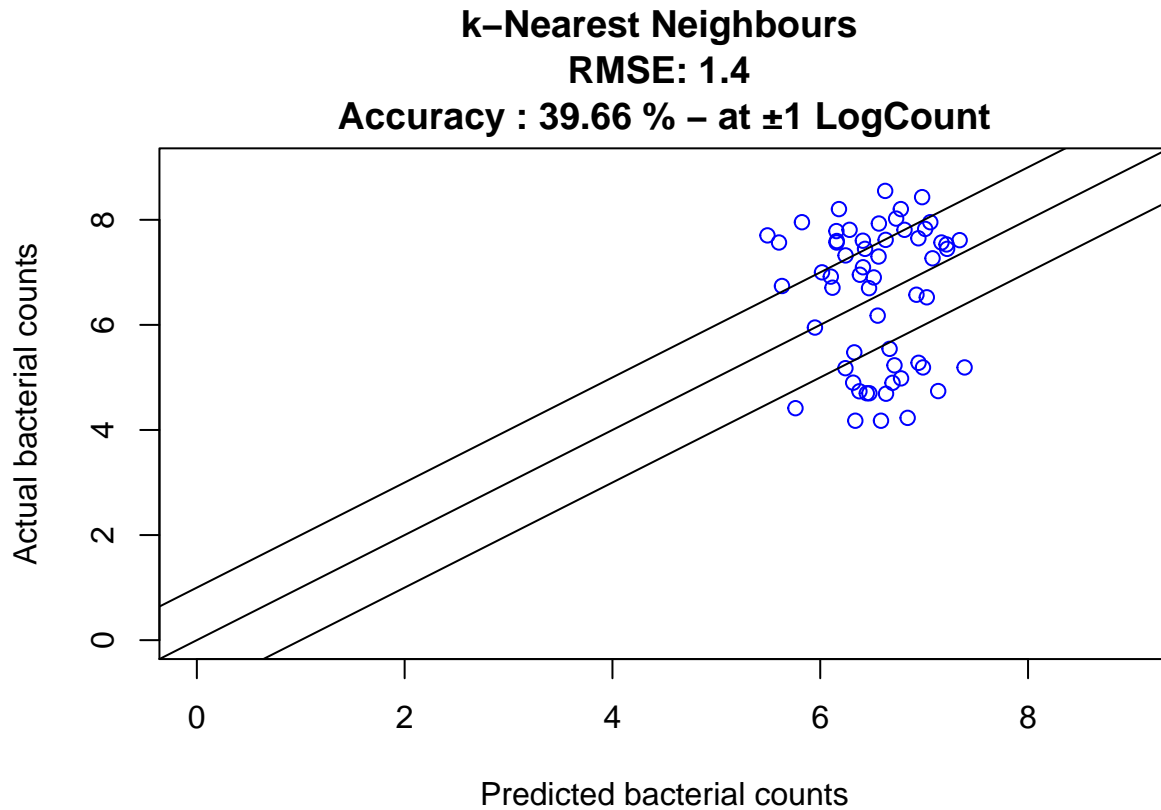


# Machine Learning Regression Model Report - Multispectral imaging

## k-Nearest Neighbours for Total Viable Counts

```
## k-Nearest Neighbors
##
## 138 samples
## 18 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 138, 138, 138, 138, 138, 138, ...
## Resampling results across tuning parameters:
##
##  k    RMSE      Rsquared    MAE
##  1  1.628877  0.07631172  1.243537
##  2  1.540528  0.06889876  1.196526
##  3  1.458107  0.07525481  1.160013
##  4  1.410955  0.07756274  1.137011
##  5  1.383676  0.07676612  1.113246
##  6  1.348109  0.08423229  1.093800
##  7  1.324584  0.09133435  1.073887
##  8  1.304858  0.10084289  1.057953
##  9  1.292287  0.10456589  1.051479
## 10  1.282264  0.10860362  1.043588
## 11  1.274284  0.11352365  1.036029
## 12  1.264728  0.11964079  1.028379
## 13  1.259579  0.12199076  1.026719
## 14  1.252244  0.12784045  1.018962
## 15  1.243733  0.13536835  1.011787
## 16  1.244774  0.13470195  1.012879
## 17  1.248375  0.12579566  1.018626
## 18  1.251891  0.12048311  1.022001
## 19  1.249402  0.12453075  1.023488
## 20  1.255870  0.11879677  1.029441
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was k = 15.
```

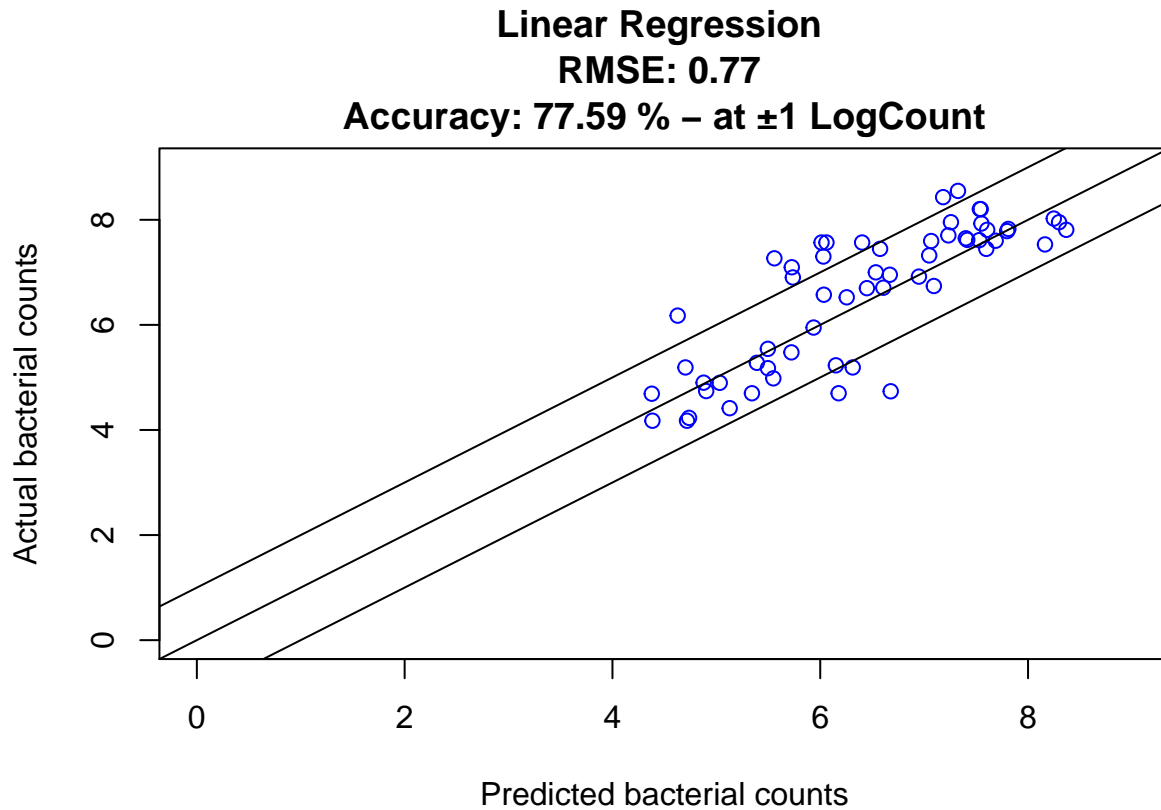
## Total Viable Counts distribution - k-Nearest Neighbours



## Linear Model for Total Viable Counts

```
## Linear Regression
##
## 138 samples
## 18 predictor
##
## No pre-processing
## Resampling: Cross-Validated (10 fold, repeated 3 times)
## Summary of sample sizes: 124, 125, 123, 126, 124, 125, ...
## Resampling results:
##
##   RMSE      Rsquared   MAE
## 0.7236853 0.7030126 0.6040175
##
## Tuning parameter 'intercept' was held constant at a value of TRUE
```

## Total Viable Counts distribution - Linear Regression



## Random Forest for Total Viable Counts

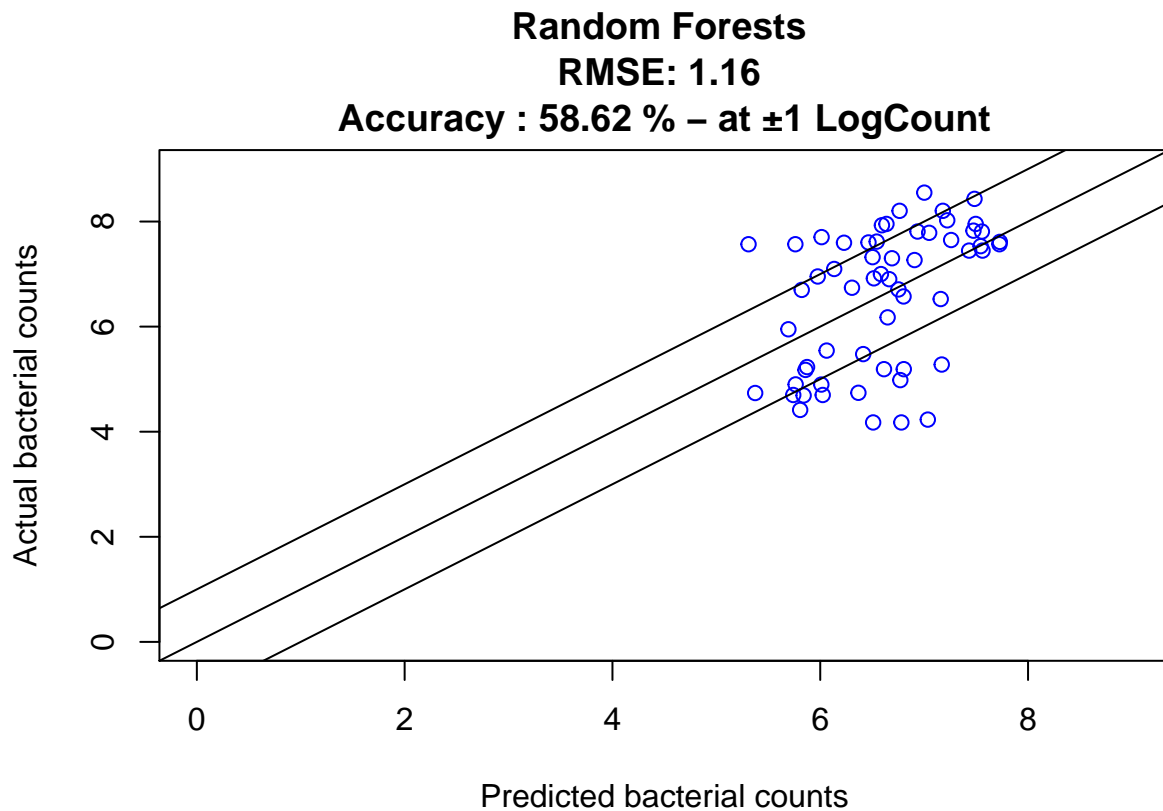
```
## Random Forest
##
## 138 samples
## 18 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 138, 138, 138, 138, 138, 138, ...
## Resampling results across tuning parameters:
##
## mtry  RMSE      Rsquared  MAE
## 1     1.188901  0.1841859  0.9630549
## 2     1.170060  0.2063113  0.9452098
## 3     1.161264  0.2201670  0.9377782
## 4     1.156754  0.2253392  0.9324010
## 5     1.150935  0.2312156  0.9266913
## 6     1.146597  0.2376249  0.9212315
## 7     1.144505  0.2405144  0.9212290
## 8     1.142046  0.2435282  0.9193930
## 9     1.141230  0.2441195  0.9176895
```

```

## 10 1.138232 0.2484838 0.9147746
## 11 1.137488 0.2497101 0.9131078
## 12 1.135517 0.2525015 0.9112199
## 13 1.132027 0.2559235 0.9079039
## 14 1.134073 0.2549491 0.9095807
## 15 1.137588 0.2508757 0.9099973
## 16 1.136646 0.2536130 0.9100582
## 17 1.136792 0.2547193 0.9095905
## 18 1.134369 0.2569851 0.9096968
## 19 1.135515 0.2557435 0.9085800
## 20 1.137550 0.2549353 0.9105189
## 21 1.137882 0.2541819 0.9112513
## 22 1.138924 0.2523766 0.9101218
## 23 1.136143 0.2557962 0.9099453
## 24 1.134835 0.2561576 0.9079689
## 25 1.137213 0.2524183 0.9108289
## 26 1.135381 0.2556186 0.9091418
## 27 1.134264 0.2554668 0.9068803
## 28 1.138132 0.2531244 0.9111146
## 29 1.135559 0.2541419 0.9080745
## 30 1.136549 0.2525486 0.9099291
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was mtry = 13.

```

## Total Viable Counts distribution - Random Forests



## Support Vector Machines with Polynomial Kernel for Total Viable Counts

```
## Support Vector Machines with Polynomial Kernel
##
## 138 samples
## 18 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 138, 138, 138, 138, 138, 138, ...
## Resampling results across tuning parameters:
##
## degree scale C      RMSE      Rsquared    MAE
## 1      1e-03 0.25  1.3676301 0.09253978 1.0526364
## 1      1e-03 0.50  1.3576378 0.11436081 1.0418330
## 1      1e-03 1.00  1.3447874 0.14142642 1.0258226
## 1      1e-03 2.00  1.3229561 0.17469934 0.9991767
## 1      1e-03 4.00  1.2892970 0.21807947 0.9676240
## 1      1e-02 0.25  1.3146656 0.18896765 0.9898902
## 1      1e-02 0.50  1.2739168 0.22930067 0.9572206
## 1      1e-02 1.00  1.2242615 0.28717548 0.9208566
```

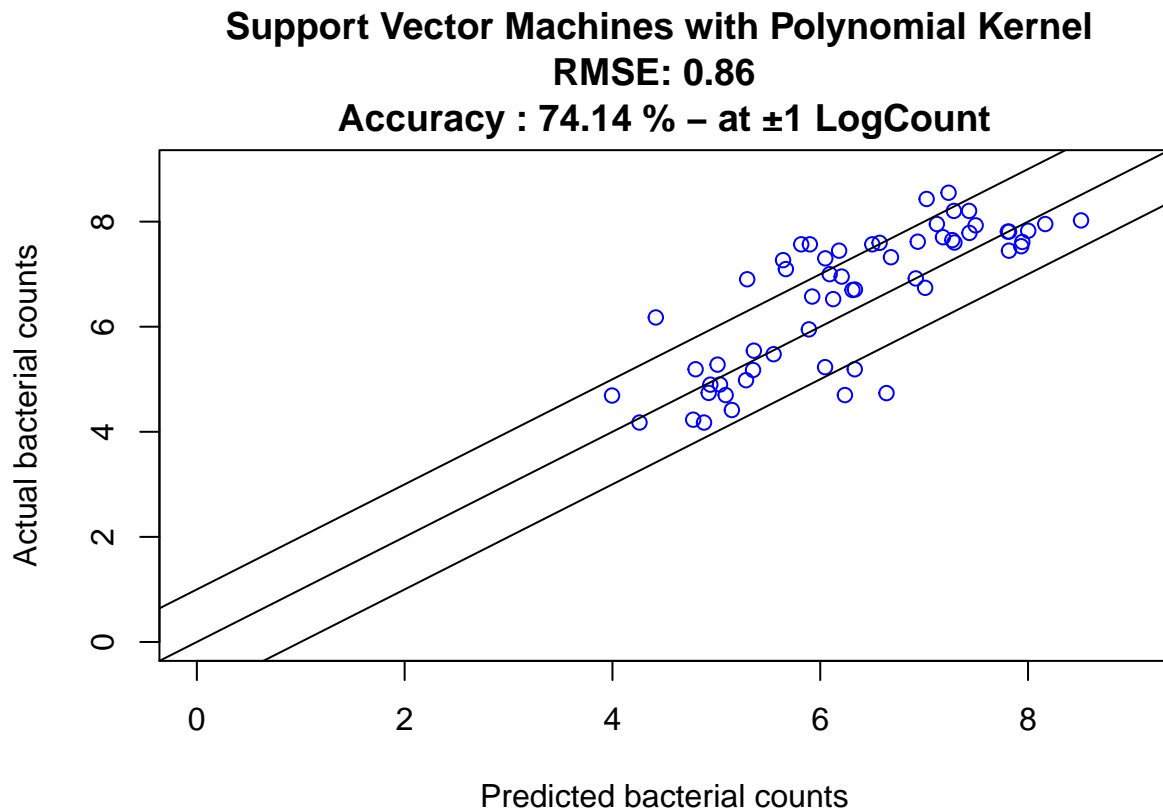
##	1	1e-02	2.00	1.1509346	0.36491743	0.8777510
##	1	1e-02	4.00	1.0547758	0.46194316	0.8189233
##	1	1e-01	0.25	1.1242524	0.39484159	0.8619770
##	1	1e-01	0.50	1.0225750	0.48711874	0.7981973
##	1	1e-01	1.00	0.9245739	0.54757236	0.7297399
##	1	1e-01	2.00	0.8628217	0.58732377	0.6867345
##	1	1e-01	4.00	0.8345130	0.61291308	0.6659258
##	1	1e+00	0.25	0.8492290	0.59791700	0.6763812
##	1	1e+00	0.50	0.8260830	0.62137573	0.6618675
##	1	1e+00	1.00	0.8035463	0.64297265	0.6493684
##	1	1e+00	2.00	0.7890749	0.65805040	0.6400528
##	1	1e+00	4.00	0.7727995	0.67239067	0.6274016
##	1	1e+01	0.25	0.7843358	0.66303873	0.6366435
##	1	1e+01	0.50	0.7692698	0.67543677	0.6246898
##	1	1e+01	1.00	0.7661456	0.67748857	0.6212147
##	1	1e+01	2.00	0.7721518	0.67654938	0.6250640
##	1	1e+01	4.00	0.7887900	0.66630463	0.6358592
##	2	1e-03	0.25	1.3574861	0.11445617	1.0418212
##	2	1e-03	0.50	1.3449095	0.14231106	1.0259957
##	2	1e-03	1.00	1.3227719	0.17647156	0.9992471
##	2	1e-03	2.00	1.2890448	0.21929010	0.9671754
##	2	1e-03	4.00	1.2410911	0.26913136	0.9327727
##	2	1e-02	0.25	1.2774312	0.22041267	0.9611724
##	2	1e-02	0.50	1.2295493	0.27600059	0.9290269
##	2	1e-02	1.00	1.1580789	0.34645258	0.8858744
##	2	1e-02	2.00	1.0614462	0.42988072	0.8209364
##	2	1e-02	4.00	0.9721790	0.49952452	0.7548686
##	2	1e-01	0.25	1.0585318	0.39047017	0.8189041
##	2	1e-01	0.50	1.0001907	0.44391312	0.7741359
##	2	1e-01	1.00	0.9760972	0.47291690	0.7601455
##	2	1e-01	2.00	0.9536689	0.50449127	0.7537555
##	2	1e-01	4.00	0.9454659	0.52377752	0.7519660
##	2	1e+00	0.25	1.0529717	0.45222153	0.8212108
##	2	1e+00	0.50	1.0799058	0.45118343	0.8419655
##	2	1e+00	1.00	1.1071367	0.45045362	0.8666394
##	2	1e+00	2.00	1.1891505	0.42618057	0.9281434
##	2	1e+00	4.00	1.2741221	0.40875016	0.9870502
##	2	1e+01	0.25	1.6977033	0.27431639	1.2640238
##	2	1e+01	0.50	1.8474663	0.25114590	1.3626904
##	2	1e+01	1.00	1.9812785	0.23054046	1.4471620
##	2	1e+01	2.00	2.0357445	0.22125060	1.4813216
##	2	1e+01	4.00	2.0610812	0.21794628	1.4975662
##	3	1e-03	0.25	1.3505266	0.13076728	1.0333781
##	3	1e-03	0.50	1.3315863	0.16101908	1.0111778
##	3	1e-03	1.00	1.3051090	0.20180973	0.9810911
##	3	1e-03	2.00	1.2615842	0.24169008	0.9490091
##	3	1e-03	4.00	1.2083516	0.30753939	0.9104506
##	3	1e-02	0.25	1.2531271	0.23573889	0.9477926
##	3	1e-02	0.50	1.1939533	0.29818407	0.9107428
##	3	1e-02	1.00	1.1114518	0.37717097	0.8559972
##	3	1e-02	2.00	1.0177586	0.45288171	0.7865700
##	3	1e-02	4.00	0.9530155	0.50118378	0.7440590
##	3	1e-01	0.25	1.0626304	0.41051738	0.8280139
##	3	1e-01	0.50	1.0852208	0.42096477	0.8440464

```

## 3      1e-01  1.00  1.1492715  0.41480486  0.8847046
## 3      1e-01  2.00  1.2214640  0.40746119  0.9234563
## 3      1e-01  4.00  1.3161489  0.38200657  0.9879433
## 3      1e+00  0.25  2.4404176  0.19424901  1.6893924
## 3      1e+00  0.50  2.5459452  0.18415106  1.7565941
## 3      1e+00  1.00  2.5567404  0.18349190  1.7616413
## 3      1e+00  2.00  2.5567404  0.18349190  1.7616413
## 3      1e+00  4.00  2.5567404  0.18349190  1.7616413
## 3      1e+01  0.25  3.7945211  0.12309847  2.4624198
## 3      1e+01  0.50  3.7945211  0.12309847  2.4624198
## 3      1e+01  1.00  3.7945211  0.12309847  2.4624198
## 3      1e+01  2.00  3.7945211  0.12309847  2.4624198
## 3      1e+01  4.00  3.7945211  0.12309847  2.4624198
##
## RMSE was used to select the optimal model using the smallest value.
## The final values used for the model were degree = 1, scale = 10 and C = 1.

```

## Total Viable Counts distribution - Support Vector Machines with Polynomial Kernel

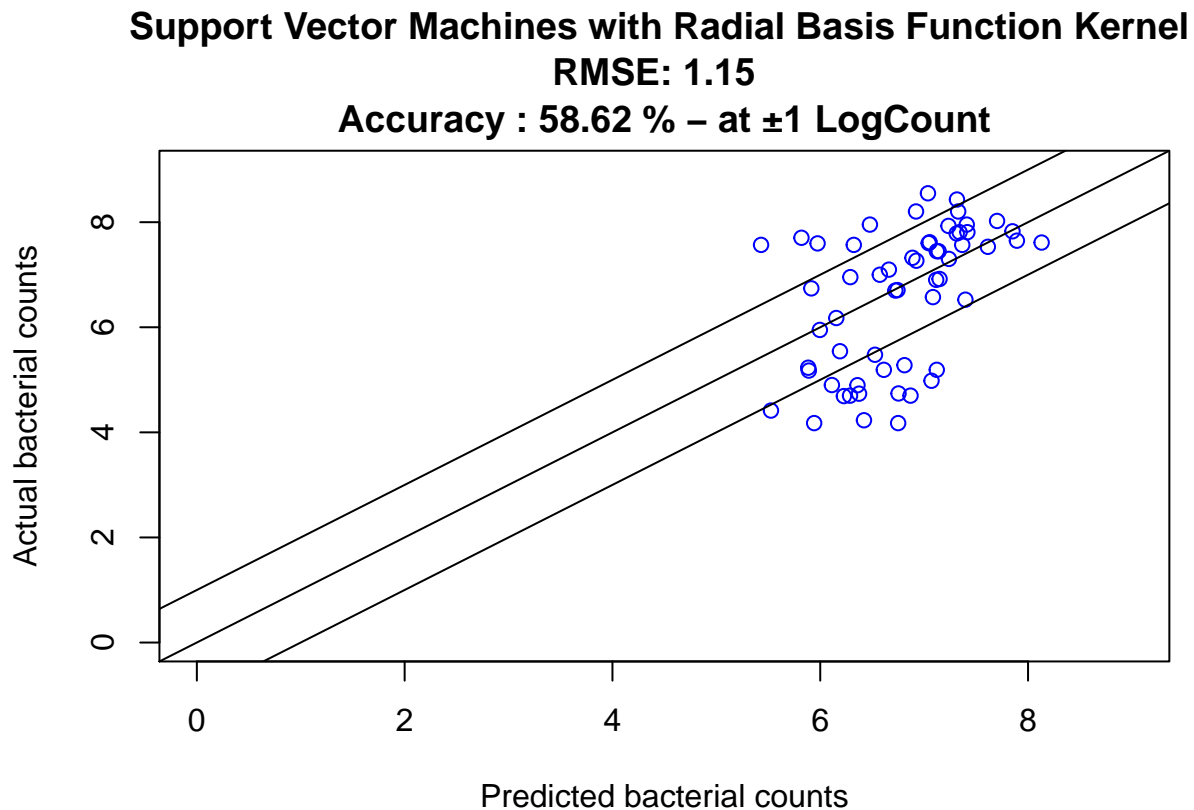


## Support Vector Machines with Radial Basis Function Kernel for Total Viable Counts

```
## Support Vector Machines with Radial Basis Function Kernel
##
## 138 samples
## 18 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 138, 138, 138, 138, 138, 138, ...
## Resampling results across tuning parameters:
##
##      C      RMSE      Rsquared    MAE
## 0.25  1.209806  0.2458085  0.9321366
## 0.50  1.166212  0.2674341  0.9041694
## 1.00  1.132044  0.2907852  0.8819976
##
## Tuning parameter 'sigma' was held constant at a value of 0.08788172
## RMSE was used to select the optimal model using the smallest value.
## The final values used for the model were sigma = 0.08788172 and C = 1.
```



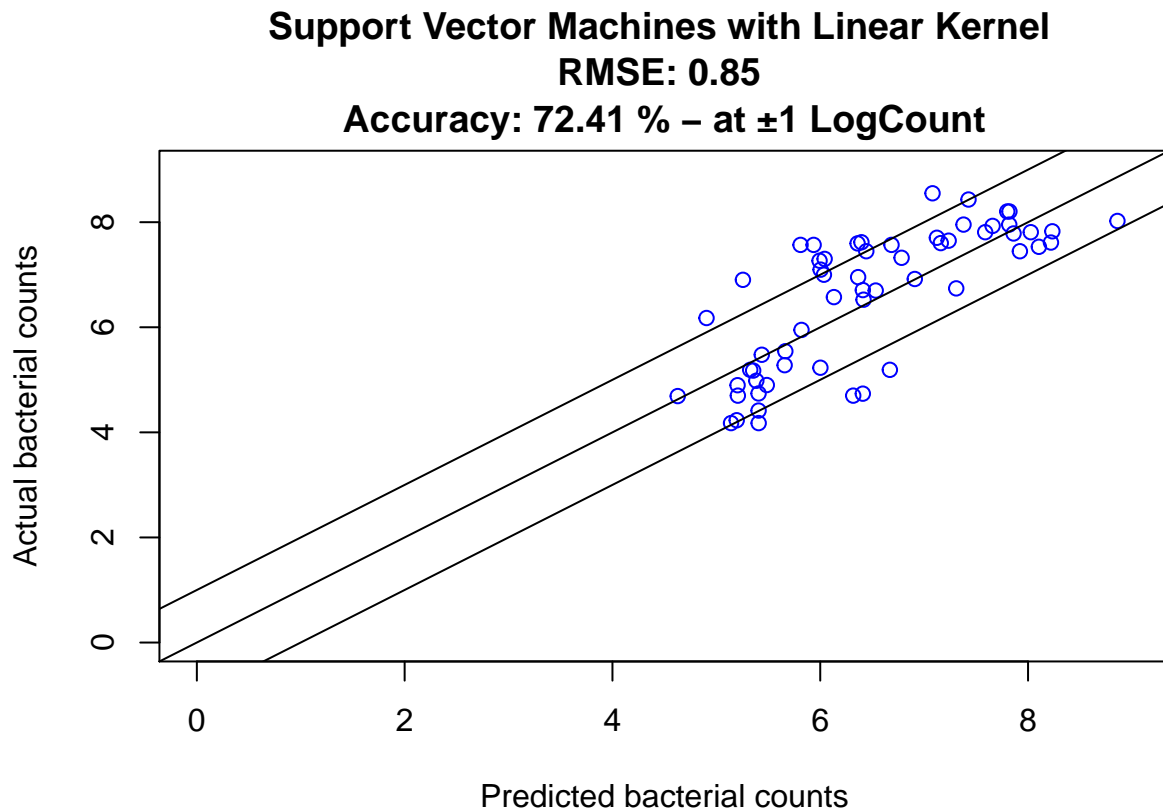
## Total Viable Counts distribution - Support Vector Machines with Radial Basis Function Kernel



## Support Vector Machines with Linear Kernel for Total Viable Counts

```
## Support Vector Machines with Linear Kernel
##
## 138 samples
## 18 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 138, 138, 138, 138, 138, 138, ...
## Resampling results:
##
##      RMSE      Rsquared   MAE
## 0.8387273 0.6353293 0.6790855
##
## Tuning parameter 'C' was held constant at a value of 1
```

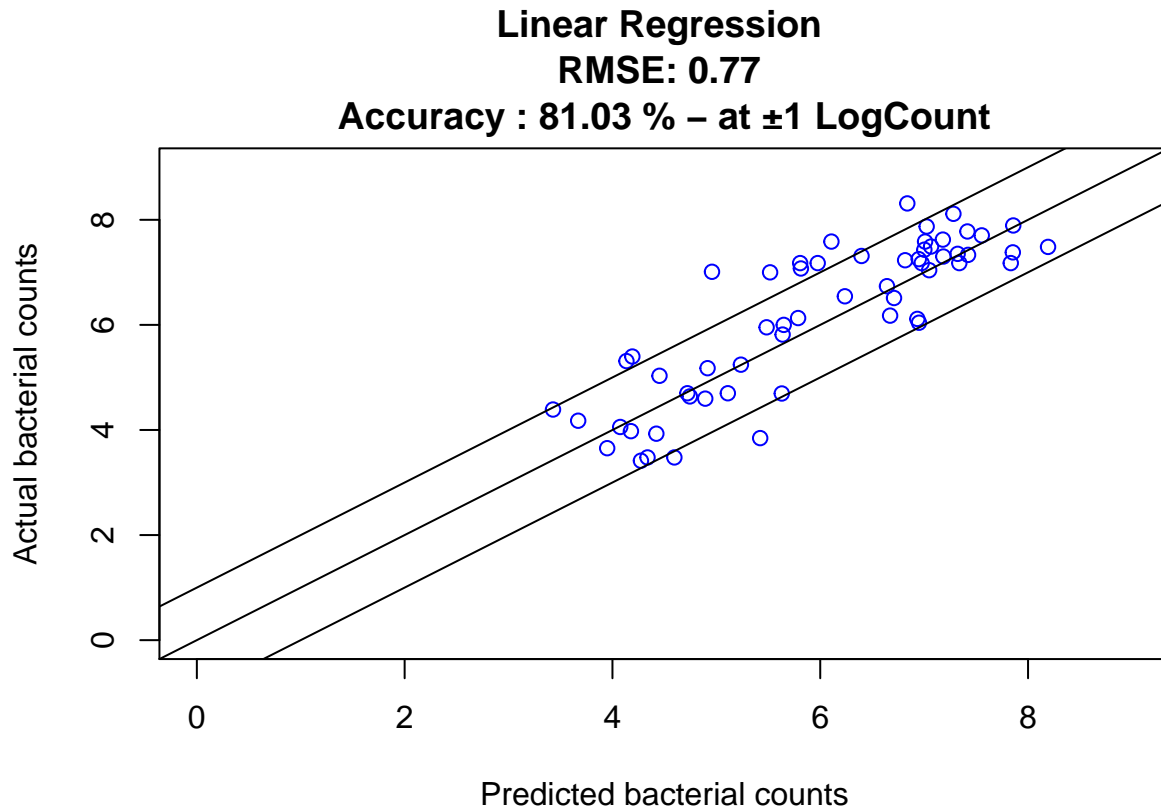
## Total Viable Counts distribution - Support Vector Machines with Linear Kernel



## Linear Regression for Pseudomonas count

```
## Linear Regression
##
## 138 samples
## 18 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 138, 138, 138, 138, 138, 138, ...
## Resampling results:
##
##      RMSE      Rsquared  MAE
## 0.9018537 0.623495 0.7242078
##
## Tuning parameter 'intercept' was held constant at a value of TRUE
```

## Pseudomonas count distribution - Linear Regression



## k-Nearest Neighbors for Pseudomonas count

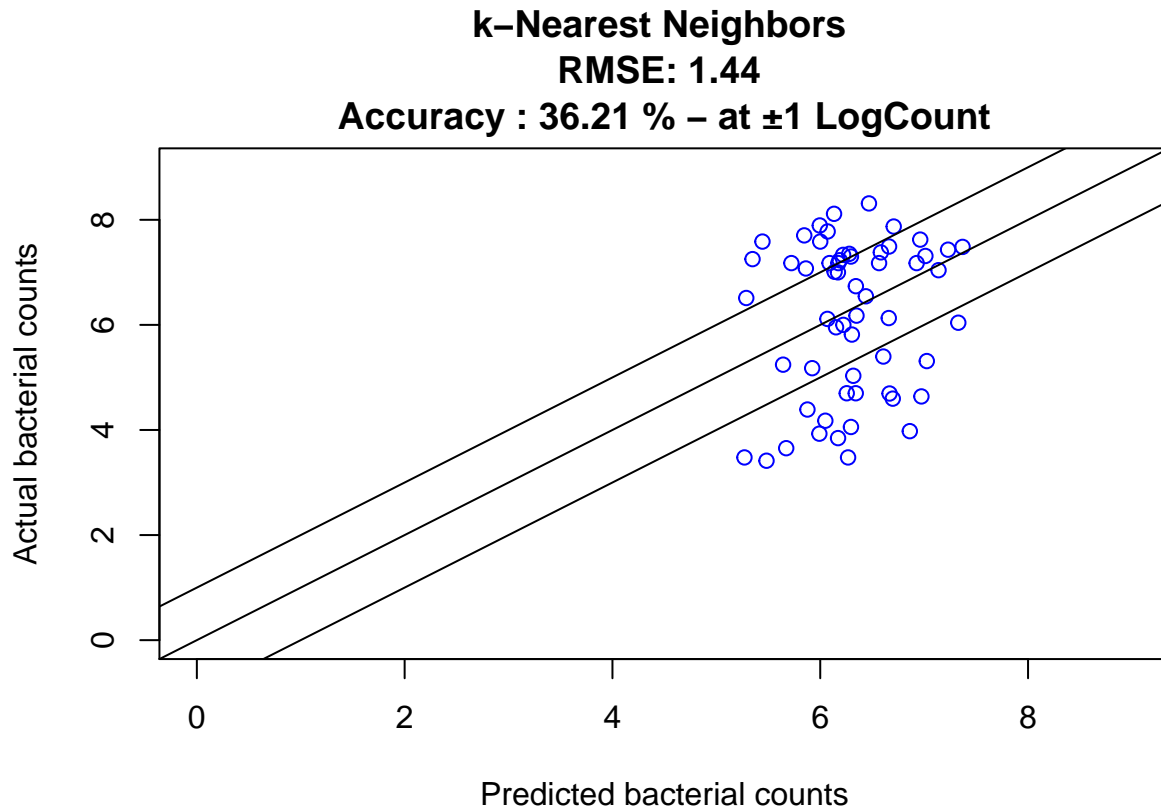
```
## k-Nearest Neighbors
##
## 138 samples
## 18 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 138, 138, 138, 138, 138, 138, ...
## Resampling results across tuning parameters:
##
## k    RMSE      Rsquared    MAE
## 1  1.802243  0.05546741  1.375047
## 2  1.655395  0.06903053  1.266341
## 3  1.579494  0.07674866  1.228527
## 4  1.500408  0.08947601  1.183179
## 5  1.464973  0.09318193  1.168453
## 6  1.423114  0.10435305  1.144412
## 7  1.409726  0.10514684  1.142439
## 8  1.405345  0.10195844  1.143662
## 9  1.395716  0.10335157  1.136140
```

```

## 10 1.391819 0.10070734 1.133982
## 11 1.381842 0.10361957 1.132392
## 12 1.365477 0.11582739 1.119393
## 13 1.373494 0.10354011 1.130199
## 14 1.367918 0.10429634 1.127995
## 15 1.362717 0.10671975 1.131448
## 16 1.367447 0.10265767 1.139331
## 17 1.373960 0.09653224 1.148381
## 18 1.372610 0.09719171 1.146859
## 19 1.380481 0.08622640 1.156282
## 20 1.382652 0.08463720 1.161897
## 21 1.384240 0.08042581 1.167150
## 22 1.387457 0.07556320 1.173572
## 23 1.387959 0.07494243 1.174797
## 24 1.394417 0.07120362 1.182651
## 25 1.397370 0.06753618 1.186496
## 26 1.396697 0.06691320 1.185932
## 27 1.397408 0.06442330 1.186580
## 28 1.400129 0.05861022 1.189993
## 29 1.402177 0.05556277 1.193019
## 30 1.399274 0.05781532 1.190742
## 31 1.399815 0.05561283 1.191632
## 32 1.401270 0.05206061 1.193260
## 33 1.401860 0.05207312 1.194276
## 34 1.403129 0.05176167 1.195678
## 35 1.403378 0.05168143 1.198561
## 36 1.402083 0.05405971 1.197215
## 37 1.401814 0.05302619 1.197706
## 38 1.403350 0.05001614 1.200067
## 39 1.402941 0.04962366 1.200396
## 40 1.404510 0.04726071 1.201995
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was k = 15.

```

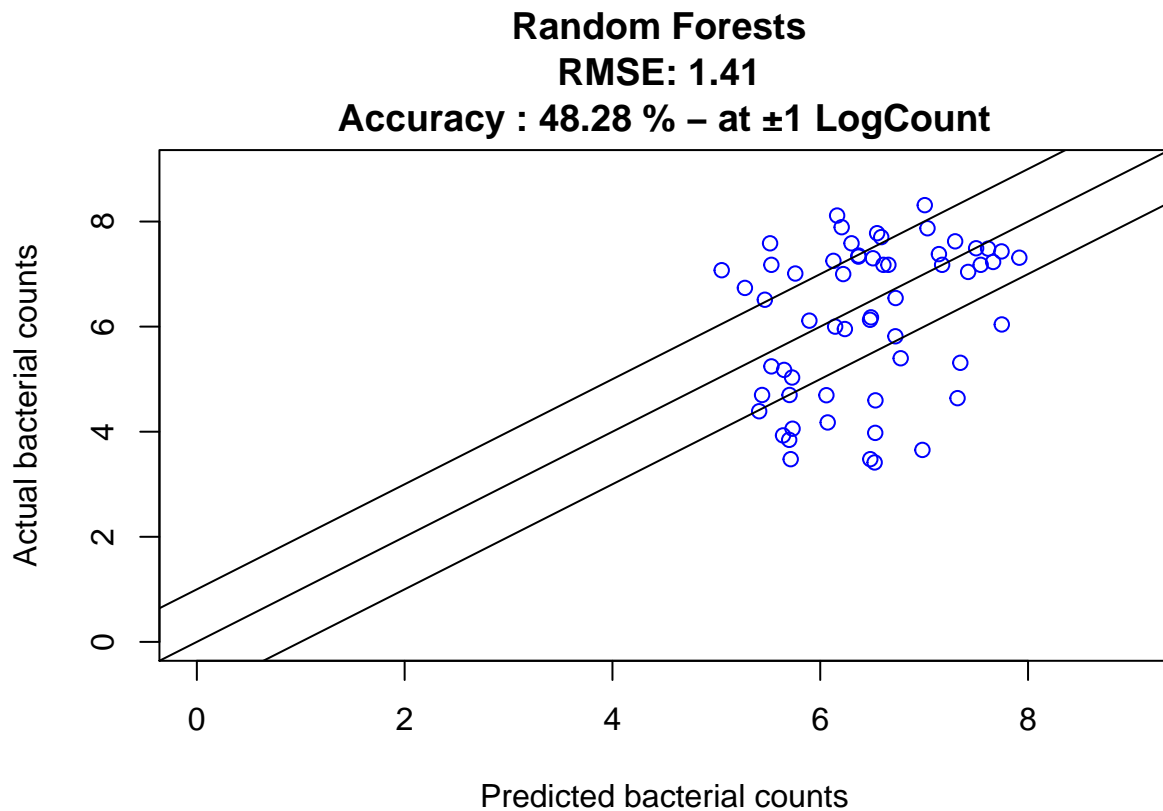
## Pseudomonas count distribution - k-Nearest Neighbors



## Random Forests for Pseudomonas count

```
## Random Forest
##
## 138 samples
## 18 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 138, 138, 138, 138, 138, 138, ...
## Resampling results across tuning parameters:
##
##  mtry  RMSE      Rsquared  MAE
##    2    1.278884  0.2073877  1.032475
##   10    1.273923  0.2258058  1.010466
##   18    1.281386  0.2272474  1.013910
##
## RMSE was used to select the optimal model using the smallest value.
## The final value used for the model was mtry = 10.
```

## Pseudomonas count distribution - Random Forests

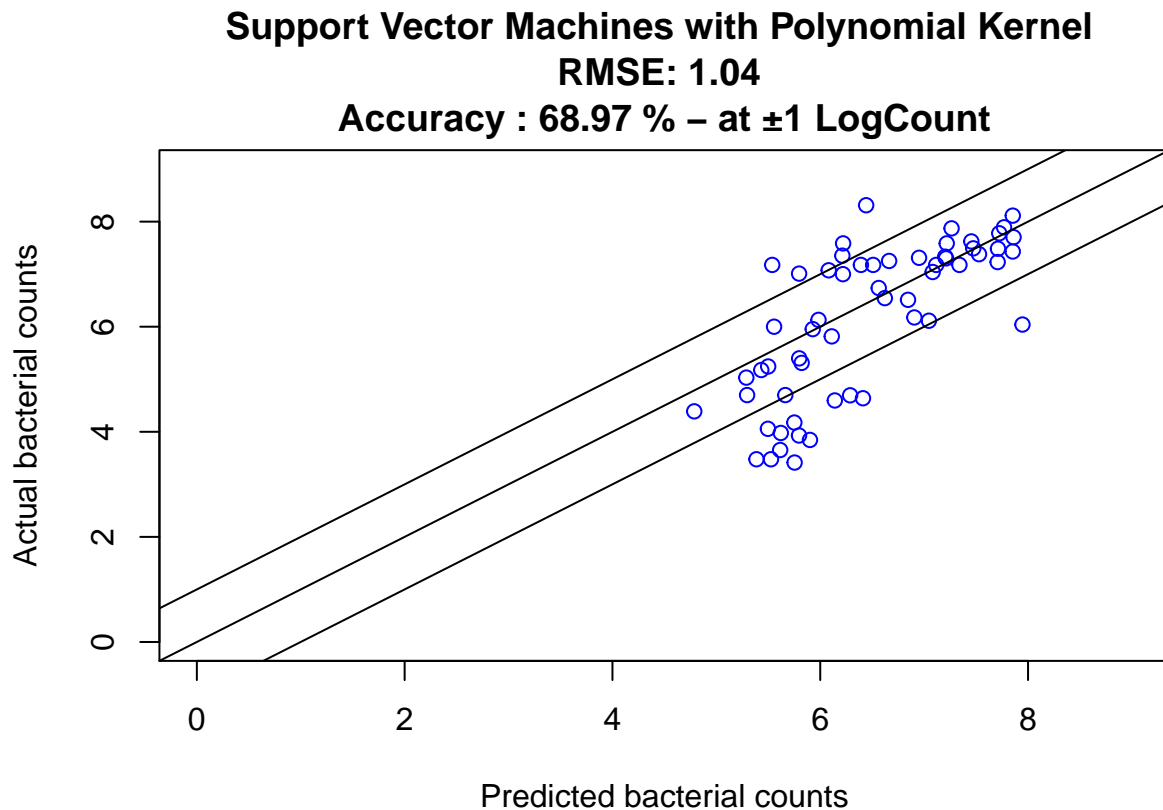


## Support Vector Machines with Polynomial Kernel for Pseudomonas count

```
## Support Vector Machines with Polynomial Kernel
##
## 138 samples
## 18 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 138, 138, 138, 138, 138, ...
## Resampling results across tuning parameters:
##
## degree scale C      RMSE      Rsquared    MAE
## 1      0.001 0.25  1.498815  0.08671280  1.2085891
## 1      0.001 0.50  1.490548  0.09157661  1.2004254
## 1      0.001 1.00  1.476151  0.11743751  1.1841817
## 1      0.010 0.25  1.441459  0.14467394  1.1412708
## 1      0.010 0.50  1.388721  0.18625010  1.0938515
## 1      0.010 1.00  1.322276  0.24445597  1.0388502
## 1      0.100 0.25  1.230259  0.33165599  0.9705950
## 1      0.100 0.50  1.155914  0.40945286  0.9159509
```

##	1	0.100	1.00	1.081890	0.47339853	0.8600469
##	2	0.001	0.25	1.490375	0.09201551	1.2002851
##	2	0.001	0.50	1.475767	0.11882243	1.1838592
##	2	0.001	1.00	1.452244	0.13749271	1.1542460
##	2	0.010	0.25	1.380201	0.20256241	1.0802025
##	2	0.010	0.50	1.318839	0.25102031	1.0245075
##	2	0.010	1.00	1.249904	0.31026211	0.9727546
##	2	0.100	0.25	1.183395	0.35613040	0.9375475
##	2	0.100	0.50	1.132691	0.40035254	0.9011227
##	2	0.100	1.00	1.109878	0.42418939	0.8799024
##	3	0.001	0.25	1.483127	0.10662421	1.1920621
##	3	0.001	0.50	1.465480	0.13270366	1.1694440
##	3	0.001	1.00	1.425583	0.16197226	1.1272731
##	3	0.010	0.25	1.346100	0.21848618	1.0481109
##	3	0.010	0.50	1.283780	0.27131575	0.9997789
##	3	0.010	1.00	1.214849	0.33417540	0.9544008
##	3	0.100	0.25	1.245176	0.32308927	0.9708177
##	3	0.100	0.50	1.261190	0.34443858	0.9784025
##	3	0.100	1.00	1.310914	0.33367670	1.0049243
##						
##	RMSE was used to select the optimal model using the smallest value.					
##	The final values used for the model were degree = 1, scale = 0.1 and C = 1.					

## Pseudomonas count distribution - Support Vector Machines with Polynomial Kernel

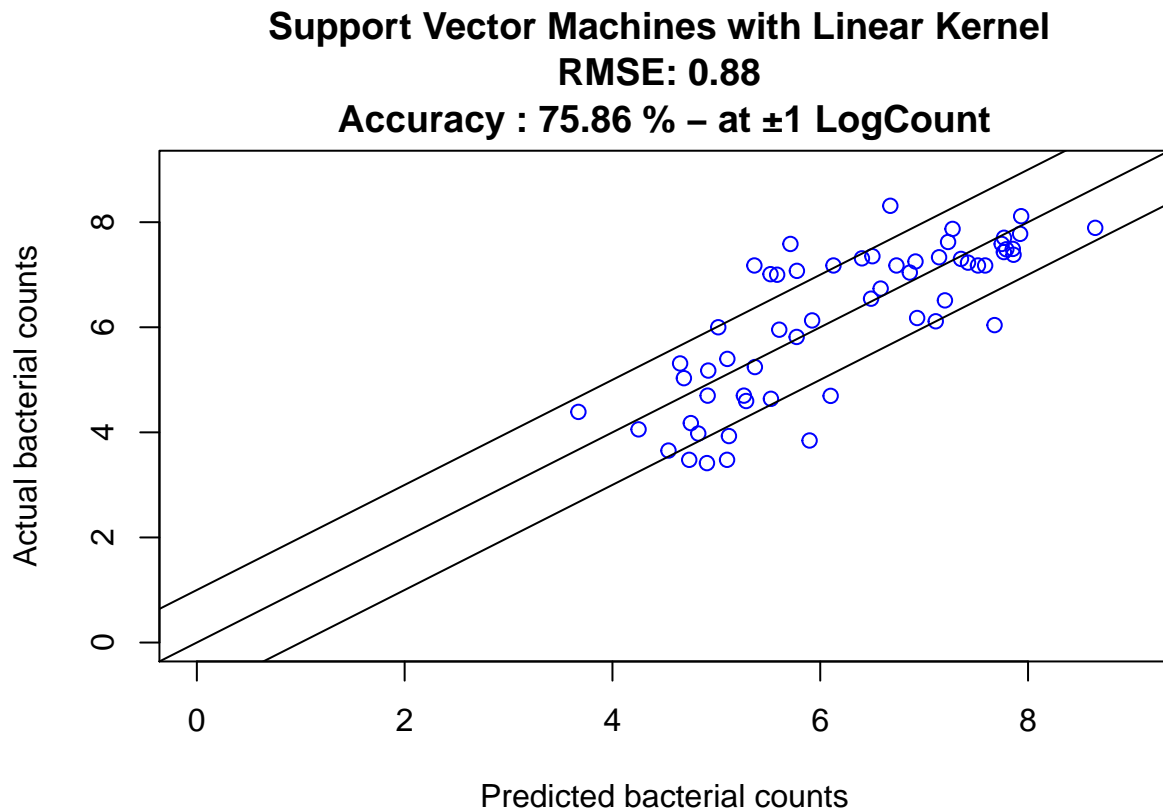


## Support Vector Machines with Linear Kernel for Pseudomonas count

```
## Support Vector Machines with Linear Kernel
##
## 138 samples
## 18 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 138, 138, 138, 138, 138, 138, ...
## Resampling results:
##
##   RMSE      Rsquared   MAE
## 0.9496852  0.5848639  0.7452106
##
## Tuning parameter 'C' was held constant at a value of 1
```



## Pseudomonas count distribution - Support Vector Machines with Linear Kernel



## Support Vector Machines with Radial Basis Function Kernel for Pseudomonas count

```
## Support Vector Machines with Radial Basis Function Kernel
##
## 138 samples
## 18 predictor
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 138, 138, 138, 138, 138, 138, ...
## Resampling results across tuning parameters:
##
## C      RMSE      Rsquared  MAE
## 0.25   1.370191   0.1973470  1.087674
## 0.50   1.323962   0.2256414  1.057805
## 1.00   1.284950   0.2587782  1.027601
##
## Tuning parameter 'sigma' was held constant at a value of 0.07967483
## RMSE was used to select the optimal model using the smallest value.
```

## The final values used for the model were  $\sigma = 0.07967483$  and  $C = 1$ .

## Pseudomonas count distribution - Support Vector Machines with Radial Basis Function Kernel

