|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm: | Classification score: | ROC Score: | Accuracy | | P | R | F1 |
| Nearest Neighbors | 0.9515151515151515 | 0.9083942442550437 | 0.95 | 0 | 0.97 | 0.98 | 0.97 |
| 1 | 0.78 | 0.67 | 0.72 |
| Logistic Regression | 0.9484848484848485 | 0.9475663164850577 | 0.95 | 0 | 0.96 | 0.99 | 0.97 |
| 1 | 0.81 | 0.59 | 0.68 |
| Linear SVM | 0.9511363636363637 | 0.9400395471464018 | 0.95 | 0 | 0.96 | 0.99 | 0.97 |
| 1 | 0.82 | 0.61 | 0.70 |
| RBF SVM | 0.9454545454545454 | 0.9375016857266157 | 0.95 | 0 | 0.97 | 0.97 | 0.97 |
| 1 | 0.70 | 0.75 | 0.72 |
| Decision Tree | 0.9640151515151515 | 0.9677748071528751 | 0.96 | 0 | 0.99 | 0.97 | 0.98 |
| 1 | 0.77 | 0.87 | 0.82 |
| Random Forest | 0.9651515151515152 | 0.9799592391304348 | 0.97 | 0 | 0.99 | 0.98 | 0.98 |
| 1 | 0.78 | 0.87 | 0.82 |
| Neural Net | 0.9496212121212121 | 0.9607840988779803 | 0.95 | 0 | 0.96 | 0.99 | 0.97 |
| 1 | 0.81 | 0.60 | 0.69 |
| Ada Boost | 0.9647727272727272 | 0.9770319748624446 | 0.96 | 0 | 0.98 | 0.98 | 0.98 |
| 1 | 0.81 | 0.82 | 0.81 |
| Naive Bayes | 0.9272727272727272 | 0.9321840611177042 | 0.93 | 0 | 0.97 | 0.94 | 0.96 |
| 1 | 0.59 | 0.77 | 0.66 |

# No class imbalance handling

# Class weight balanced

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm: | Classification score: | ROC Score: | Accuracy | | P | R | F1 |
| Logistic Regression | 0.934469696969697 | 0.9580043356888553 | 0.93 | 0 | 0.98 | 0.95 | 0.96 |
| 1 | 0.61 | 0.82 | 0.70 |
| Linear SVM | 0.931060606060606 | 0.95633968065595 | 0.93 | 0 | 0.98 | 0.94 | 0.96 |
| 1 | 0.60 | 0.83 | 0.69 |
| RBF SVM | 0.9208333333333333 | 0.9477382605998489 | 0.92 | 0 | 0.99 | 0.92 | 0.95 |
| 1 | 0.55 | 0.88 | 0.68 |
| Decision Tree | 0.9541666666666667 | 0.9673719184917465 | 0.95 | 0 | 1.00 | 0.95 | 0.97 |
| 1 | 0.68 | 0.96 | 0.80 |
| Random Forest | 0.9568181818181818 | 0.9785272143704824 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.70 | 0.94 | 0.80 |

# Balanced Resampling (Class 0 down sample, Class 1 up sample)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm: | Classification score: | ROC Score: | Accuracy | | P | R | F1 |
| Nearest Neighbors | 0.928409090909091 | 0.9247660211457546 | 0.93 | 0 | 0.98 | 0.94 | 0.96 |
| 1 | 0.58 | 0.83 | 0.69 |
| Logistic Regression | 0.9337121212121212 | 0.9573620738483117 | 0.93 | 0 | 0.98 | 0.95 | 0.96 |
| 1 | 0.61 | 0.81 | 0.70 |
| Linear SVM | 0.931060606060606 | 0.956305966123638 | 0.93 | 0 | 0.98 | 0.94 | 0.96 |
| 1 | 0.60 | 0.83 | 0.69 |
| RBF SVM | 0.9196969696969697 | 0.9462716784442766 | 0.92 | 0 | 0.99 | 0.92 | 0.95 |
| 1 | 0.55 | 0.88 | 0.67 |
| Decision Tree | 0.9575757575757575 | 0.9740988105513 | 0.96 | 0 | 1.00 | 0.96 | 0.98 |
| 1 | 0.70 | 0.96 | 0.81 |
| Random Forest | 0.9556818181818182 | 0.9790641182975509 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.70 | 0.94 | 0.80 |
| Neural Net | 0.9318181818181818 | 0.9663672254288487 | 0.93 | 0 | 0.98 | 0.94 | 0.96 |
| 1 | 0.60 | 0.85 | 0.70 |
| Ada Boost | 0.9556818181818182 | 0.9776826653360664 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.70 | 0.94 | 0.80 |
| Naive Bayes | 0.925 | 0.9322675045851764 | 0.93 | 0 | 0.97 | 0.94 | 0.96 |
| 1 | 0.58 | 0.76 | 0.66 |

# Under Sample

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm: | Classification score: | ROC Score: | Accuracy | | P | R | F1 |
| Nearest Neighbors | 0.95 | 0.928961794691984 | 0.95 | 0 | 0.98 | 0.97 | 0.97 |
| 1 | 0.71 | 0.79 | 0.75 |
| Logistic Regression | 0.9507575757575758 | 0.9445640373826735 | 0.95 | 0 | 0.97 | 0.98 | 0.97 |
| 1 | 0.77 | 0.67 | 0.72 |
| Linear SVM | 0.9541666666666667 | 0.9548570840975293 | 0.95 | 0 | 0.97 | 0.98 | 0.97 |
| 1 | 0.80 | 0.69 | 0.74 |
| RBF SVM | 0.9352272727272727 | 0.9634745185564786 | 0.94 | 0 | 0.98 | 0.95 | 0.96 |
| 1 | 0.61 | 0.83 | 0.71 |
| Decision Tree | 0.9598484848484848 | 0.9569751995900313 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.72 | 0.92 | 0.81 |
| Random Forest | 0.9621212121212122 | 0.9792175194195707 | 0.96 | 0 | 0.99 | 0.97 | 0.98 |
| 1 | 0.74 | 0.93 | 0.82 |
| Neural Net | 0.9507575757575758 | 0.9612999312223541 | 0.95 | 0 | 0.97 | 0.98 | 0.97 |
| 1 | 0.77 | 0.68 | 0.72 |
| Ada Boost | 0.9636363636363636 | 0.977267133725321 | 0.96 | 0 | 0.99 | 0.97 | 0.98 |
| 1 | 0.77 | 0.88 | 0.82 |
| Naive Bayes | 0.9215909090909091 | 0.9321655181249326 | 0.92 | 0 | 0.97 | 0.95 | 0.96 |
| 1 | 0.57 | 0.69 | 0.62 |

# Class weight Balanced

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm: | Classification score: | ROC Score: | Accuracy | | P | R | F1 |
| Logistic Regression | 0.9363636363636364 | 0.9579605067968497 | 0.94 | 0 | 0.98 | 0.95 | 0.96 |
| 1 | 0.62 | 0.83 | 0.71 |
| Linear SVM | 0.9314393939393939 | 0.955723547577948 | 0.93 | 0 | 0.98 | 0.94 | 0.96 |
| 1 | 0.60 | 0.83 | 0.69 |
| RBF SVM | 0.9193181818181818 | 0.954027706602654 | 0.92 | 0 | 0.99 | 0.92 | 0.95 |
| 1 | 0.54 | 0.89 | 0.67 |
| Decision Tree | 0.946969696969697 | 0.9578340773006796 | 0.95 | 0 | 1.00 | 0.95 | 0.97 |
| 1 | 0.65 | 0.96 | 0.77 |
| Random Forest | 0.9568181818181818 | 0.9782187263998274 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.70 | 0.94 | 0.80 |

# SMOTE

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm: | Classification score: | ROC Score: | Accuracy | | P | R | F1 |
| Nearest Neighbors | 0.9340909090909091 | 0.9298282581724024 | 0.93 | 0 | 0.98 | 0.95 | 0.96 |
| 1 | 0.61 | 0.82 | 0.70 |
| Logistic Regression | 0.9363636363636364 | 0.9529269271226671 | 0.94 | 0 | 0.98 | 0.95 | 0.96 |
| 1 | 0.63 | 0.80 | 0.70 |
| Linear SVM | 0.9329545454545455 | 0.9572870590139176 | 0.93 | 0 | 0.98 | 0.94 | 0.96 |
| 1 | 0.60 | 0.83 | 0.70 |
| RBF SVM | 0.9272727272727272 | 0.9458350752508362 | 0.93 | 0 | 0.99 | 0.93 | 0.96 |
| 1 | 0.57 | 0.88 | 0.70 |
| Decision Tree | 0.9556818181818182 | 0.971741321879383 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.70 | 0.93 | 0.80 |
| Random Forest | 0.9583333333333334 | 0.9781243257093538 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.71 | 0.94 | 0.81 |
| Neural Net | 0.9318181818181818 | 0.9639380933757685 | 0.93 | 0 | 0.98 | 0.94 | 0.96 |
| 1 | 0.60 | 0.85 | 0.70 |
| Ada Boost | 0.959469696969697 | 0.9761469683892547 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.72 | 0.94 | 0.81 |
| Naive Bayes | 0.925 | 0.9366925369511275 | 0.93 | 0 | 0.97 | 0.94 | 0.96 |
| 1 | 0.58 | 0.77 | 0.66 |

# Borderline Smote

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Algorithm: | Classification score: | ROC Score: | Accuracy | | P | R | F1 |
| Nearest Neighbors | 0.9371212121212121 | 0.9235590408889848 | 0.94 | 0 | 0.98 | 0.95 | 0.96 |
| 1 | 0.62 | 0.83 | 0.71 |
| Logistic Regression | 0.9181818181818182 | 0.9448826397130221 | 0.92 | 0 | 0.98 | 0.93 | 0.95 |
| 1 | 0.54 | 0.85 | 0.66 |
| Linear SVM | 0.9053030303030303 | 0.9440414621318374 | 0.91 | 0 | 0.98 | 0.91 | 0.95 |
| 1 | 0.50 | 0.85 | 0.63 |
| RBF SVM | 0.8984848484848484 | 0.9367422658862876 | 0.90 | 0 | 0.99 | 0.90 | 0.94 |
| 1 | 0.48 | 0.90 | 0.62 |
| Decision Tree | 0.959469696969697 | 0.9725740708274895 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.71 | 0.95 | 0.82 |
| Random Forest | 0.9575757575757575 | 0.9734885775164526 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.71 | 0.94 | 0.81 |
| Neural Net | 0.9015151515151515 | 0.959432146132269 | 0.90 | 0 | 0.99 | 0.90 | 0.94 |
| 1 | 0.49 | 0.88 | 0.63 |
| Ada Boost | 0.9579545454545455 | 0.9766467863307801 | 0.96 | 0 | 0.99 | 0.96 | 0.98 |
| 1 | 0.71 | 0.94 | 0.81 |
| Naive Bayes | 0.9181818181818182 | 0.930945894918546 | 0.92 | 0 | 0.98 | 0.93 | 0.95 |
| 1 | 0.54 | 0.79 | 0.64 |