**Balanced Resampling (Class 0 down sample, Class 1 up sample):**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Algorithm:** | **Classification score:** | **ROC Score:** | **Accuracy** | | **P** | **R** | **F1** |
| Nearest Neighbors | 0.9673038229376257 | 0.9809991994016262 | 0.97 | 0 | 0.99 | 0.95 | 0.97 |
| 1 | 0.95 | 0.99 | 0.97 |
| Logistic Regression | 0.9009054325955734 | 0.9707938474369213 | 0.90 | 0 | 0.86 | 0.96 | 0.91 |
| 1 | 0.95 | 0.84 | 0.90 |
| Linear SVM | 0.8948692152917505 | 0.9678100749890437 | 0.89 | 0 | 0.86 | 0.94 | 0.90 |
| 1 | 0.94 | 0.85 | 0.89 |
| RBF SVM | 0.9290744466800804 | 0.9759931822621409 | 0.93 | 0 | 0.92 | 0.94 | 0.93 |
| 1 | 0.94 | 0.92 | 0.93 |
| Decision Tree | 0.9698189134808853 | 0.9847289149614733 | 0.97 | 0 | 0.98 | 0.96 | 0.97 |
| 1 | 0.96 | 0.98 | 0.97 |
| Random Forest | 0.9668008048289738 | 0.9886028596847606 | 0.97 | 0 | 0.98 | 0.96 | 0.97 |
| 1 | 0.96 | 0.98 | 0.97 |
| Neural Net | 0.9175050301810865 | 0.9762310338650076 | 0.92 | 0 | 0.89 | 0.95 | 0.92 |
| 1 | 0.95 | 0.89 | 0.92 |
| Ada Boost | 0.9587525150905433 | 0.9879388994656944 | 0.96 | 0 | 0.96 | 0.96 | 0.96 |
| 1 | 0.96 | 0.96 | 0.96 |
| Naive Bayes | 0.8415492957746479 | 0.9494944894338221 | 0.84 | 0 | 0.77 | 0.96 | 0.86 |
| 1 | 0.95 | 0.72 | 0.82 |

**No Class Imbalance Handling:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Algorithm:** | **Classification score:** | **ROC Score:** | **Accuracy** | | **P** | **R** | **F1** |
| Nearest Neighbors | 0.95 | 0.9141838723163231 | 0.95 | 0 | 0.96 | 0.98 | 0.97 |
| 1 | 0.79 | 0.64 | 0.71 |
| Logistic Regression | 0.9481060606060606 | 0.9526218106052431 | 0.95 | 0 | 0.96 | 0.99 | 0.97 |
| 1 | 0.81 | 0.58 | 0.68 |
| Linear SVM | 0.9488636363636364 | 0.9513490870104648 | 0.95 | 0 | 0.96 | 0.99 | 0.97 |
| 1 | 0.82 | 0.58 | 0.68 |
| RBF SVM | 0.9443181818181818 | 0.935740944276621 | 0.94 | 0 | 0.97 | 0.96 | 0.97 |
| 1 | 0.68 | 0.76 | 0.72 |
| Decision Tree | 0.9647727272727272 | 0.970931330240587 | 0.96 | 0 | 0.98 | 0.98 | 0.98 |
| 1 | 0.80 | 0.83 | 0.82 |
| Random Forest | 0.9670454545454545 | 0.9798294381810335 | 0.97 | 0 | 0.99 | 0.98 | 0.98 |
| 1 | 0.79 | 0.88 | 0.83 |
| Neural Net | 0.9503787878787879 | 0.959595661613982 | 0.95 | 0 | 0.96 | 0.98 | 0.97 |
| 1 | 0.81 | 0.62 | 0.70 |
| Ada Boost | 0.9606060606060606 | 0.9781732117812061 | 0.96 | 0 | 0.98 | 0.98 | 0.98 |
| 1 | 0.79 | 0.79 | 0.79 |
| Naive Bayes | 0.928030303030303 | 0.9344041630704498 | 0.93 | 0 | 0.97 | 0.95 | 0.96 |
| 1 | 0.60 | 0.69 | 0.64 |

**Class weight = balanced**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Algorithm:** | **Classification score:** | **ROC Score:** | **Accuracy** | | **P** | **R** | **F1** |
| Logistic Regression | 0.9238636363636363 | 0.9596024045204445 | 0.92 | 0 | 0.98 | 0.93 | 0.96 |
|  | 1 | 0.57 | 0.82 | 0.67 |
| Linear SVM | 0.9083333333333333 | 0.9575525609558744 | 0.91 | 0 | 0.98 | 0.92 | 0.95 |
| 1 | 0.51 | 0.83 | 0.63 |
| RBF SVM | 0.9128787878787878 | 0.948796896914446 | 0.91 | 0 | 0.99 | 0.91 | 0.95 |
| 1 | 0.52 | 0.90 | 0.66 |
| Decision Tree | 0.9522727272727273 | 0.9758393232819074 | 0.95 | 0 | 1.00 | 0.95 | 0.97 |
| 1 | 0.67 | 0.96 | 0.79 |
| Random Forest | 0.956060606060606 | 0.9821102262919408 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.70 | 0.94 | 0.80 |

**Under Sample**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Algorithm:** | **Classification score:** | **ROC Score:** | **Accuracy** | | **P** | **R** | **F1** |
| Nearest Neighbors | 0.9122383252818036 | 0.9360315116505484 | 0.91 | 0 | 0.93 | 0.96 | 0.95 |
| 1 | 0.81 | 0.73 | 0.77 |
| Logistic Regression | 0.9106280193236715 | 0.9522335626663205 | 0.91 | 0 | 0.92 | 0.97 | 0.95 |
| 1 | 0.84 | 0.69 | 0.75 |
| Linear SVM | 0.9114331723027376 | 0.9640463750243395 | 0.91 | 0 | 0.92 | 0.98 | 0.95 |
| 1 | 0.88 | 0.65 | 0.75 |
| RBF SVM | 0.9122383252818036 | 0.9634845362497566 | 0.91 | 0 | 0.95 | 0.93 | 0.94 |
| 1 | 0.76 | 0.82 | 0.79 |
| Decision Tree | 0.9565217391304348 | 0.9738471149477511 | 0.96 | 0 | 0.99 | 0.96 | 0.97 |
| 1 | 0.85 | 0.95 | 0.90 |
| Random Forest | 0.9549114331723028 | 0.9870959628740184 | 0.95 | 0 | 0.99 | 0.95 | 0.97 |
| 1 | 0.83 | 0.97 | 0.90 |
| Neural Net | 0.9178743961352657 | 0.9647035438437075 | 0.92 | 0 | 0.94 | 0.96 | 0.95 |
| 1 | 0.83 | 0.74 | 0.78 |
| Ada Boost | 0.9476650563607085 | 0.9779321087817225 | 0.95 | 0 | 0.98 | 0.95 | 0.97 |
| 1 | 0.83 | 0.93 | 0.88 |
| Naive Bayes | 0.8816425120772947 | 0.9240199260076589 | 0.88 | 0 | 0.91 | 0.94 | 0.93 |
| 1 | 0.74 | 0.63 | 0.68 |

**Class weight = balanced**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Algorithm:** | **Classification score:** | **ROC Score:** | **Accuracy** | | **P** | **R** | **F1** |
| Logistic Regression | 0.9122383252818036 | 0.9623101512299603 | 0.91 | 0 | 0.95 | 0.94 | 0.94 |
| 1 | 0.76 | 0.81 | 0.79 |
| Linear SVM | 0.9033816425120773 | 0.9627847731550594 | 0.90 | 0 | 0.96 | 0.92 | 0.94 |
| 1 | 0.73 | 0.83 | 0.77 |
| RBF SVM | 0.9009661835748792 | 0.9581541669371065 | 0.90 | 0 | 0.97 | 0.90 | 0.94 |
| 1 | 0.70 | 0.90 | 0.78 |
| Decision Tree | 0.9533011272141707 | 0.9731473518530538 | 0.95 | 0 | 0.99 | 0.95 | 0.97 |
| 1 | 0.82 | 0.98 | 0.89 |
| Random Forest | 0.9516908212560387 | 0.9852664373336796 | 0.95 | 0 | 0.99 | 0.95 | 0.97 |
| 1 | 0.82 | 0.98 | 0.89 |

**SMOTE for Class Imbalance:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Algorithm:** | **Classification score:** | **ROC Score:** | **Accuracy** | | **P** | **R** | **F1** |
| Nearest Neighbors | 0.9314393939393939 | 0.9092817793181573 | 0.93 | 0 | 0.98 | 0.95 | 0.96 |
| 1 | 0.60 | 0.79 | 0.68 |
| Logistic Regression | 0.9306818181818182 | 0.9556055467148561 | 0.93 | 0 | 0.98 | 0.94 | 0.96 |
| 1 | 0.59 | 0.83 | 0.69 |
| Linear SVM | 0.9272727272727272 | 0.9505062237026649 | 0.93 | 0 | 0.98 | 0.94 | 0.96 |
| 1 | 0.58 | 0.84 | 0.69 |
| RBF SVM | 0.918939393939394 | 0.9474348298090409 | 0.92 | 0 | 0.99 | 0.92 | 0.95 |
| 1 | 0.54 | 0.90 | 0.68 |
| Decision Tree | 0.9568181818181818 | 0.9716519783687562 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.70 | 0.95 | 0.80 |
| Random Forest | 0.9571969696969697 | 0.9841196124177366 | 0.96 | 0 | 1.00 | 0.96 | 0.98 |
| 1 | 0.70 | 0.96 | 0.81 |
| Neural Net | 0.9272727272727272 | 0.9662627103786817 | 0.93 | 0 | 0.99 | 0.93 | 0.96 |
| 1 | 0.57 | 0.87 | 0.69 |
| Ada Boost | 0.9583333333333334 | 0.9791905477937211 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.71 | 0.93 | 0.81 |
| Naive Bayes | 0.9253787878787879 | 0.9371906691660374 | 0.93 | 0 | 0.97 | 0.95 | 0.96 |
| 1 | 0.58 | 0.73 | 0.65 |

**Borderline SMOTE:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Algorithm:** | **Classification score:** | **ROC Score:** | **Accuracy** | | **P** | **R** | **F1** |
| Nearest Neighbors | 0.9325757575757576 | 0.9062820288056964 | 0.93 | 0 | 0.98 | 0.95 | 0.96 |
| 1 | 0.61 | 0.79 | 0.69 |
| Logistic Regression | 0.9068181818181819 | 0.9483046647426909 | 0.91 | 0 | 0.98 | 0.91 | 0.95 |
| 1 | 0.50 | 0.85 | 0.63 |
| Linear SVM | 0.8988636363636363 | 0.9465675234653146 | 0.90 | 0 | 0.98 | 0.90 | 0.94 |
| 1 | 0.48 | 0.86 | 0.62 |
| RBF SVM | 0.8931818181818182 | 0.914547989265293 | 0.89 | 0 | 0.99 | 0.89 | 0.94 |
| 1 | 0.47 | 0.91 | 0.62 |
| Decision Tree | 0.9598484848484848 | 0.9755358924910994 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.72 | 0.94 | 0.81 |
| Random Forest | 0.9571969696969697 | 0.9839510397561765 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.70 | 0.95 | 0.81 |
| Neural Net | 0.8810606060606061 | 0.9626257552055238 | 0.88 | 0 | 0.99 | 0.88 | 0.93 |
| 1 | 0.44 | 0.93 | 0.59 |
| Ada Boost | 0.9571969696969697 | 0.9768027160427231 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.71 | 0.92 | 0.80 |
| Naive Bayes | 0.9162878787878788 | 0.93369531502859 | 0.92 | 0 | 0.97 | 0.93 | 0.95 |
| 1 | 0.54 | 0.75 | 0.63 |