

Mike Dowd  
lab05

Linear SVM classifier with 7/13 features ("Malic acid" "Total phenols" "Flavanoids"  
"Nonflavanoid Phenols" "Hue" "Od280/od315 of diluted wines" "Proline") on test data:

	Predicted			
Actual	1	2	3	
1	17	1	0	
2	3	20	0	
3	0	0	13	

	precision	recall	f1
1	0.850000	0.944444	0.894736
2	0.952381	0.869565	0.909090
3	1.000000	1.000000	1.000000

^^ This is the performance on the test data, the training data was classified perfectly

Polynomial SVM classifier with 7/13 features ("Malic acid" "Total phenols" "Flavanoids"  
"Nonflavanoid Phenols" "Hue" "Od280/od315 of diluted wines" "Proline") on test data:

	Predicted			
Actual	1	2	3	
1	15	3	0	
2	3	19	1	
3	0	0	13	

	precision	recall	f1
1	0.833333	0.833333	0.833333
2	0.863636	0.826087	0.844444
3	0.928571	1.000000	0.962963

It seems the polynomial model performs worse than the linear. It was also slightly worse on the training data

Now I used a KNN model with k=10. This is the performance over all data points:

	Predicted			
Actual	1	2	3	
1	55	1	3	
2	3	53	15	
3	2	15	31	

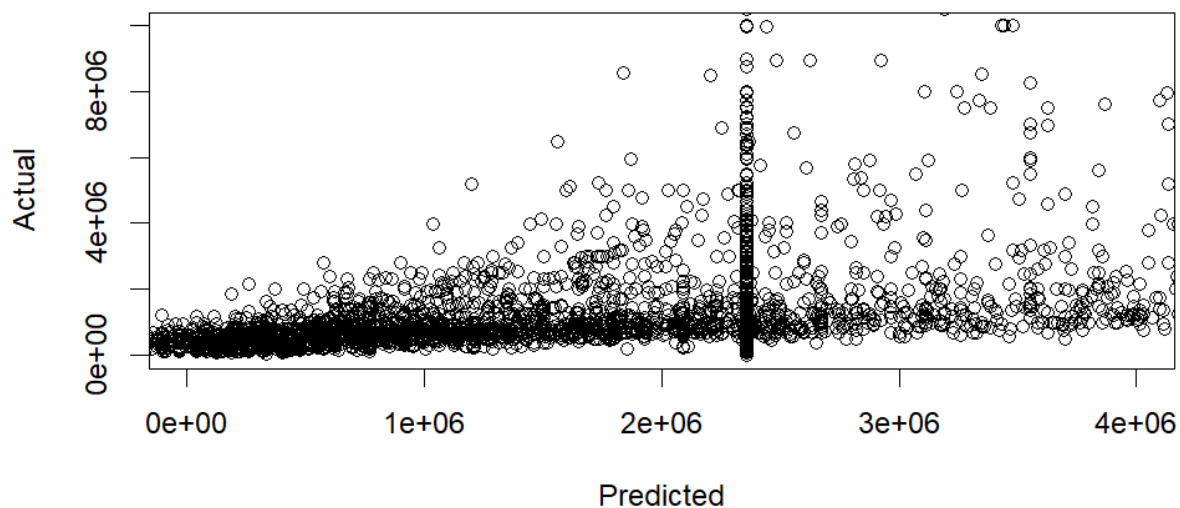
	precision	recall	f1
--	-----------	--------	----

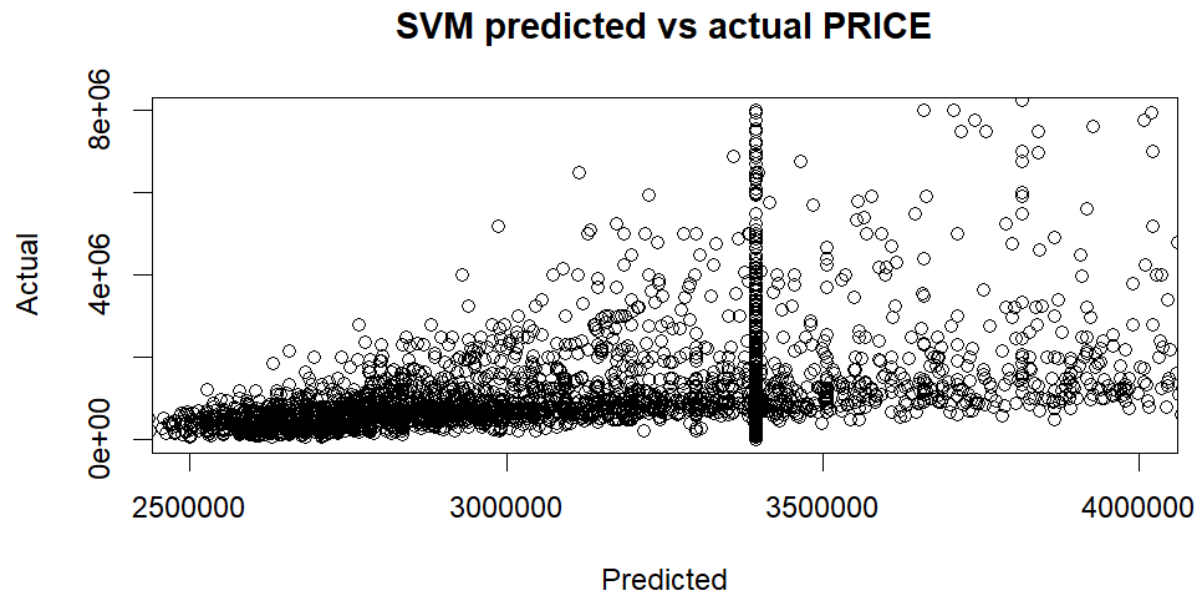
```
1 0.9166667 0.9322034 0.9243697
2 0.7681159 0.7464789 0.7571429
3 0.6326531 0.6458333 0.6391753
```

The KNN model performs the worst overall. However interestingly it performs better than both SVM models at predicting the 1st class. The linear SVM model performed the best over the three models on both the training and test data, with the highest overall accuracy and better f1 scores for all classes than the polynomial SVM model.

## SVM Regression

**Linear model predicted vs actual PRICE**





The main difference I see here is that the SVM has much less lower values predicted, as you can see from the graph I had to start the SVM plot at 2.5e+06, as there were no predicted values under it. However for the linear model, there are plenty of values predicted under that, which is not likely for the price of a house. Other than that, there aren't too many extreme differences, I would say the SVM seems to be closer to a 45/45 line which would be perfect predictions.