

Homework 1: Naive Bayes

After using a combination of the Naive Bayes Algorithm and the N-Fold (leave-one-out cross), I was able to get the following results:

Actual	Predicted	probSoft	probHard	probHard
none	none	0	0	1
soft	soft	0.58931377684 6517	0	0.41068622315 3483
none	none	0	0	1
hard	hard	0	0.77002053388 0903	0.22997946611 9096
none	none	0	0	1
soft	soft	0.71519389701 2079	0	0.28480610298 7921
none	none	0	0	1
hard	none	0	0	1
none	none	0	0	1
soft	soft	0.53444180522 5653	0	0.46555819477 4347
none	none	0	0	1
hard	none	0	0	1
none	none	0	0	1
soft	soft	0.66765578635 0148	0	0.33234421364 9852
none	none	0	0	1
none	hard	0	0.63636363636 3636	0.36363636363 6364
none	none	0	0	1
none	soft	0.75301866081 2294	0	0.24698133918 7706

none	none	0	0	1
hard	none	0	0	1
none	none	0	0	1
soft	none	0	0	1
none	none	0	0	1
none	hard	0	0.5833333333333333	0.4166666666666667

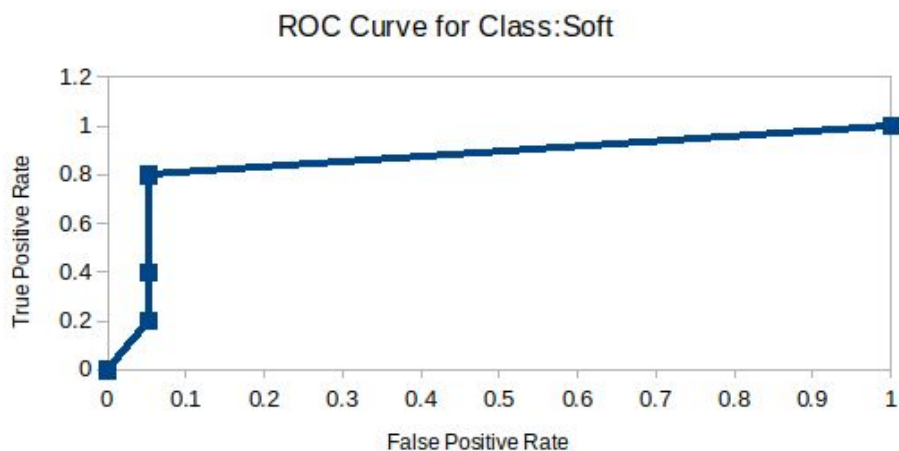
With this information, I was able to create a confusion matrix and determine the accuracy of my algorithm:

	Predicted Soft	Predicted Hard	Predicted None
Actual Soft	4	0	1
Actual Hard	0	1	3
Actual None	1	2	12

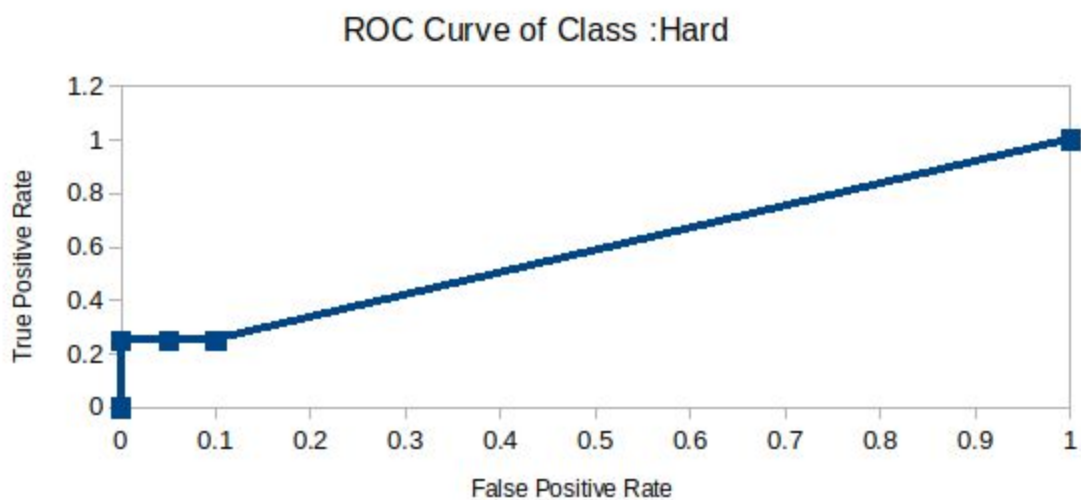
Where the each row represents the actual categories and the columns represents the predicted categories. The calculated accuracy is 0.8055555555555556.

After determining this information, I was able to create the ROC curve and AUC for each class:

Soft class ROC with AUC= 0.957894736842105:



Hard class ROC with AUC=0.925:



None class ROC with AUC=0.874074074074074 :

