1. Summarize the random forest model results
   1. The optimal parameter mtry is 6 with an accuracy of 0.9568 and a kappa of 0.8257.

Table

Description automatically generated with medium confidence

* 1. Overall, the variable importance ranking does make sense to me. I think the number of rooms in a house has a significant impact on the value of a home, especially in the northeast where space is limited. Also, the % lower status of the population is important because people of lower socioeconomic status might not be able to afford properties above the median value and may influence the outcome of the model. However, I am not sure how relevant the pupil-teach ratio by town might be to initial question over the full-value-property-tax rate per $10,000 for example.

A picture containing text, receipt

Description automatically generated

Table

Description automatically generated

1. In the default setting, the linear SVM outperformed the radial SVM with an accuracy of .9735. This suggests that in the default setting, the linear SVM is better at classifying whether a property exceeds the median value of owner-occupied home in Boston. Furthermore, the linear SVM had a higher sensitivity rate at .9200, indicating that the linear SVM had more true positives than the radial SVM. Ultimately, this suggests that the linear SVM can predict when a property exceeds the median value better. However, the radial SVM had a higher specificity rate, suggesting that the radial was more accurate in terms of predicting when a home did not exceed the median value. Overall, in the default setting, the linear SVM had better model performance. **A screenshot of a computer

   Description automatically generated with low confidenceTable

   Description automatically generated**

Linear SVM Radial SVM

1. After tuning the linear SVM, the optimal parameter was c = 0.75. When c = 0.75, the accuracy decreased to .9351 when compared to the default setting. Furthermore, kappa also decreased to .7358, indicating that the classification accuracy lowered as well. Overall, the parameters that I chose were not better than the default parameters.

**Text

Description automatically generatedText

Description automatically generated**

1. In all, I believe that the Random Forest model has better performance over the SVM. The Random Forest model has a higher accuracy than both the SVMs in the default setting, as well as the Linear SVM after tuning. While the Random Forest model has the same sensitivity rate as the Linear SVM, it has a higher specificity rate even after the tune. Furthermore, the Random Forest model outperforms the Radial SVM in sensitivity but matches the Radial SVM in specificity. Overall, the Random Forest has shown that it is more accurate, and that it is better at predicting true positives while besting or matching in terms of true negatives.