**Machine Learning Classifiers**

**Dataset 1: Credit dataset**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Accuracy** | | | | | | | | | | |
| **Method** | **Best Parameters** | **Sample** | | | | | | | | | | **Average of 10 Sample** |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **Decision Tree** | Split | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **Support Vector Machines** | Default kernal | 97.3 | 97.8 | 98.3 | 97.5 | 99 | 98.4 | 99 | 96 | 98.2 | 99 | 98.05 |
| **Naïve Bayesian** | Gaussian | 98.2 | 96 | 97.5 | 97 | 98.6 | 98 | 97 | 97.5 | 98 | 96 | 97.38 |
| **kNN** | K =11 | 78 | 78 | 77 | 79 | 81.5 | 80 | 85 | 83.2 | 79.7 | 81.5 | 80.29 |
| **Logistic Regression** | Family Binomial | 76.55 | 78.95 | 79.75 | 77.25 | 74.25 | 80.45 | 78.65 | 73.5 | 75.55 | 75.65 | 77.05 |
| **Neural Network** | Threshold | 99 | 84.5 | 80.5 | 84.5 | 80.5 | 99 | 80.5 | 84.5 | 80.5 | 80.5 | 85.4 |
| **Random Forest** | Trees | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **Bagging** | No of Iterations | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| **Boosting** | No of Iterations | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Note – For the first dataset Decision Trees, Random Forest, Bagging and Boosting Classifiers are the best.

**Dataset 2: Graduate School admission dataset**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Accuracy** | | | | | | | | | | |
| **Method** | **Best Parameters** | **Sample** | | | | | | | | | | **Average of 10 Sample** |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **Decision Tree** | Split | 62.5 | 72.5 | 75 | 77.5 | 62.5 | 67.2 | 62.5 | 75 | 72.5 | 62.5 | 68.97 |
| **Support Vector Machines** | Default kernal | 80 | 82.5 | 62.5 | 75 | 65 | 72.5 | 65 | 75 | 67.2 | 77.5 | 72.22 |
| **Naïve Bayesian** | Gaussian | 57 | 72.5 | 77.5 | 67.5 | 57.5 | 77.5 | 72.5 | 67 | 72.5 | 77.5 | 69.9 |
| **kNN** | K =11 | 92.5 | 90 | 87.5 | 82.5 | 85 | 92.5 | 90 | 80 | 82.5 | 95 | 87.75 |
| **Logistic Regression** | Family Binomial | 69.05 | 66.55 | 67.25 | 64.45 | 66.75 | 66.25 | 69.55 | 63.75 | 69.55 | 65.35 | 66.85 |
| **Neural Network** | Threshold | 70 | 57.5 | 70 | 57.5 | 57.5 | 62.5 | 70 | 57.5 | 62.5 | 70 | 63.3 |
| **Boosting** | Trees | 70 | 65 | 65 | 70 | 55 | 72.5 | 72.5 | 65 | 62.5 | 67.5 | 66.5 |
| **Bagging** | No of Iterations | 75 | 75 | 57.5 | 72.5 | 60 | 65 | 57.5 | 67.5 | 57.5 | 52.5 | 64 |
| **Random Forest** | No of Iterations | 72.5 | 80 | 65 | 75 | 62.5 | 80 | 70 | 62.5 | 67.5 | 57.5 | 69.25 |

Note – For the second dataset kNN is the best classifier.

**Dataset 3: Wisconsin Prognostic Breast Cancer (WPBC) dataset**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Accuracy** | | | | | | | | | | |
| **Method** | **Best Parameters** | **Sample** | | | | | | | | | | **Average of 10 Sample** |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **Decision Tree** | Split | 65 | 75 | 60 | 75 | 60 | 75 | 65 | 75 | 60 | 75 | 68.5 |
| **Support Vector Machines** | Default kernal | 70 | 65 | 60 | 75 | 60 | 85 | 70 | 85 | 60 | 65 | 69.5 |
| **Naïve Bayesian** | Gaussian | 65 | 85 | 70 | 65 | 50 | 60 | 65 | 60 | 50 | 85 | 65.5 |
| **kNN** | K =11 | 55 | 60 | 55 | 60 | 40 | 75 | 55 | 75 | 40 | 60 | 57.5 |
| **Logistic Regression** | Family Binomial | 80 | 80 | 80 | 75 | 70 | 85 | 80 | 85 | 70 | 80 | 78.5 |
| **Neural Network** | Threshold | 65 | 65 | 60 | 70 | 60 | 85 | 65 | 85 | 60 | 65 | 68 |
| **Random Forest** | Trees | 75 | 60 | 65 | 70 | 70 | 85 | 75 | 85 | 70 | 60 | 71.5 |
| **Bagging** | No of Iterations | 65 | 75 | 60 | 65 | 65 | 70 | 65 | 70 | 65 | 75 | 67.5 |
| **Boosting** | No of Iterations | 75 | 80 | 75 | 75 | 60 | 70 | 75 | 70 | 60 | 80 | 72 |

Note – For the third dataset, Logistic Regression is the best classifier.

**Dataset 4: Wisconsin Diagnostic Breast Cancer (WDBC) dataset**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Accuracy** | | | | | | | | | | |
| **Method** | **Best Parameters** | **Sample** | | | | | | | | | | **Average of 10 Sample** |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **Decision Tree** | Split | 94.73 | 92.98 | 91.22 | 96.49 | 98.24 | 98.24 | 92.98 | 98.24 | 92.98 | 94.73 | 95.083 |
| **Support Vector Machines** | Default kernal | 98.24 | 100 | 98.24 | 96.49 | 100 | 100 | 100 | 98.24 | 98.24 | 96.49 | 98.594 |
| **Naïve Bayesian** | Gaussian | 96.49 | 92.98 | 94.73 | 89.47 | 92.98 | 96.49 | 94.73 | 98.24 | 92.98 | 94.73 | 94.382 |
| **kNN** | K =11 | 84.21 | 82.45 | 92.98 | 85.96 | 85.96 | 82.45 | 91.22 | 84.21 | 82.45 | 87.70 | 85.959 |
| **Logistic Regression** | Family Binomial | 96.49 | 94.73 | 89.47 | 92.98 | 96.49 | 92.98 | 94.73 | 96.49 | 96.49 | 96.49 | 94.734 |
| **Neural Network** | Threshold | 98.24 | 96.49 | 57.89 | 91.22 | 82.45 | 96.49 | 70.17 | 94.73 | 68.24 | 64.91 | 82.083 |
| **Random Forest** | Trees | 98.24 | 96.49 | 96.49 | 98.24 | 94.73 | 98.24 | 96.49 | 98.24 | 92.98 | 94.73 | 96.487 |
| **Bagging** | No of Iterations | 98.24 | 96.49 | 96.49 | 92.98 | 96.49 | 98.24 | 92.98 | 96.49 | 92.98 | 94.73 | 95.611 |
| **Boosting** | No of Iterations | 98.24 | 94.73 | 92.98 | 96.49 | 94.73 | 100 | 100 | 98.24 | 91.22 | 96.49 | 96.312 |

Note – For the fourth dataset, SVM Dataset is the best classifier.

**Dataset 5: Ionosphere dataset**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Accuracy** | | | | | | | | | | |
| **Method** | **Best Parameters** | **Sample** | | | | | | | | | | **Average of 10 Sample** |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **Decision Tree** | Split | 91.66 | 86.11 | 97.22 | 86.11 | 86.11 | 83.33 | 83.33 | 86.11 | 100 | 88.88 | 88.886 |
| **Support Vector Machines** | Default kernal | 91.66 | 88.88 | 97.22 | 91.66 | 86.11 | 88.88 | 97.22 | 97.22 | 100 | 91.66 | 93.051 |
| **Naïve Bayesian** | Gaussian | 83.33 | 80.55 | 94.44 | 80.55 | 86.11 | 83.33 | 91.66 | 91.66 | 97.22 | 91.66 | 88.051 |
| **kNN** | K =11 | 80.55 | 77.77 | 88.88 | 86.11 | 77.77 | 77.77 | 88.88 | 86.11 | 91.66 | 77.77 | 83.327 |
| **Logistic Regression** | Family Binomial | 88.88 | 75 | 80.55 | 86.11 | 77.77 | 88.88 | 83.33 | 88.88 | 83.33 | 83.33 | 83.606 |
| **Neural Network** | Threshold | 91.66 | 77.77 | 88.88 | 88.88 | 83.33 | 83.33 | 80.55 | 86.11 | 83.33 | 83.33 | 84.717 |
| **Random Forest** | Trees | 94.44 | 94.44 | 97.22 | 88.88 | 86.11 | 91.66 | 94.44 | 94.44 | 100 | 91.66 | 93.329 |
| **Bagging** | No of Iterations | 94.44 | 91.66 | 97.22 | 88.88 | 80.55 | 88.88 | 88.88 | 91.66 | 97.22 | 94.44 | 91.383 |
| **Boosting** | No of Iterations | 86.11 | 86.11 | 91.66 | 91.66 | 88.88 | 80.55 | 91.66 | 97.22 | 94.44 | 91.66 | 89.995 |

Note – For the fifth dataset, Random Forest is the best classifier.