CMSC 409: Artificial Intelligence

Project No. 2

Due Thursday, October 3rd, 2019, noon

Student certification:

Team member 1:

Print Name: Peter J George

Date: 10/03/2019

I have contributed by doing the following: Wrote the hard/soft activation functions, equations, and plotted data for training sets. Normalized data and graph parameters such as label, point size, etc.

Signed:

Team member 2:

Print Name: Daniel Webster

Date: 10/30/2019

I have contributed by doing the following:

implemented testing function that test's error of training set and plotted the tested data set with the perceptron

Signed :

Team member3:

Print Name: Joseph Longo

Date: 10/2/19

I have contributed by doing the following: debugging code, created the project write-up.

Joseph Longo

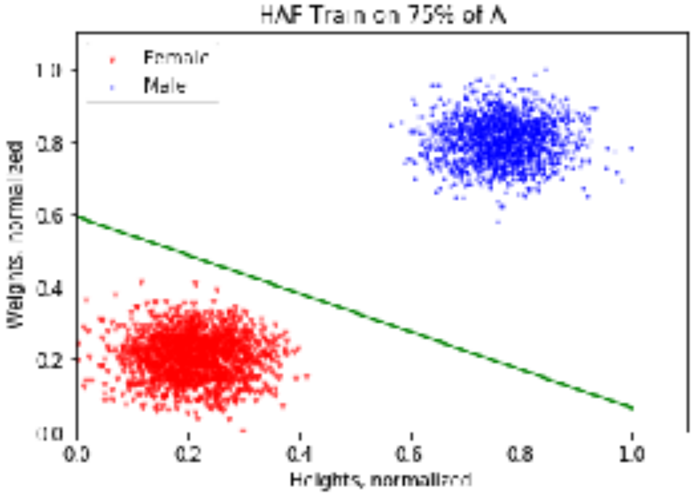
Scenario A)

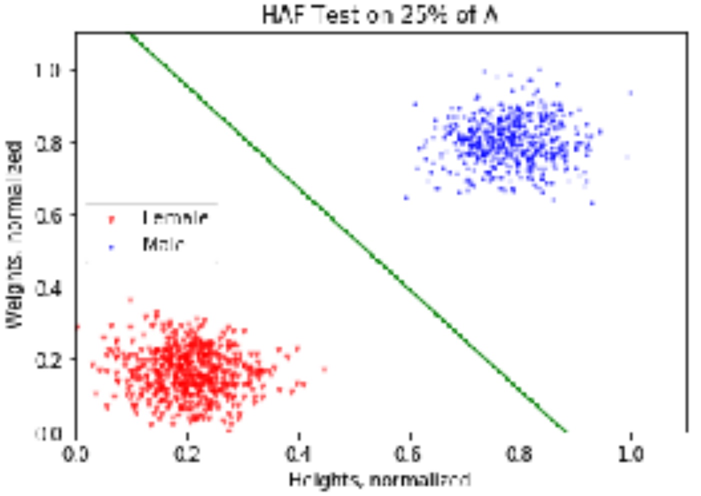
Hard Activation Function

1.) 75% of the data is chosen for training

25% chosen for testing

Dataset A





Error: 0

Error: 0

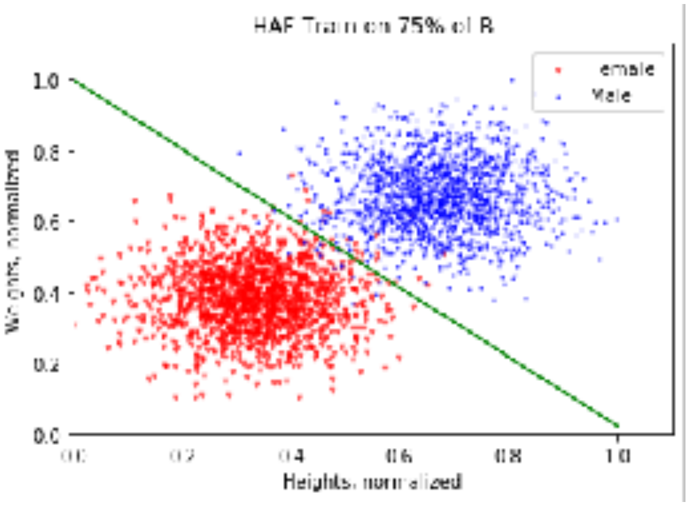
Scenario A)

HAF

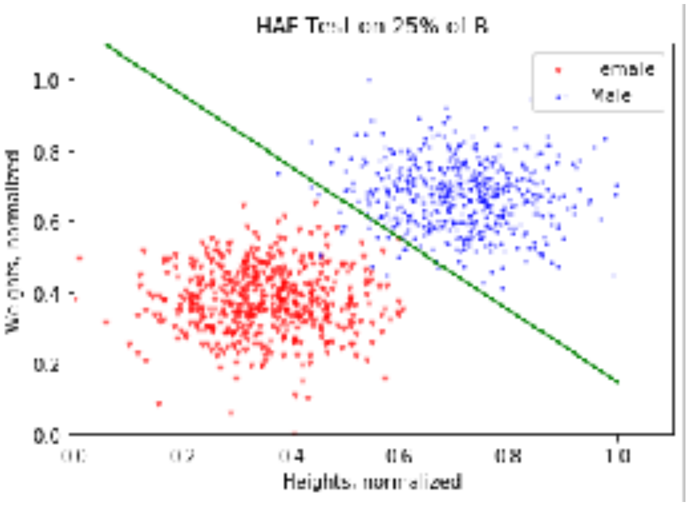
1.) 75% of Data is Trained

25% of Data is Tested

Dataset B:



Error: 96/3000 OR .032



Error: 38/1000 OR .038

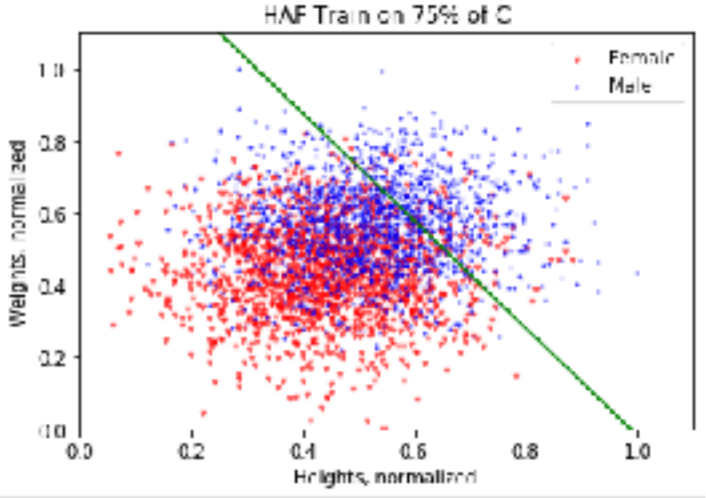
Scenario A)

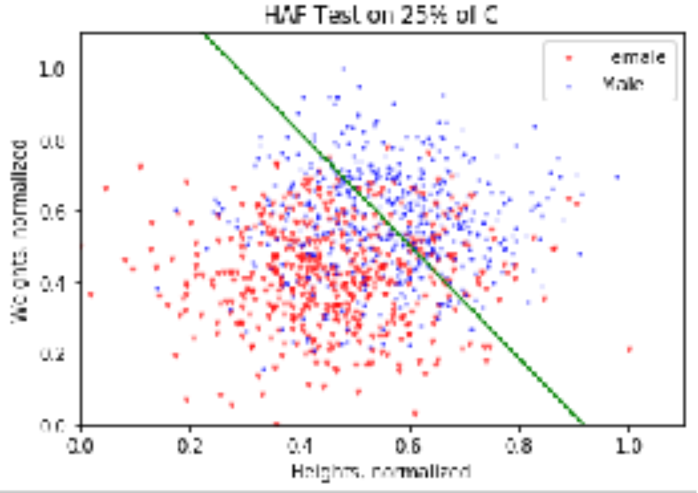
HAF

1.) 75% of Data is Trained

25% of Data is Tested

Dataset C:





Error: 1444/3000 OR .4813

Error: 298/1000 OR .298

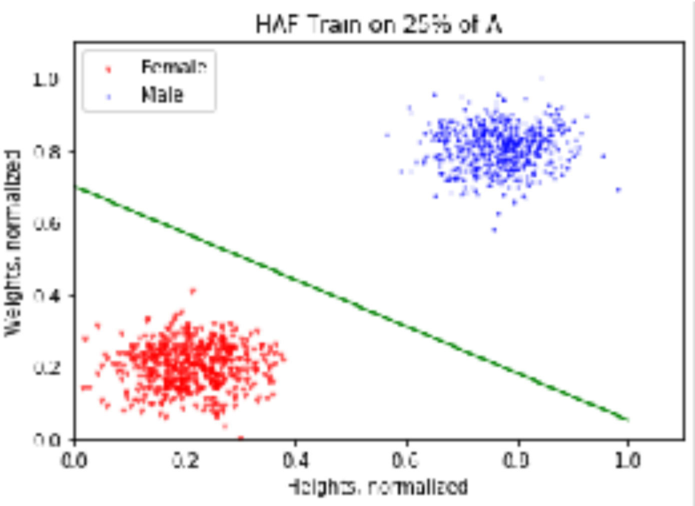
Scenario A)

HAF

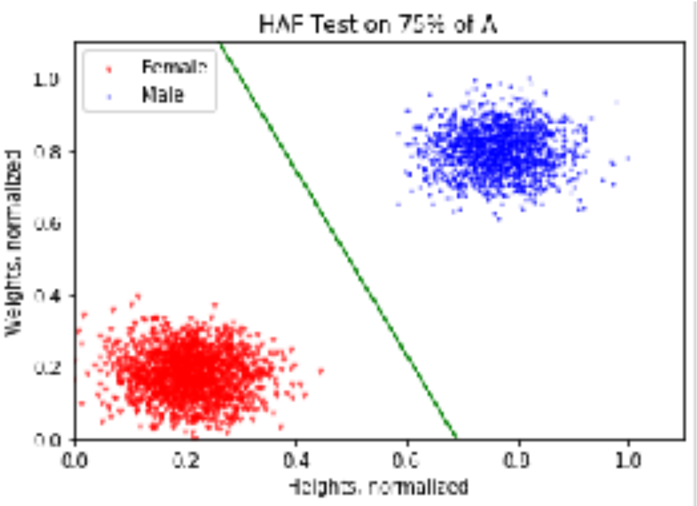
2.) 25% of Data is Trained

75% of Data is Tested

Dataset A:



Error: 0



Error: 0

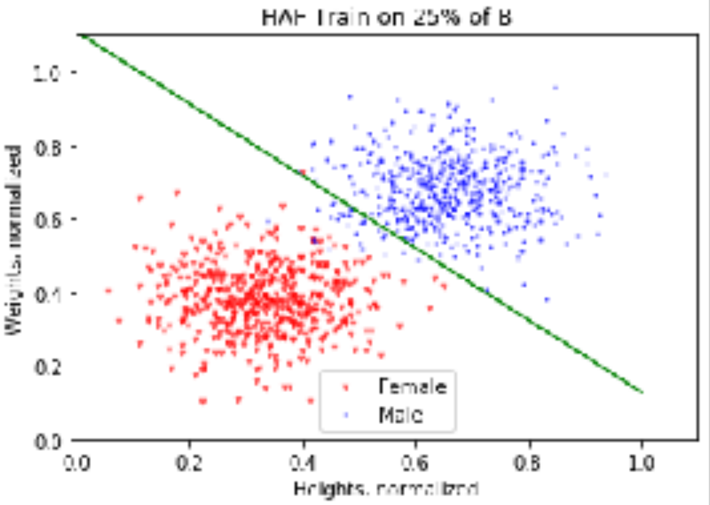
Scenario A)

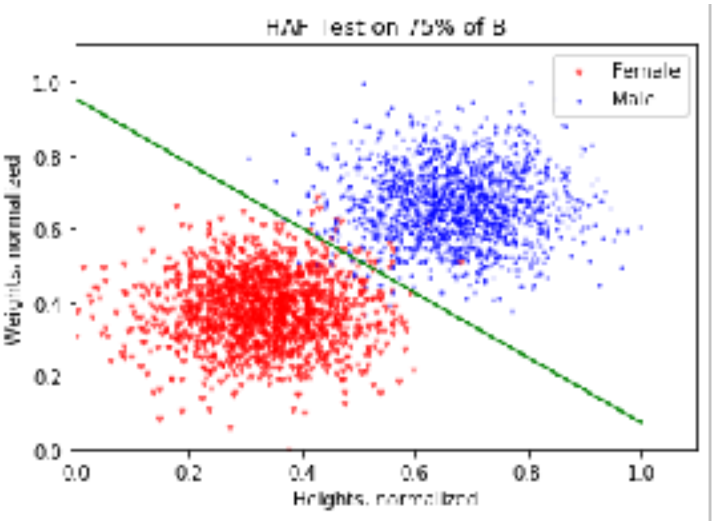
HAF

2.) 25% of Data is Trained

75% of Data is Tested

Dataset B:





Error: 59/1000 OR .059

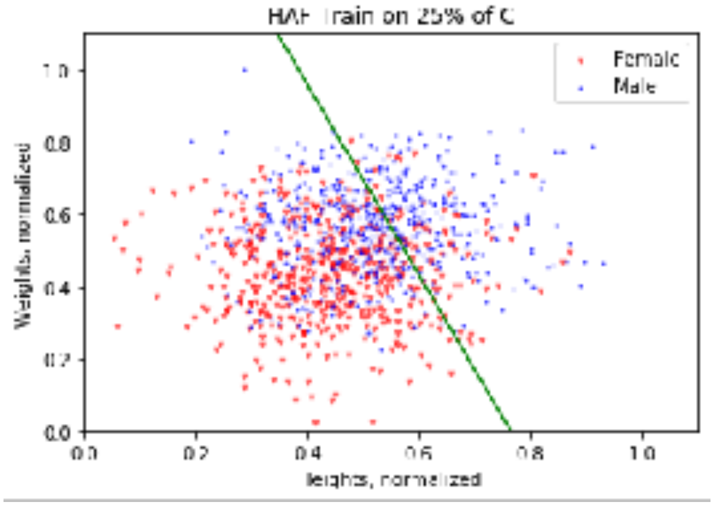
Error: 40/3000 OR .013

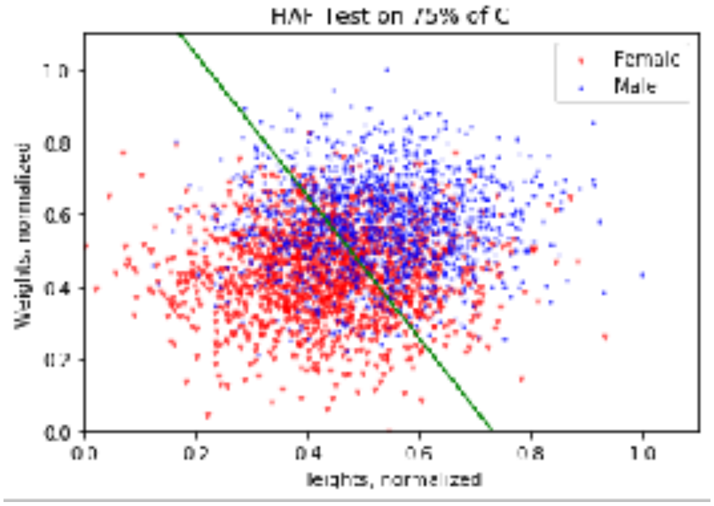
Scenario A)

HAF

2.) 25% of Data is Trained 75% of Data is Tested

Dataset C:





Error: 521/1000 OR .521 Error: 821/3000 OR .2736

3.) Discuss Results:

When comparing the results from Scenario 1 and Scenario 2 you can actually see that the errors, while slightly different, are actually quite similar for all three of the datasets provided. Perhaps this is just a result of the Hard Activation Function. What can be seen quite clearly is that as the points in a dataset get closer and closer, the errors for both the training and testing sets rise significantly. However, when you look at the errors for each scenario, the amount of errors in the testing group is actually less when you train from only 25% of the data. Since the activation function is hard, it might be beneficial to train on a smaller group so that you have more points in your test group to lower the error percentage.

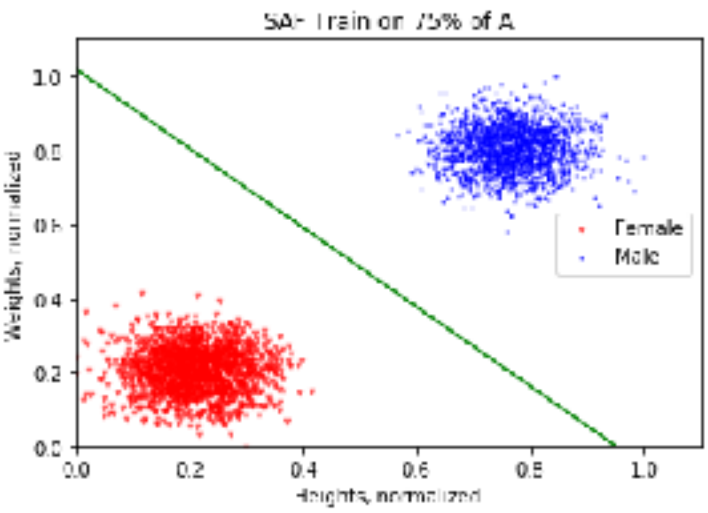
Scenario B)

SAF

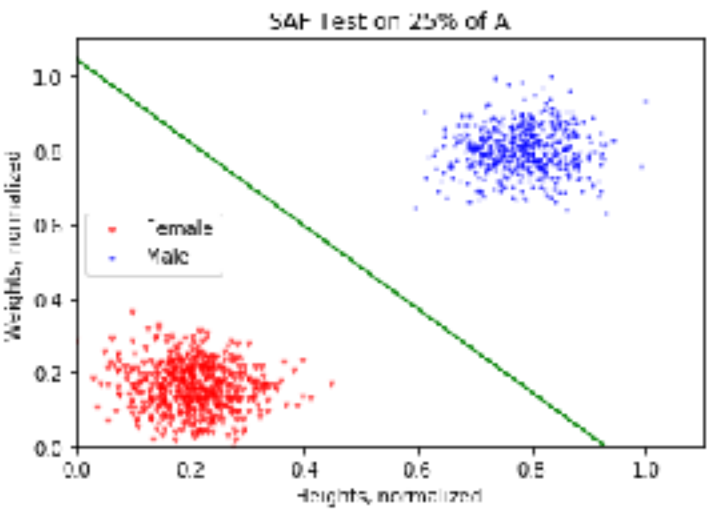
1.) 75% of Data is Trained

25% of Data is Tested

Dataset A:


SAF Train on 75% of A

Error: 0


SAF Test on 25% of A

Error: 0

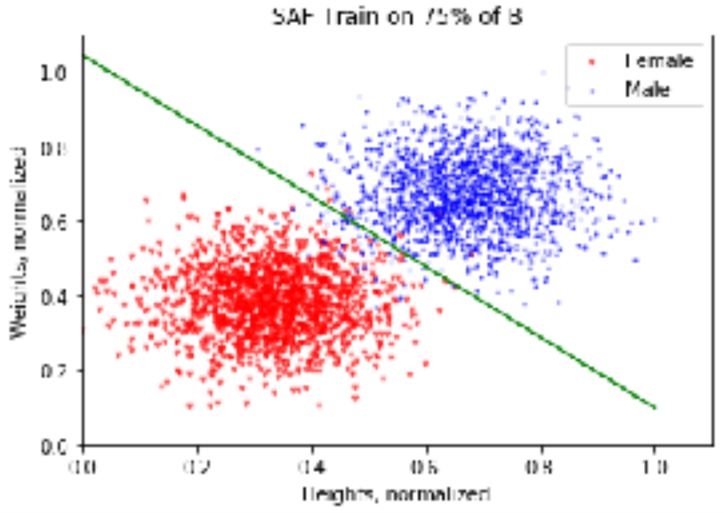
Scenario B)

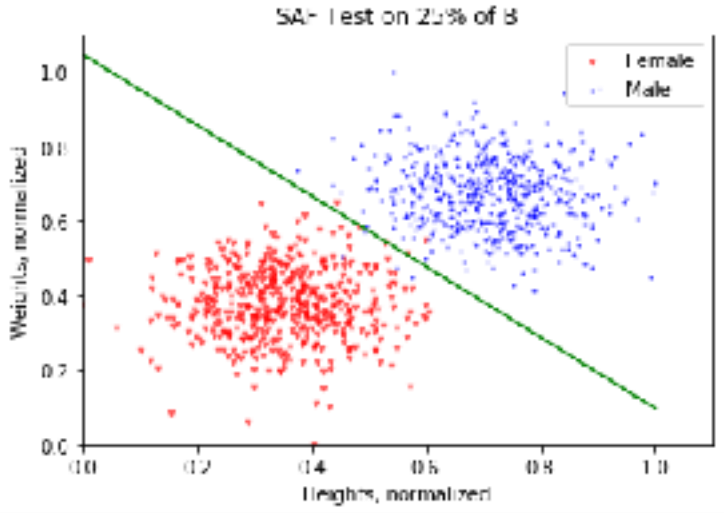
SAF

1.) 75% of Data is Trained

25% of Data is Tested

Dataset B:





Error: 72/3000 OR .024

Error: 7/1000 OR .007

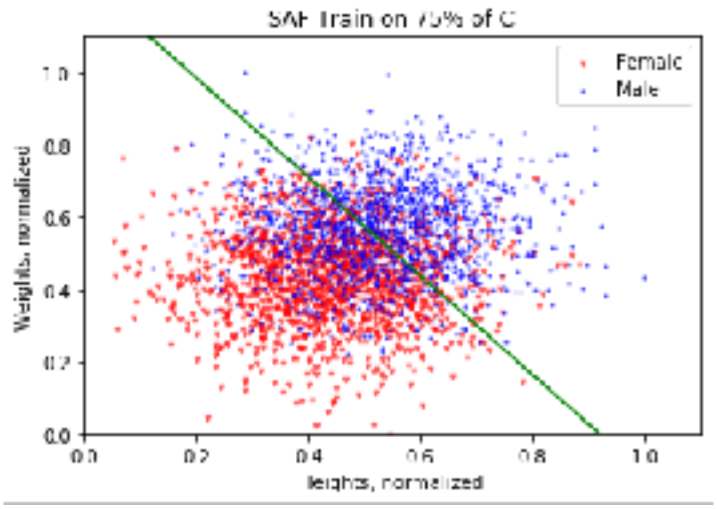
Scenario B)

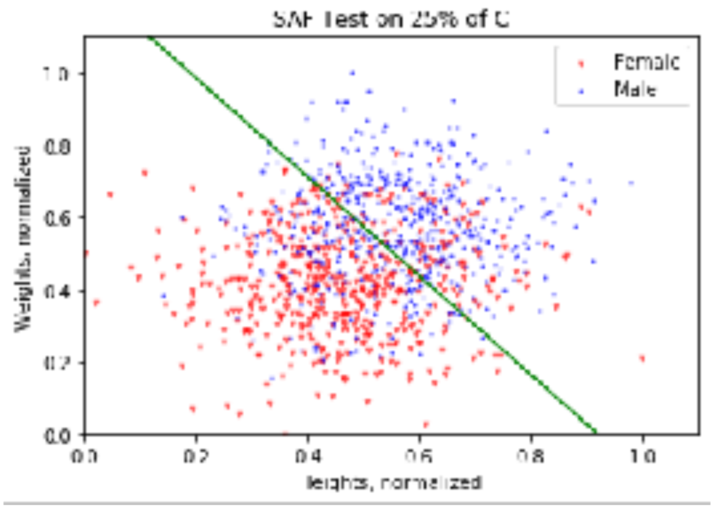
SAF

1.) 75% of Data is Trained

25% of Data is Tested

Dataset C:





Error: 998/3000 OR .3326

Error: 287/1000 OR .287

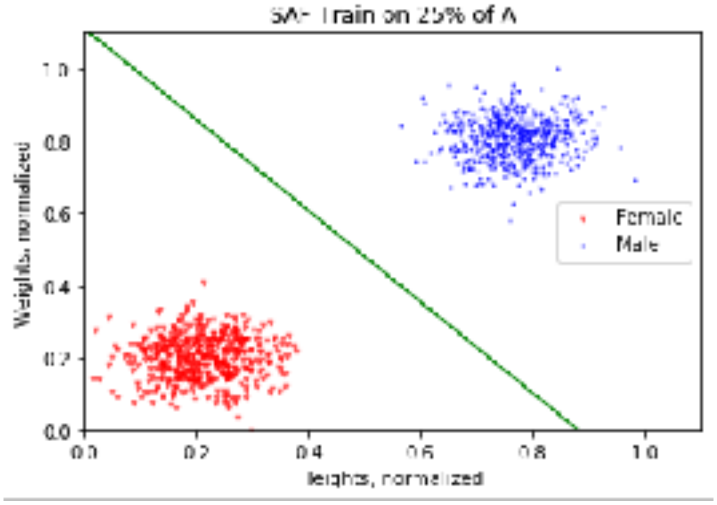
Scenario B)

SAF

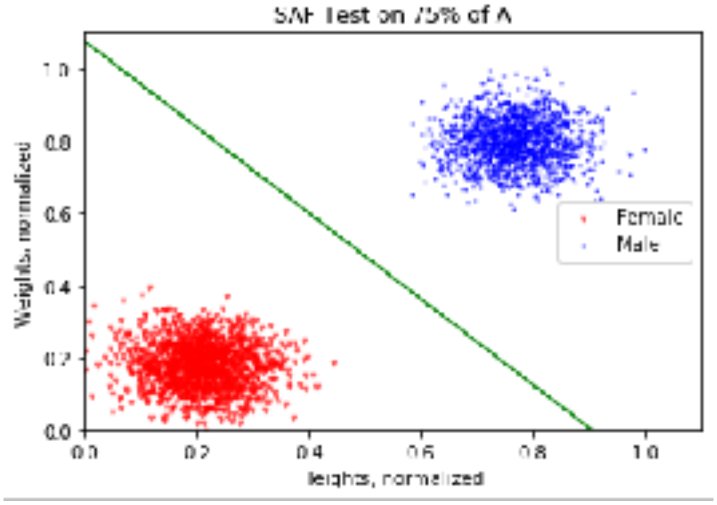
2.) 25% of Data is Trained

75% of Data is Tested

Dataset A:



Error: 0



Error: 0

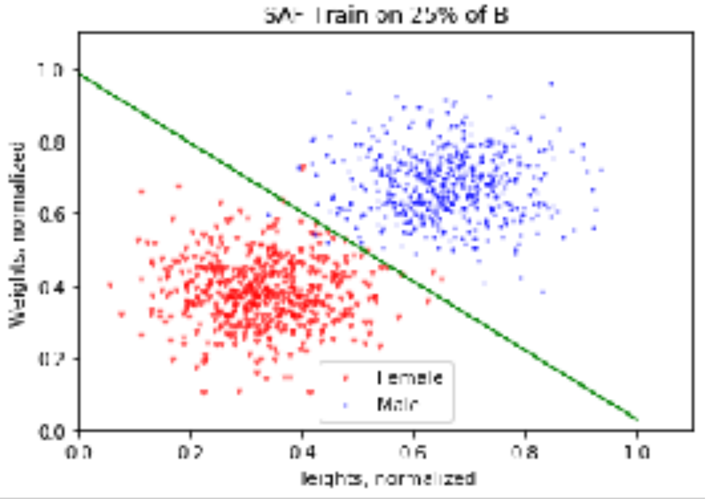
Scenario B)

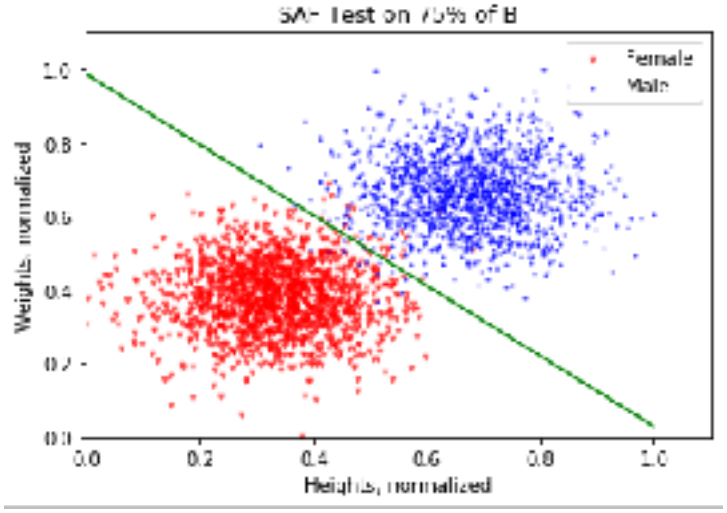
SAF

2.) 25% of Data is Trained

75% of Data is Tested

Dataset B:





Error: 40/1000 OR .04

Error: 39/3000 OR .013

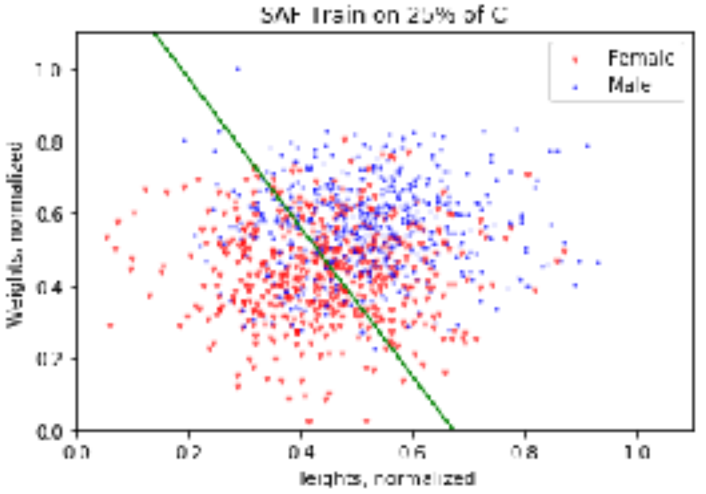
Scenario B)

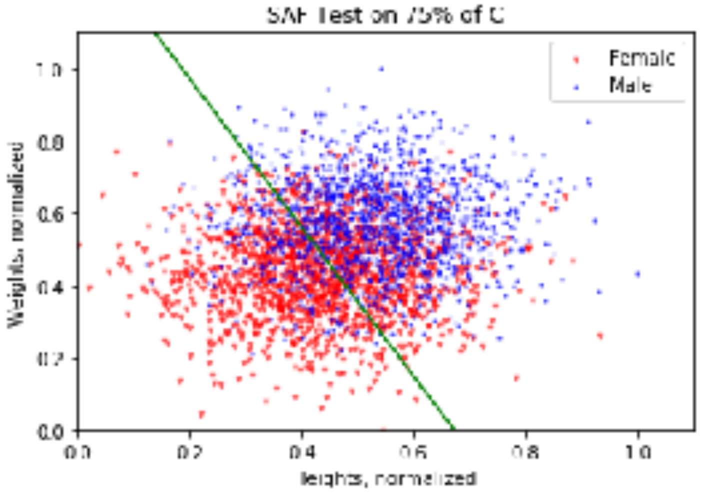
SAF

2.) 25% of Data is Trained

75% of Data is Tested

Dataset C:





Error: 365/1000 OR .365

Error: 874/3000 OR .291

3.) Discuss Results:

Similarly when looking at the errors and graph plots for the datasets, you can see that changing between Scenario 1 and Scenario 2 didn’t have the greatest impact, however it did have one. You can see that when you train on 75% of the data, the amount of errors both in the training set and the testing set are significantly reduced compared to the amount of errors when only 25% of the data is used to train. For a soft activation function, it makes perfect sense to train on a larger amount of data; it provides less errors in every scenario that we tested.

Pr2.2 Soft vs Hard Activation:

Hard Activation:

When a hard activation function was used, the amount of errors produced when training on 75% of the data vs training on 25% of the data varied for each of the 3 datasets.

For dataset A, amount of errors produced did not change in response to the amount of data that was trained on.

For dataset B, the percentage of errors produced actually changed quite a bit in response to the amount of data that was trained on. The error percentage dropped quite a bit when only training on 25% of the data. The error for testing on 25% of the data was .038 while the error for testing on 75% of the data was only .013.

For dataset C, similarly, the percentage of errors produced dropped when only training on 25% of the data. When testing on 25% of the data, the error was .298. When testing on 75% of the data, the error was .2736.

Soft Activation:

When the soft activation function was used, it was very clear that the amount of errors produced dropped drastically when training on 75% of the data.

For dataset A, amount of errors produced did not change in response to the amount of data that was trained on, as there were no errors between them.

For dataset B, when training on 75% of the data the error for the training set was .024, while the error for the testing set was only .007. When training on 25% of the data the error for the training set was .04 while the error for the testing set was .013. Clear advantage for training on a larger data set.

For dataset C, when training on 75% of the data the error for the training set was .3326, while the error for the testing set was only .287. When training on 25% of the data the error for the training set was .365, while the testing set was .291. Again, advantage for training on a larger data set.

In conclusion, it’s very clear that if your aim is to reduce the total amount of errors then you should train a soft activation function on a large portion of the dataset (in this case 75%).