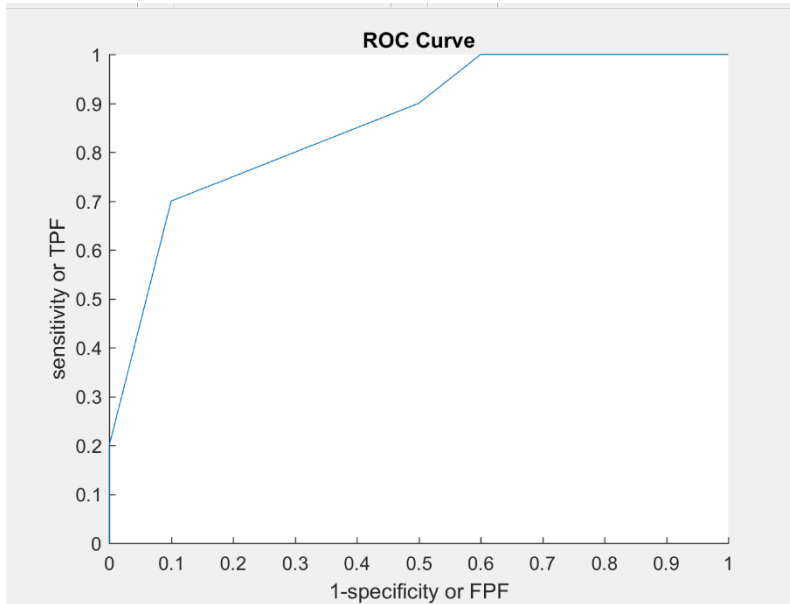


## BIOINFORMATICS PROJECT

### RESULTS:



threshold = 2  
matrix =  
10 6  
0 4  
accuracy = 0.7000  
prevalence = 0.5000  
sensitivity = 1  
specificity = 0.4000  
PPV = 0.6250  
NPV = 1  
area = 0.8600

### Conclusion:

With the best threshold being at the rating 2, the samples were accurately predicted 70% of the time. The curve matches an ideal ROC curve and the area under the curve of 0.86 shows that it was good, but not excellent in predicting the disease. According to the sensitivity value of 1.0, my ratings were very accurate at correctly predicting if there was a disease; however, with the specificity value of 0.4, my ratings were significantly less accurate at correctly predicting if there wasn't a disease. Furthermore, the positive predictive value of 0.625 shows that my ratings predicted more samples as having the disease than there were. On the other hand, the negative predictive value of 1.0 shows that none of the samples that truly had the disease were predicted to not have the disease. Overall, since the threshold is 2 meaning most samples were counted as predictive truth, it showed that more values were predicted as having the disease than there truly were. Because this threshold is on the lower end, it allowed for more correct predictions about samples truly having the disease, but also created more error and predicted more samples as having the disease even if they truly didn't. While there are some tradeoffs about this threshold value, it is more accurate and has less error than other threshold values, and therefore this threshold is the best one to use for this test.