**Chemical Predictors**

**(Same preprocessing has been done as the Ex 12 and biological predictors)**

**1. Mixture Discriminant Analysis**   
   
 225 samples  
 105 predictors  
 3 classes: 'Mild', 'None', 'Severe'   
   
 Pre-processing : Center and Scale  
 Resampling: Bootstrapped (25 reps)   
   
 Summary of sample sizes: 225, 225, 225, 225, 225, 225, ...   
   
 Resampling results across tuning parameters:  
   
 subclasses Accuracy Kappa Accuracy SD Kappa SD   
 1 0.4768575 0.1407148 0.04461386 0.06389823  
 2 0.4809074 0.1380243 0.04511030 0.06184408  
 3 0.4790951 0.1497179 0.04332225 0.06019336  
 4 0.4930027 0.1261193 0.02066666 0.06079511  
 5 0.4935065 0.1365727 NA NA  
   
 Kappa was used to select the optimal model using the largest value.  
 The final value used for the model was subclasses = 3.

Confusion Matrix and Statistics **Testing set**  
   
 Reference  
 Prediction Mild None Severe  
 Mild 19 12 6  
 None 7 4 0  
 Severe 3 5 0  
   
 Overall Statistics  
   
 Accuracy : 0.4107   
 95% CI : (0.281, 0.5502)  
 No Information Rate : 0.5179   
 P-Value [Acc > NIR] : 0.95910   
   
 Kappa : -0.0359   
 Mcnemar's Test P-Value : 0.06249

Statistics by Class:  
   
 Class: Mild Class: None Class: Severe  
 Sensitivity 0.6552 0.19048 0.0000  
 Specificity 0.3333 0.80000 0.8400

**2. Neural Network**   
   
 225 samples  
 105 predictors  
 3 classes: 'Mild', 'None', 'Severe'   
   
 Pre-processing: spatial sign transformation, scaled, centered   
 Resampling: Bootstrapped (25 reps)   
   
 Summary of sample sizes: 225, 225, 225, 225, 225, 225, ...   
   
 Resampling results across tuning parameters:  
   
 size decay Accuracy Kappa Accuracy SD Kappa SD   
 1 0.0 0.4362578 0.0643256215 0.09454080 0.086041032  
 1 0.1 0.5397015 0.1535189601 0.05579580 0.090870709  
 1 1.0 0.5319810 0.1108969983 0.05090369 0.075931592  
 1 2.0 0.5029356 0.0000000000 0.07226298 0.000000000  
 2 0.0 0.4656351 0.1232237996 0.07902498 0.083677834  
 2 0.1 0.5220552 0.1522592843 0.05154941 0.080289755  
 2 1.0 0.5276783 0.1028643224 0.05041837 0.088945147  
 2 2.0 0.5029356 0.0000000000 0.07226298 0.000000000  
 3 0.0 0.4837104 0.1179377366 0.05773742 0.083631879  
 3 0.1 0.5289381 0.1632786716 0.05074972 0.086573708  
 3 1.0 0.5262686 0.1016492490 0.05092069 0.088742193  
 3 2.0 0.5038245 -0.0011502054 0.06931202 0.005751027  
 4 0.0 0.4652746 0.0940609344 0.06512006 0.089810182  
 4 0.1 0.5360214 0.1756546192 0.05396036 0.088262583  
 4 1.0 0.5276264 0.1036408713 0.05074551 0.088828979  
 4 2.0 0.5038245 -0.0011502054 0.06931202 0.005751027  
 5 0.0 0.4855815 0.1182206895 0.04985448 0.080637664  
 6 0.1 0.5326444 0.1667321278 0.04741452 0.079481441  
 6 1.0 0.5276311 0.1037629926 0.05048460 0.087904638  
 6 2.0 0.5042690 -0.0007327586 0.06789757 0.003663793  
 7 0.0 0.4965067 0.1325307242 0.05740434 0.087491481  
   
 Kappa was used to select the optimal model using the largest value.  
 The final values used for the model were size = 4 and decay = 0.1.

Confusion Matrix and Statistics **Testing set**  
   
 Reference  
 Prediction Mild None Severe  
 Mild 20 7 6  
 None 5 10 0  
 Severe 4 4 0

Overall Statistics for Testing set  
   
 Accuracy : 0.5357   
 95% CI : (0.3974, 0.6701)  
 No Information Rate : 0.5179   
 P-Value [Acc > NIR] : 0.4475   
   
 Kappa : 0.1982   
 Mcnemar's Test P-Value : 0.1924   
   
 Statistics by Class:  
   
 Class: Mild Class: None Class: Severe  
 Sensitivity 0.6897 0.4762 0.0000  
 Specificity 0.5185 0.8571 0.8400  
 Pos Pred Value 0.6061 0.6667 0.0000  
 Neg Pred Value 0.6087 0.7317 0.8750  
 Prevalence 0.5179 0.3750 0.1071  
 Detection Rate 0.3571 0.1786 0.0000  
 Detection Prevalence 0.5893 0.2679 0.1429  
 Balanced Accuracy 0.6041 0.6667 0.4200

**3. Flexible Discriminant Analysis**   
   
 225 samples  
 105 predictors  
 3 classes: 'Mild', 'None', 'Severe'   
   
 Pre-processing : Center and Scale  
 Resampling: Bootstrapped (25 reps)   
   
 Summary of sample sizes: 225, 225, 225, 225, 225, 225, ...   
   
 Resampling results across tuning parameters:  
   
 nprune Accuracy Kappa Accuracy SD Kappa SD   
 2 0.5015368 0.03112774 0.04447235 0.06885291  
 37 0.4923910 0.12518750 0.05192218 0.06916835  
 72 0.4839631 0.12240884 0.05371793 0.07259106  
   
 Tuning parameter 'degree' was held constant at a value of 1  
 Kappa was used to select the optimal model using the largest value.  
 The final values used for the model were degree = 1 and nprune = 37.

Confusion Matrix and Statistics for Testing set  
   
 Reference  
 Prediction Mild None Severe  
 Mild 18 10 5  
 None 7 10 1  
 Severe 4 1 0  
   
 Overall Statistics  
   
 Accuracy : 0.5   
 95% CI : (0.3634, 0.6366)  
 No Information Rate : 0.5179   
 P-Value [Acc > NIR] : 0.6562   
   
 Kappa : 0.1146   
 Mcnemar's Test P-Value : 0.8871   
   
 Statistics by Class:  
   
 Class: Mild Class: None Class: Severe  
 Sensitivity 0.6207 0.4762 0.00000  
 Specificity 0.4444 0.7714 0.90000

**4.Support Vector Machines with Radial Basis Function Kernel**   
   
 225 samples  
 105 predictors  
 3 classes: 'Mild', 'None', 'Severe'   
   
 Pre-processing : Center and Scale  
 Resampling: Bootstrapped (25 reps)   
   
 Summary of sample sizes: 225, 225, 225, 225, 225, 225, ...   
   
 Resampling results across tuning parameters:  
   
 C Accuracy Kappa Accuracy SD Kappa SD   
 0.0625 0.5145799 0.000000000 0.03848358 0.00000000  
 0.1250 0.5145799 0.000000000 0.03848358 0.00000000  
 0.2500 0.5136192 0.001083215 0.03693514 0.00978591  
 0.5000 0.5183040 0.035416932 0.03409514 0.04878818  
 1.0000 0.5396650 0.111770968 0.03903374 0.07316135  
 2.0000 0.5581065 0.171586409 0.05365454 0.09837376  
 4.0000 0.5638349 0.201461792 0.04848978 0.08507142  
 8.0000 0.5527656 0.195408110 0.04879124 0.08445991  
 16.0000 0.5491559 0.194837858 0.05111048 0.08646477  
   
 Tuning parameter 'sigma' was held constant at a value of 0.002809725  
 Kappa was used to select the optimal model using the largest value.  
 The final values used for the model were sigma = 0.002809725 and C = 4.

Confusion Matrix and Statistics **Testing set**  
   
 Reference  
 Prediction Mild None Severe  
 Mild 23 13 6  
 None 6 8 0  
 Severe 0 0 0

Overall Statistics  
   
 Accuracy : 0.5536   
 95% CI : (0.4147, 0.6866)  
 No Information Rate : 0.5179   
 P-Value [Acc > NIR] : 0.3448   
   
 Kappa : 0.1379   
 Mcnemar's Test P-Value : NA   
   
 Statistics by Class:  
   
 Class: Mild Class: None Class: Severe  
 Sensitivity 0.7931 0.3810 0.0000  
 Specificity 0.2963 0.8286 1.0000

Confusion Matrix and Statistics **Testing set**  
   
 Reference  
 Prediction Mild None Severe  
 Mild 19 9 3  
 None 10 12 2  
 Severe 0 0 1  
   
 Overall Statistics  
   
 Accuracy : 0.5714   
 95% CI : (0.4322, 0.7029)  
 No Information Rate : 0.5179   
 P-Value [Acc > NIR] : 0.2524   
   
 Kappa : 0.2218   
 Statistics by Class:  
   
 Class: Mild Class: None Class: Severe  
 Sensitivity 0.6552 0.5714 0.16667  
 Specificity 0.5556 0.6571 1.00000

**5. k-Nearest Neighbors**

225 samples

105 predictors

3 classes: 'Mild', 'None', 'Severe'

Pre-processing : Center and Scale

Resampling: Bootstrapped (25 reps)

Summary of sample sizes: 225, 225, 225, 225, 225, 225, ...

Resampling results across tuning parameters:

k Accuracy Kappa Accuracy SD Kappa SD

3 0.5454926 0.2131093798 0.04922864 0.085427893

5 0.4854528 0.1079461141 0.04748937 0.069487519

9 0.4861334 0.0839134159 0.05149011 0.089374177

13 0.4918352 0.0866274922 0.04852894 0.080135955

17 0.4907008 0.0854200474 0.06183916 0.092388082

21 0.4873642 0.0756337978 0.05701704 0.084778743

Kappa was used to select the optimal model using the largest value.

The final value used for the model was k = 3

Confusion Matrix and Statistics **Testing set**  
   
 Reference  
 Prediction Mild None Severe  
 Mild 23 12 3  
 None 5 5 2  
 Severe 1 4 1  
   
 Overall Statistics  
   
 Accuracy : 0.5179   
 95% CI : (0.3803, 0.6534)  
 No Information Rate : 0.5179   
 P-Value [Acc > NIR] : 0.5537   
   
 Kappa : 0.134   
 Statistics by Class:  
   
 Class: Mild Class: None Class: Severe  
 Sensitivity 0.7931 0.23810 0.16667  
 Specificity 0.4444 0.80000 0.90000

**6. Naive Bayes**

225 samples

105 predictors

3 classes: 'Mild', 'None', 'Severe'

Pre-processing : Center and Scale

Resampling: Bootstrapped (25 reps)

Summary of sample sizes: 225, 225, 225, 225, 225, 225, ...

Resampling results across tuning parameters:

usekernel Accuracy Kappa Accuracy SD Kappa SD

FALSE NaN NaN NA NA

TRUE 0.5012535 0.06762174 0.05173594 0.07834722

Tuning parameter 'fL' was held constant at a value of 0

Kappa was used to select the optimal model using the largest value.

The final values used for the model were fL = 0 and usekernel = TRUE.

Confusion Matrix and Statistics **Testing set**

Reference

Prediction Mild None Severe

Mild 25 15 5

None 4 5 1

Severe 0 1 0

Overall Statistic

Accuracy : 0.5357

Kappa : 0.0985

**For Testing set:**

|  |  |  |
| --- | --- | --- |
| LINEAR | Accuracy | Kappa |
| LDA | 0.5179 | 0.102 |
| PLSDA | 0.5357 | 0.145 |
| NSC | 0.51 | 0 |

|  |  |  |
| --- | --- | --- |
| Non Linear Model | Accuracy | Kappa |
| MDA | 0.4107 | -0.012 |
| NNEt | 0.5389 | 0.1982 |
| FDA | 0.5 | 0.1146 |
| SVM | 0.5536 | 0.1379 |
| KNN | 0.5179 | 0.134 |
| Naïve Bayes | 0.5357 | 0.0985 |

Non Lenear model Neural network is better for the Chemical predictor comparing all the models linear and non linear.