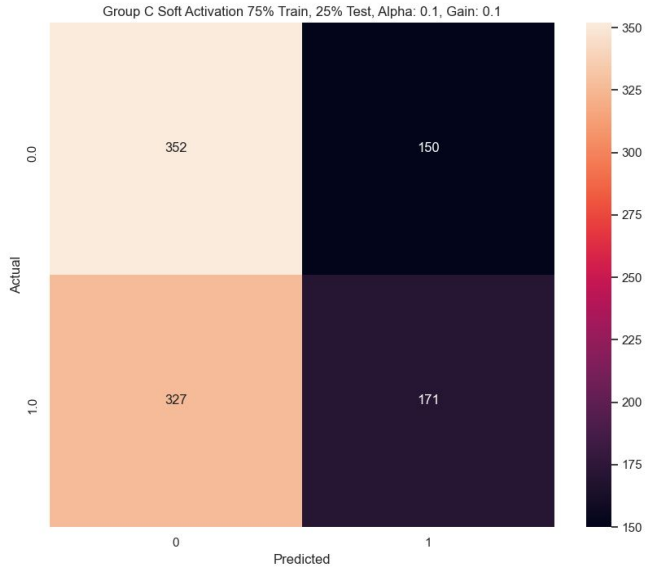

[Testing] Group B Soft Activation 25% Train, 75% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

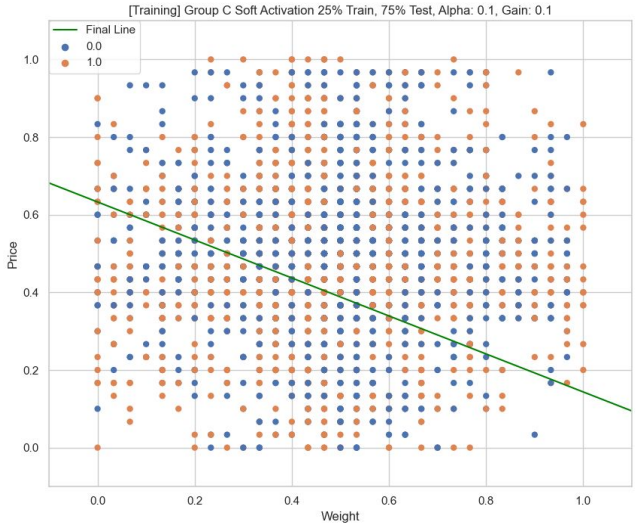

Group B Soft Activation 25% Train, 75% Test, Alpha: 0.3, Gain: 0.2
1071
433
395
1101
Actual
Predicted
0.0
1.0
0
1
1100
1000
900
800
700
600
500
400

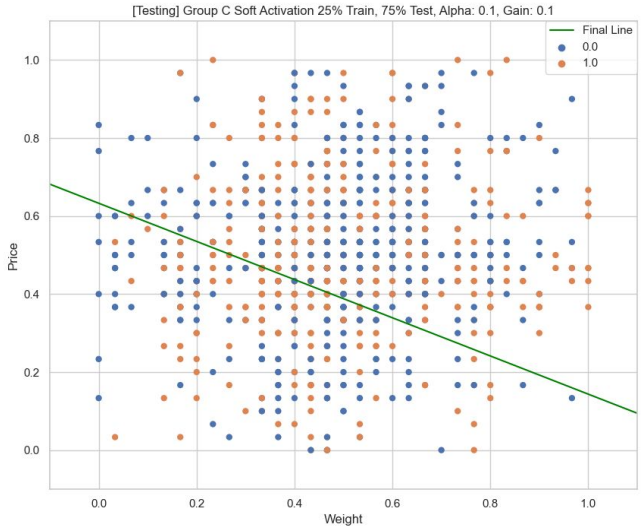
Group C

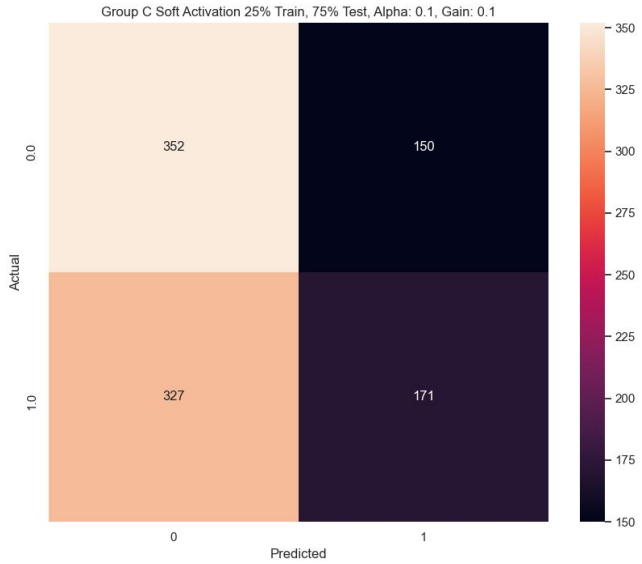

[Training] Group C Soft Activation 75% Train, 25% Test, Alpha: 0.1, Gain: 0.1
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0


[Testing] Group C Soft Activation 75% Train, 25% Test, Alpha: 0.1, Gain: 0.1
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

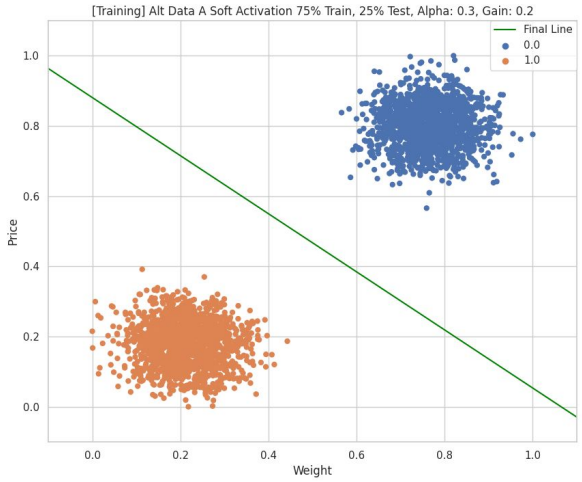

Group C Soft Activation 75% Train, 25% Test, Alpha: 0.1, Gain: 0.1
352
150
327
171
0.0
1.0
Actual
0
1
Predicted
350
325
300
275
250
225
200
175
150

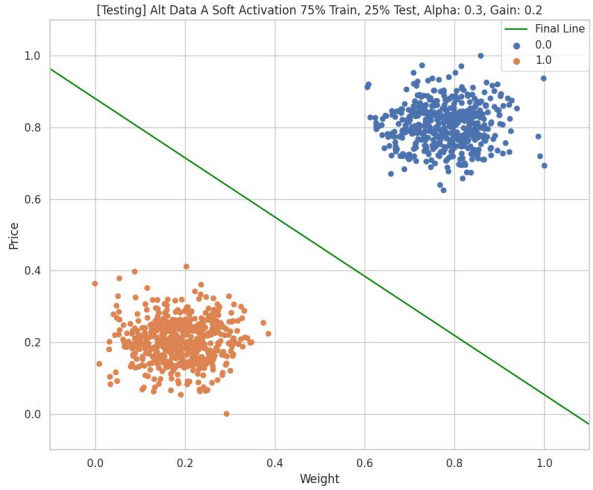

[Training] Group C Soft Activation 25% Train, 75% Test, Alpha: 0.1, Gain: 0.1
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

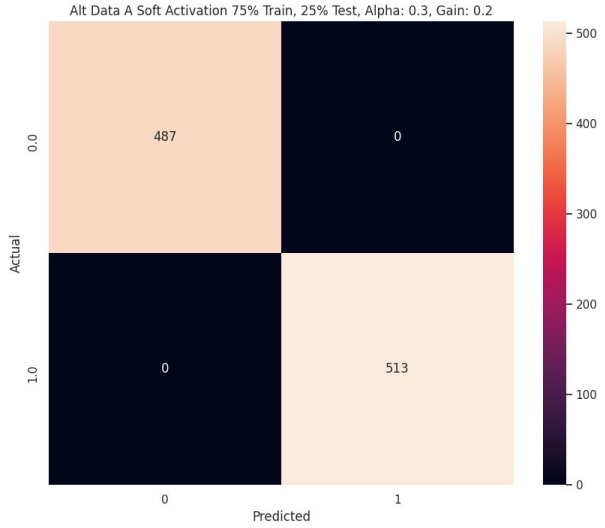

[Testing] Group C Soft Activation 25% Train, 75% Test, Alpha: 0.1, Gain: 0.1
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

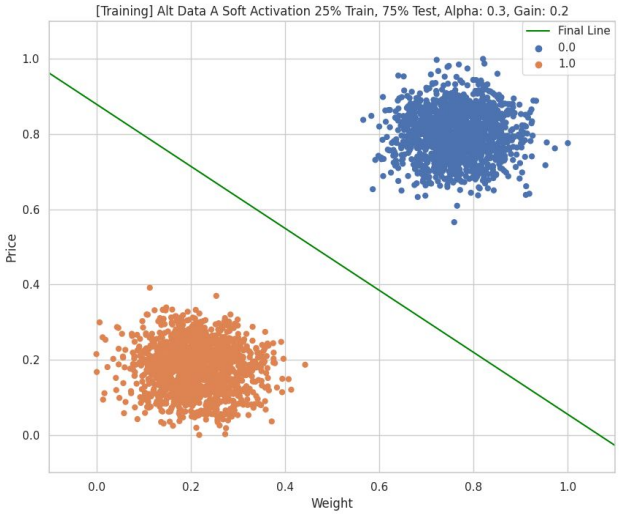

Group C Soft Activation 25% Train, 75% Test, Alpha: 0.1, Gain: 0.1
352
150
327
171
0.0
1.0
Actual
0
1
Predicted
350
325
300
275
250
225
200
175
150

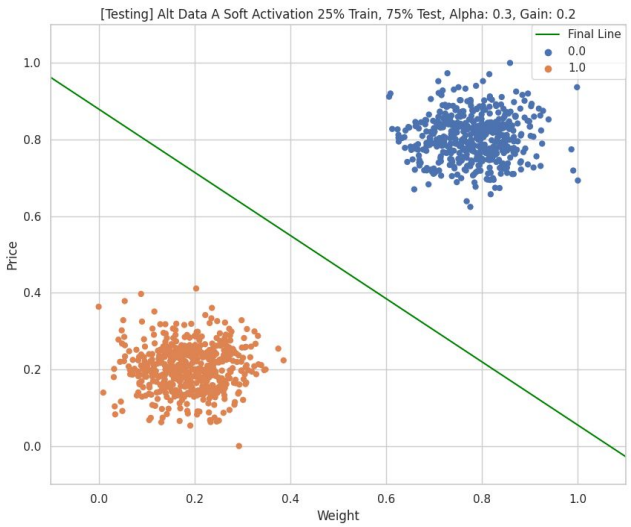
Group A Alt

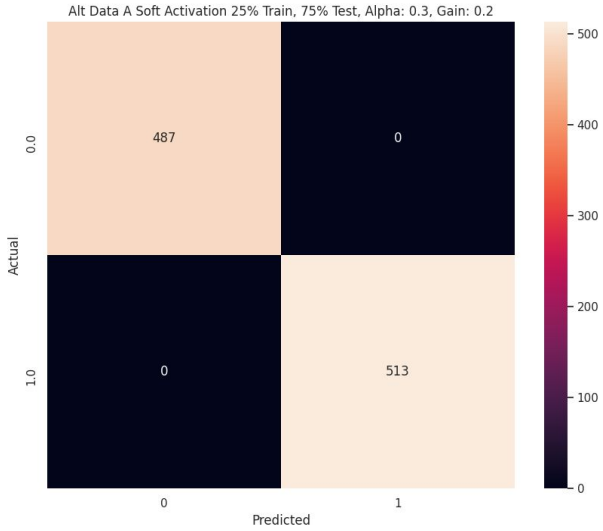

[Training] Alt Data A Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0


[Testing] Alt Data A Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

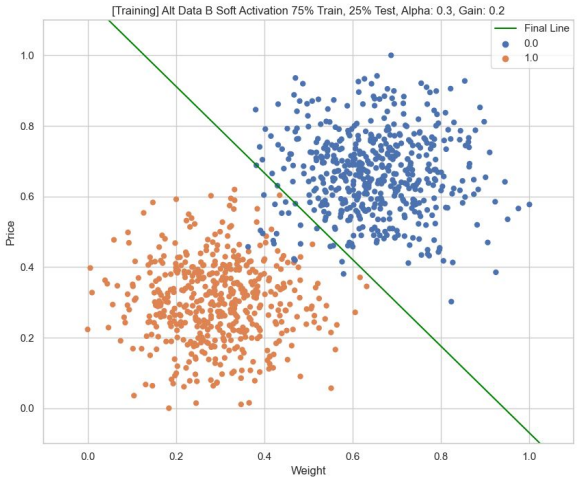

Alt Data A Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
487
0
0
513
0.0
1.0
Actual
0
1
Predicted
500
400
300
200
100
0

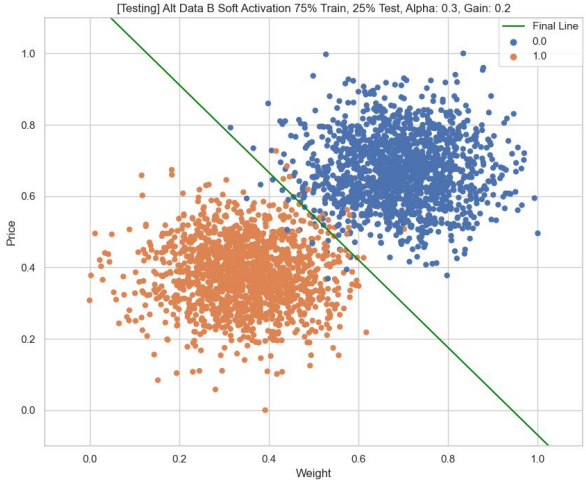

[Training] Alt Data A Soft Activation 25% Train, 75% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

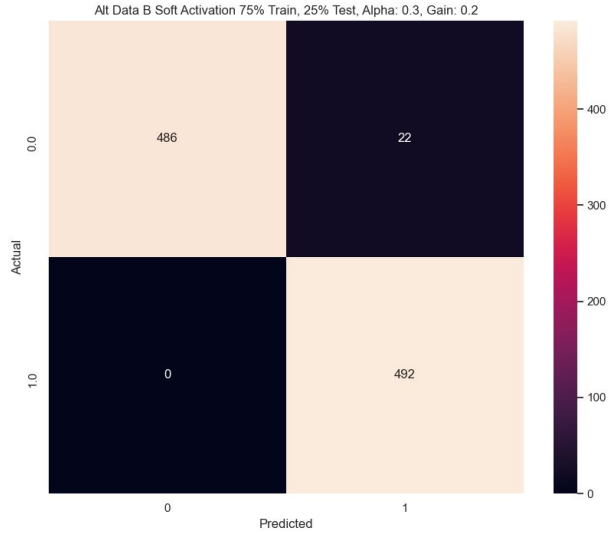

[Testing] Alt Data A Soft Activation 25% Train, 75% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

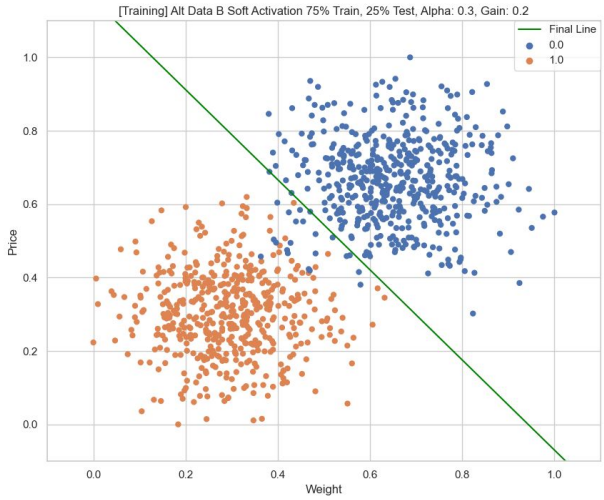

Alt Data A Soft Activation 25% Train, 75% Test, Alpha: 0.3, Gain: 0.2
487
0
0
513
Actual
Predicted
0.0
1.0
0
1
500
400
300
200
100
0

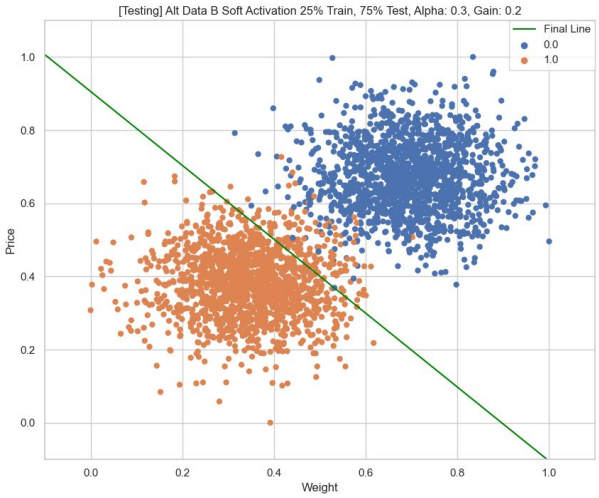
Group B Alt

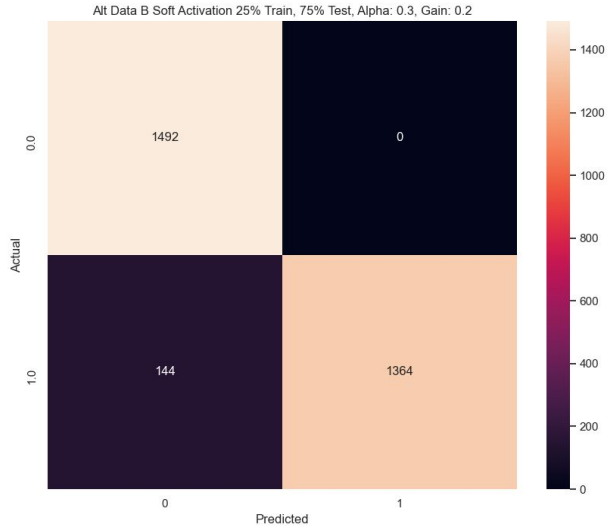

[Training] Alt Data B Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0


[Testing] Alt Data B Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

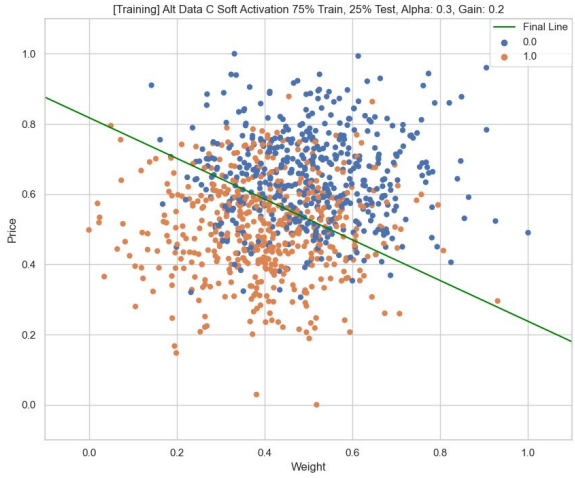

Alt Data B Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
486
22
0
492
0.0
1.0
Actual
0
1
Predicted
400
300
200
100
0

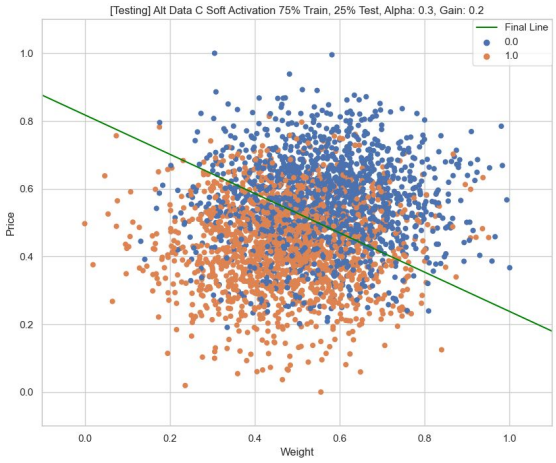

[Training] Alt Data B Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

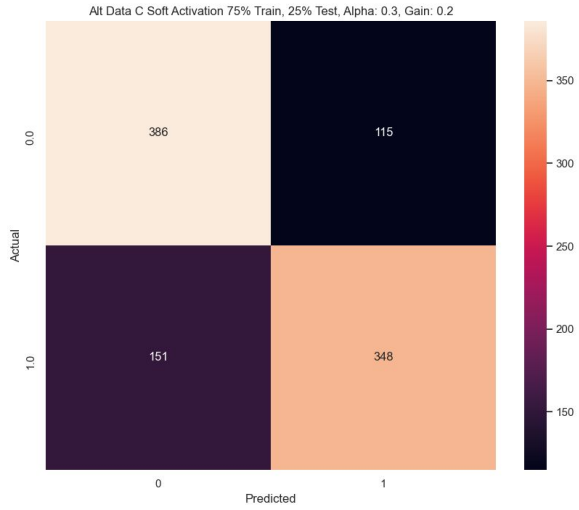

[Testing] Alt Data B Soft Activation 25% Train, 75% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0

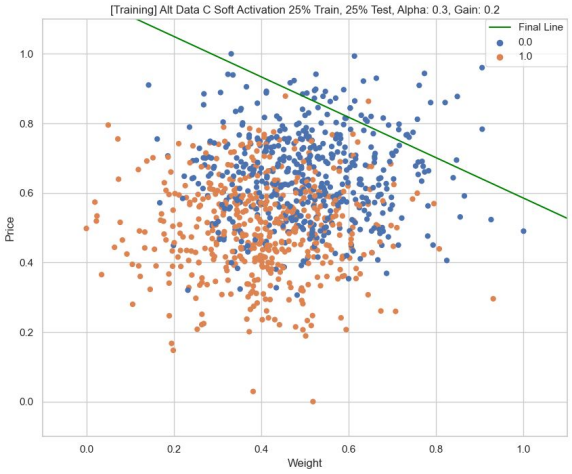

Alt Data B Soft Activation 25% Train, 75% Test, Alpha: 0.3, Gain: 0.2
1492
0
144
1364
Actual
Predicted
0.0
1.0
0
1
1400
1200
1000
800
600
400
200
0

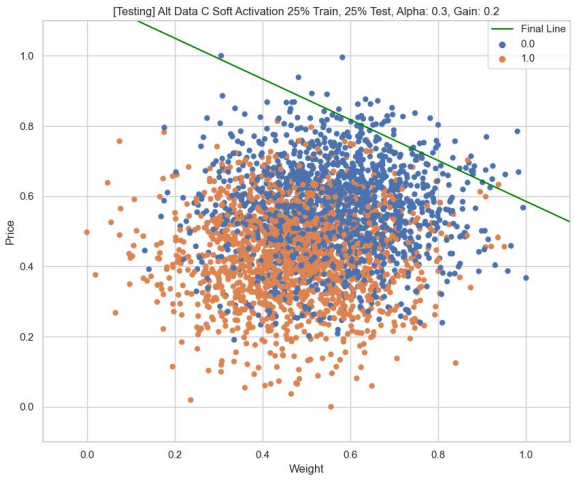
Group C Alt

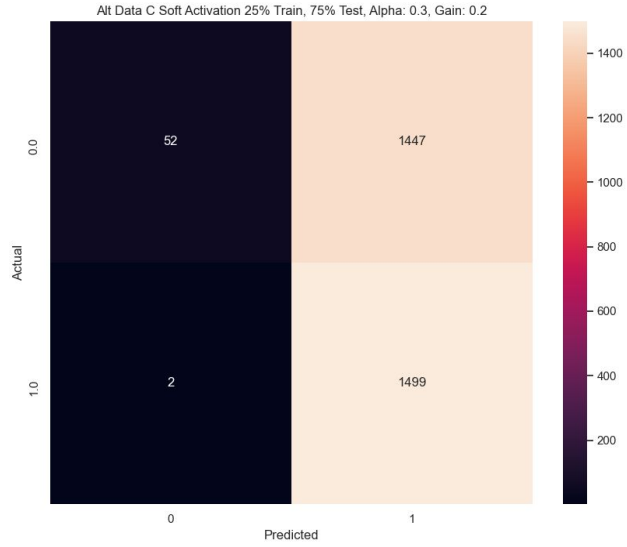

[Training] Alt Data C Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0


[Testing] Alt Data C Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0


Alt Data C Soft Activation 75% Train, 25% Test, Alpha: 0.3, Gain: 0.2
386
115
151
348
0.0
1.0
Actual
0
1
Predicted
350
300
250
200
150


[Training] Alt Data C Soft Activation 25% Train, 25% Test, Alpha: 0.3, Gain: 0.2
Final Line
0.0
1.0
Price
Weight
0.0
0.2
0.4
0.6
0.8
1.0





RESULTS:

Graph A:

Hard:

75:

Test Error: 0

Total Error: 0

Comparison: Compared to assignment 1, these results are almost identical, because this data set is easily divided by a decision line. The testing and total error is 0 because all points are classified correctly.

25:

Test Error: 0

Total Error: 0

Comparison: Compared to assignment 1, these results are almost identical because this data set is easily divided by a decision line. The testing and total error is 0 because all points are classified correctly.

75 v. 25: For Group A with a hard activation function, the 75% training vs 25% training made no difference in the outcome, even when randomly selected, and shuffling the data before sampling. This data shows 2 distinct clusters and is easily separated.

Soft:

75:

Test Error: 0

Total Error: 0.000107

Comparison: Assignment 1 had no soft activation to compare directly to, but the resultant decision line of the soft activation function is almost identical to that of the manual line drawn. While training, the total error while training reached 0.000107, however the decision line still splits the data perfectly. The error with the soft activation is higher than the hard activation because the output of the soft activation function will never be 0. The test error is 0 because all points are classified correctly.

25:

Test Error: 0

Total Error: 0.000189

Comparison: Assignment 1 had no soft activation to compare directly to, but the resultant decision line of the soft activation function is almost identical to that of the manual line drawn. While training, the total error while training reached 0.000189, however the decision line still splits the data perfectly. The error with the soft activation is higher than the hard activation because the output of the soft activation function will never be 0. The test error is 0 because all points are classified correctly.

75 v. 25: Changing from a hard to soft activation function made no difference with Group A. Since this data is cleanly separated, it's easy for the algorithm to determine a separator. Once again, the 75% split vs the 25% split made no difference in this dataset.

Soft v. Hard: For both the 75% and 25% training split, for this dataset, the use of a hard or a soft activation function made little difference. The biggest difference between the two was that the Soft activation function ran slightly slower than the Hard activation function, since it converges at a slower rate

Graph Alt A:

Hard:

75:

Test Error: 0

Total Error: 0

Comparison: Just as in the GroupA dataset, the decision line for this data was drawn almost identically to that of assignment 1. With disjoint clusters of data it is easy to draw a separator line between the clusters. Again, the test error is 0 because all points are classified correctly.

25:

Test Error: 0

Total Error: 0

Comparison: The 25% split of Group Alt A resulted in an almost identical line as to that from project 1. Again, the test error is 0 because all points are classified correctly.

75 v. 25: Comparing the 75% training to the 25% training, there is little difference. The distribution of training and testing data had no effect on this data set since the data forms 2 distinct clusters.

Soft:

75:

Test Error: 0

Total Error: 0.000013

Comparison: The soft activation function for 75% training split produced a near identical line to that of project 1. The total error is non-zero because the output of the soft activation function can never be 0. The test error is 0 because all data points are identified correctly.

25:

Test Error: 0

Total Error: 0.000013

Comparison: The soft activation function for 25% training split produced a near identical line to that of project 1. The total error is non-zero because the output of the soft activation function can never be 0. The test error is 0 because all data points are identified correctly.

75 v. 25: Comparing the 75% training to the 25% training when using the soft activation function still shows little difference. The distribution of training and testing data had no effect on this data set since the data forms 2 distinct clusters.

Soft v. Hard: For both the 75% and 25% training split, for this dataset, the use of a hard or a soft activation function made little difference. The biggest difference between the two was that the Soft activation function ran slightly slower than the Hard activation function, since it converges at a slower rate.

Graph B:

Hard:

75:

Test Error: 33.100000%

Total Error: 1121

Comparison: This did worse than the accuracy from project 1, but not by much. However, this had a much better true positive result, but worse false positive result.

25:

Test Error: 28.333333%
Total Error: 338
Comparison: This had almost the same error that the project 1 dataset B had, with a higher true positive rate and almost exact true negative rate. This is more similar than 75% trained group B.

75 V. 25: In this instance, 25% train actually performed better than the 75% train. This could be due to the fact that 75% training contributed to over training the model, or that the 25% training had less noise in the data, resulting in a more pure model.

Soft:

75:

Test Error: 25.3%
Total Error: 559.01337
Comparison: This result gave a better result than the result from project 1, having a drop in the error rate by about 2%. This is probably due to the type of activation function as well as the amount of data being trained on.

25:

Test Error: 27.6%
Total Error: 670.166815
Comparison: This result gave a similar result to that from project 1, having the same basic error rate. Additionally, the true positive and true negative rate were consistent with each other, hovering around 28% each.

75 V. 25: Here the 75 percent train performed slightly better than the 25 percent train. This is probably because training on more data typically yields better results, and the soft activation function has more gradual convergence than the hard, allowing for more nuanced training function.

Soft V. Hard: Overall, the soft activation function performed much better than the hard activation function, reducing the error by almost 5%. This is primarily due to the soft activation function's better capability with noisy, overlapping data.

Graph Alt B:

Hard:

75:

Test Error: 1.2%
Total Error: 77
Comparison: This perceptron had exactly the same performance as the corresponding model in project 1. Because the data had less noise, this perceptron was able to find the division between the two classes with relative ease.

25:

Test Error: 16.46666%

Total Error: 59

Comparison: This data had less capability than the corresponding data from project 1. Training on less data and testing on more typically should yield a worse result if the data isn't noisy (which in this case it's not noisy). Thus, it just didn't have enough data to actually get to the final result needed.

75 V. 25: These had significantly different results. Since the 75% train had more data to train on, and the data wasn't noisy, it had a better context of optimizing the perceptron's performance in comparison to that of the 25% train.

Soft:

75:

Test Error: 2.2%

Total Error: 74.978599

Comparison: This had a worse performance to that of the corresponding project 1 dataset. Since we drew the lines from last time, this could be purely coincidental, since out level of precision was not on scale with this neuron's. So I suspect by coincidence that our decision line did better. It is worth noting that all the error percentage for this model was grouped into false negative rate. This indicates that the perceptron completely excluded one side of the data set.

25:

Test Error: 4.8%

Total Error: 67.79312

Comparison: This also had worse performance to that of the corresponding project 1 dataset, but it still had a decent decision rate. Further, it is more similar to the 75% train than the Hard activation 25% train was. This indicates that the soft activation function outperformed the hard activation function in total error.

Soft V. Hard: In this instance, the best performing perceptron was trained using the hard activation function, however, the soft activation function performed much better on average overall. The soft activation function had a better overall performance because of the more gradual convergence as opposed to the hard function, allowing a more consistent training, eventually converging to a decent result.

Graph C:

Hard:

75:

Test Error: 50 %

Total Error: 1460.000000

Comparison: For the 75% training split of GroupC, the test error is only slightly worse than that of project 1 (49.85%) However, the decision line that was drawn is considerably worse. The decision line that was drawn by the hard activation function barely touches the graph of the points, as opposed to the manually drawn line in project 1. We believe this to be in part due to the random assignment of starting weights. Had we started with weights corresponding to the y=-x+1line, then the resulting line probably would have intersected the data more. The test error

is 50% because the algorithm is essentially classifying all points as one class. We can tell from the Total Error from training that the decision line is not performing well. This poor performance is likely due to the data being noisy, and not forming an easily separable pattern, even to the human eye.

25:

Test Error: 49.933333 %
Total Error: 530.000000

Comparison: When running the algorithm on the 25% split, it was rare that we got a decision that intersected the data points at all. The Test Error is pretty much 50% because the decision line is labelling all of the points as 1 class. Once again, we believe this is due to the distribution of the data points not forming any sort of distinguishable pattern. We believe the lower Total Error in comparison to the 75% split is due to the reduced number of data points. Since the algorithm iterates over less points, the sum of the squared error will be less even if they exhibit the same amount of error.

Soft:

75:

Test Error: 47.7 %
Total Error: 750.180790

Comparison: With the soft activation function, the generated decision line is less steep than the line we drew in project 1, but instead of missing the graph entirely as the hard activation function did, this line almost passes through the middle of the data. The test error achieved was less than the 49% from project 1.

25:

Test Error: 47.7 %
Total Error: 750.180790

Comparison: The 25% and 75% splits actually drew an identical decision line. This could be due to the weights being assigned or the data simply pushing the line to the same place. Once again, as opposed to the hard activation function, this line almost passes through the middle of the graph, and has less error than the line we drew in project 1.

75 V. 25: With this dataset, the difference in training volume made little difference to the result. The dataset shows no distinguishable pattern. While the errors are similar, the decision line of the 75% hard function barely touches the graph, and the hard function for the 25% split missed the data entirely.

Soft v Hard: The soft activation function clearly performs better than the hard activation function for both splits of data. The soft activation function drew a decision line through almost the middle of the data points, whereas the har activation function almost missed, or missed the data entirely.

Graph Alt C:

Hard:

75:

Test Error: 29.2 %

Total Error: 1118.000000

Project 1 and 75 v 25 Comparison: The 75% trained hard activation did slightly worse compared to Group C from Project 1 as the Project 1's error was 28.2%. This makes sense because we trained on 75% of the data and only tested on 25%. The 25% trained hard activation function did worse than the 75 due to training on less data. The 25% line is barely on the graph.

25:

Test Error: 49.93 %

Total Error: 384.000000

Project 1 and 75 v 25 Comparison: The 25% trained hard activation did much worse than the 75% trained one and Project 1 - Group C's implementation. Training on less data and testing it against more for hard activation is going to generate worse results since there is less representative data.

Soft:

75:

Test Error: 26.6 %

Total Error: 557.504019

Project 1 and 75 v 25 Comparison: The 75% trained soft activated did a little better than Project 1 Group C. The soft activated's growth is not as steep as the hard activation. Compared to the 25% trained soft, it performs better as there is more data being trained from, so the learning is more accurate.

25:

Test Error: 48.3 %

Total Error: 237.229658

Project 1 and 75 v 25 Comparison: The 25% trained soft activation performed worse than Project 1 Group C's implementation. This is due to the fact that it was trained on a small subset of the data and tested against a much larger set. The 25% trained method performed worse than 75% trained method for Alt Group C because of this same reason.