

4. I looked at the importance of variables to find if there were any that significantly helped in reducing the misclassification error rate for validation during the multiple 10-fold CV. I found that many different models selected a few different variables each, with only a few being common throughout all the models. This helped my analysis as I found the B class was the most difficult to decipher of all classes. There were only a select few variables used in separating B from other classes so I tried to target these variables to improve the B classification aside from models like Random Forest trees and found only marginal improvements.