

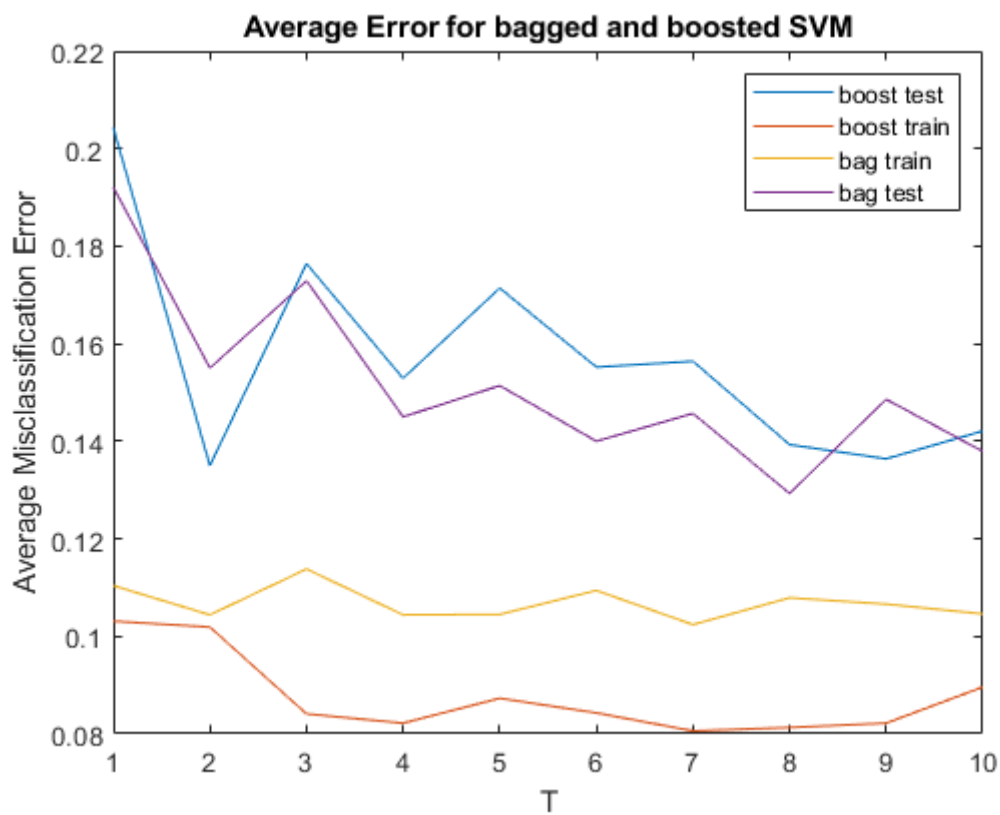
HW 10

Adam Karl

April 15, 2021

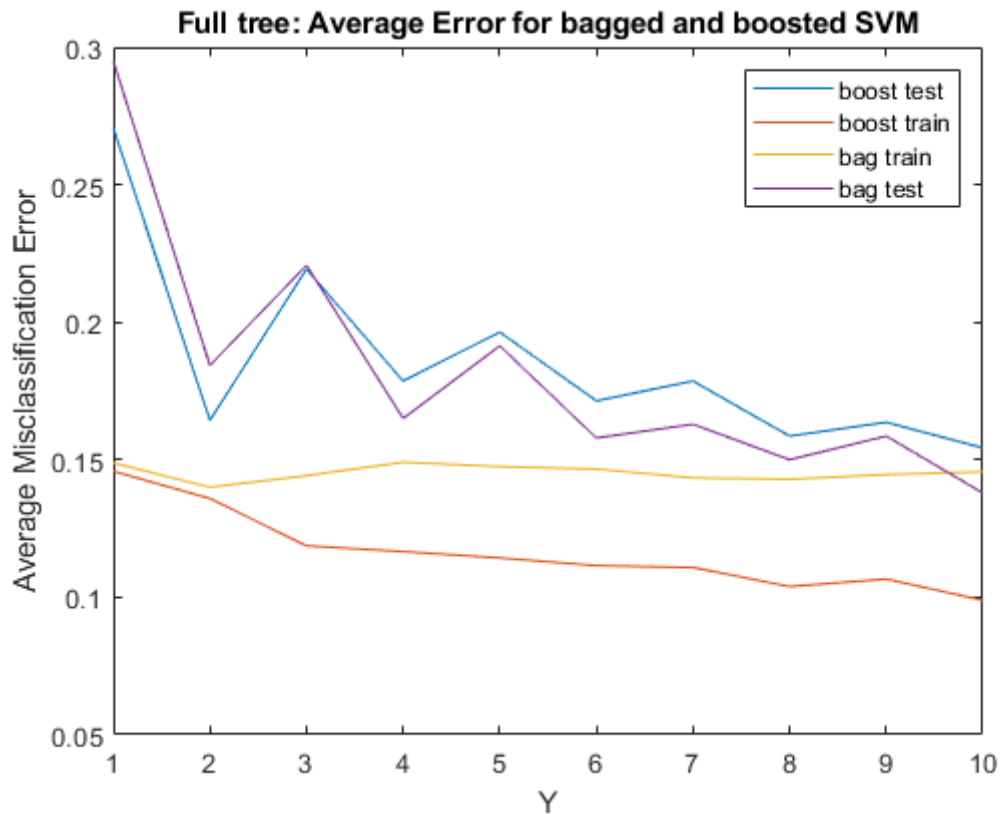
1 Bagging and Boosting SVMs

a.



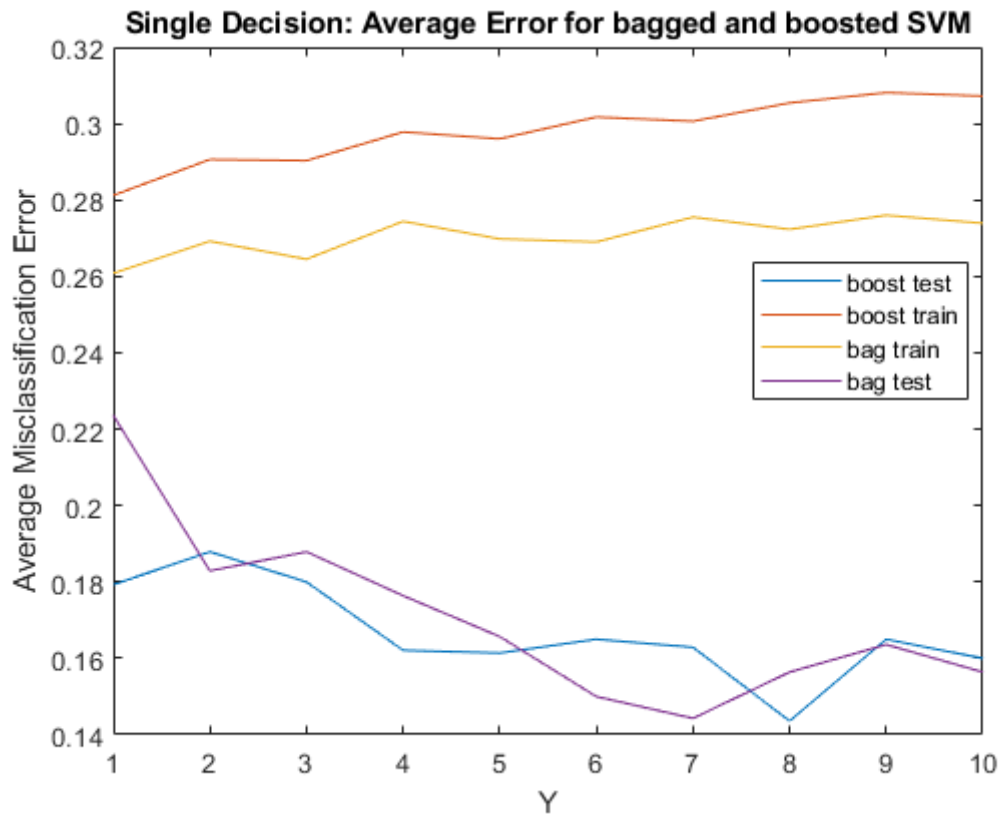
When increasing the T -value, both the boosted SVM and the bagged SVM models show small but relatively consistent improvement for the testing data. Both methods significantly improve over the $T=1$ models. For the training data, the bagged SVM struggles to make any significant improvement over the $T=1$ model, while the boosted model seems to make a 2% improvement before leveling off.

b.



With $T=1$, the full tree model results in significantly higher error than the model from part a. Other than having a slightly inflated error overall, the full tree results graph seems to have the same shape as the graph from part a. With this in mind (along with more data points) it's hard to ignore the trend that the test error improves significantly on even values of Y and worsens on odd values of Y . I'm not entirely certain why this is, but I can no longer chalk it up to testing coincidence. Compared to the results from part a, the boosted training error seems to improve more steadily now, while the bag training error stagnates to much that the bag testing error surpasses it at $Y=10$. In general, the training and testing curves are closer together for this graph compared to part a.

c.



The single decision models are severely different from any model from part a or b. First of all, for every T-value I found the testing error to be lower than the training error, which is quite surprising. Also, as for both the bagged and boosted models the training error increases slowly as T increases. Meanwhile, the bagged and boosted testing errors both improve as T increases. I'm not sure how a single decision can find properly classify over 80% of inputs with 65 variables, so I would say the testing data errors exceeded my expectations by far.