**Implementing Classifiers for a Data set**

1. **Classifiers to Train**

|  |  |
| --- | --- |
| **Classifier** | **Package** |
| Decision Tree | Tree |
| Perceptron | neuralnet |
| Neural Net | neuralnet |
| SVM | kernlab |
| Naïve Bayes | e1071 |

1. **Classifier Parameters**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Experiment # | Classifier | Train/Test Ratio | P1 | P2 | P3 | P4 | Outcome  (Accuracy)  Training  Data | Outcome  (Accuracy)  Testing  Data |
| 1 | SVM  (rbfdot) | 258/145 | Sigma  =0.01 |  |  |  | 96.5 | 97.9 |
| 2 | SVM  (polydot) | 258/145 | Degree  =1 | Offset  =1 | Scale  =2 |  | 96.1 | 98.6 |
| 3 | SVM  (tanhdot) | 258/145 | Scale  =1 | Offset  =1 |  |  | 55.4 | 52.4 |
| 4 | SVM  (vanilladot) | 258/145 |  |  |  |  | 41.4 | 34.4 |
| 5 | SVM  (besseldot) | 258/145 | Sigma  =1 | Order  =1 | Degree=1 |  | 97.6 | 95.1 |
| 6 | Decision Tree | 258/145 |  |  |  |  | 96.1 | 92.4 |
| 7 | Decision Tree  (Pruned) | 258/145 | Best  =10 |  |  |  | 96.1 | 92.4 |
| 8 | Decision Tree  (Pruned) | 258/145 | Best  =8 |  |  |  | 94.9 | 93.1 |
| 9 | Decision Tree  (Pruned) | 258/145 | Best  =6 |  |  |  | 92.2 | 89.6 |
| 10 | Naïve Bayes | 258/145 | Target label | Features |  |  | 90.3 | 84.1 |
| 11 | Perceptron | 258/145 | Hidden  =0 | Error  =8.58 |  |  |  |  |
| 12 | Perceptron  (Caret library) | 258/145 | Maxit  =1000 | tuneGrid  =my.grid | Trace  =F | Linout  =1 | 23.5 |  |
| 13 | Neural net | 248/145 | Hidden=2 | Error  =5.84 | Error  =2.5 |  |  |  |
| 14 | Neural net | 248/145 | Hidden=4 |  |  |  |  |  |

P=Parameter

1. **Experimental Methodology**

The data set chosen by us has two different files for training and testing attached in the file.

1. **Results**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sampling # | Train/Test  Ratio | Decision Tree Accuracy | Naïve Bayes Accuracy | SVM Accuracy | Perceptron Accuracy | Neural  Net  Accuracy |
| 1 | 258/145 | 92.4 | 84.1 | 97.9 | 23.5 |  |
| 2 | 258/145 | 93.1 |  | 98.6 |  |  |
| 3 | 258/145 | 89.6 |  | 52.4 |  |  |
| 4 | 258/145 |  |  | 34.4 |  |  |
| 5 |  |  |  | 95.1 |  |  |

From the table above we can see that SVM gives the highest accuracy of 97.9 % on our data set and hence it is the most preferred one.