For Question 1, I worked with six different classifiers: k-Nearest Neighbor, Support vector classifier, Extra tree classifier, Bagging, Decision Tree Classifier, and Logistic Regression. The following tables are the results for Question 2.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Accuracy | Balanced Accuracy | Matthews Correlation Coefficient | Sensitivity | Specificity | F-1 Score | Confusion  Matrix |
| 96.7% | 96.7% | 95% | 96.7% | 98.3% | 96.6% | [[50 0 0]  [ 0 47 3]  [ 0 2 48]] |

**kNN**

**SVC**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Accuracy | Balanced Accuracy | Matthews Correlation Coefficient | Sensitivity | Specificity | F-1 Score | Confusion  Matrix |
| **96.7%** | **96.7%** | **95%** | **96.7%** | **98.3%** | **96.6%** | **[[50 0 0]**  **[ 0 47 3]**  **[ 0 2 48]]** |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Accuracy | Balanced Accuracy | Matthews Correlation  Coefficient | Sensitivity | Specificity | F-1 Score | Confusion Matrix |
| **97.3%** | **97.3%** | **96%** | **97.3%** | **98.7%** | **97.3%** | **[[50 0 0]**  **[ 0 47 3]**  **[ 0 1 49]]** |

**ETC**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Accuracy | Balanced Accuracy | Matthews Correlation  Coefficient | Sensitivity | Specificity | F-1 Score | Confusion Matrix |
| **93.3%** | **93.3%** | **90%** | **93.3%** | **96.7%** | **93.3%** | **[[48 2 0]**  **[ 0 45 5]**  **[ 0 3 47]]** |

**Bagging**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Accuracy | Balanced Accuracy | Matthews Correlation  Coefficient | Sensitivity | Specificity | F-1 Score | Confusion Matrix |
| **95.3%** | **95.3%** | **93%** | **95.3%** | **97.7%** | **95.3%** | **[[50 0 0]**  **[ 0 47 3]**  **[ 0 4 46]]** |

**DTC**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Accuracy | Balanced Accuracy | Matthews Correlation  Coefficient | Sensitivity | Specificity | F-1 Score | Confusion Matrix |
| **95.3%** | **94.7%** | **93.1%** | **95.3%** | **97.7%** | **95.3%** | **[[50 0 0]**  **[ 0 48 2]**  **[ 0 5 45]]** |

**Logistic Regression**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Accuracy | Balanced Accuracy | Matthews Correlation  Coefficient | Sensitivity | Specificity | F-1 Score | Confusion Matrix |
| **95.3%** | **95.3%** | **93.1%** | **95.3%** | **97.7%** | **95.3%** | **[[50 0 0]**  **[ 0 45 5]**  **[ 0 2 48]]** |

For my first stacked model, I used kNN, Bagging, and ETC as my beta classifiers. Logistic Regression was used in combination as my meta classifier. The following table showcases the results.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Accuracy | Balanced Accuracy | Matthews Correlation  Coefficient | Sensitivity | Specificity | F-1 Score | Confusion Matrix |
| **97%** | **97%** | **95%** | **97%** | **98.3%** | **97%** | **[[50 0 0]**  **[ 0 47 3]**  **[ 0 2 48]]** |

For my second stacked model, I used kNN, Decision Tree, and ETC as my beta classifiers. Logistic Regression was used in combination as my meta classifier. The following table showcases the results.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Accuracy | Balanced Accuracy | Matthews Correlation  Coefficient | Sensitivity | Specificity | F-1 Score | Confusion Matrix |
| **96%** | **96%** | **94%** | **96%** | **98%** | **96%** | **[[50 0 0]**  **[ 0 46 4]**  **[ 0 2 48]]** |