For LS

Raw DataSet

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Accuracy | Recall | F1 | Auc | Mcc |
| Decision Tree （CART） | 0.8999 | 0.8238 | 0.8303 | 0.8785 | 0.7701 |
| Random Forest | 0.9001 | 0.8071 | 0.8265 | 0.9587 | 0.7834 |
| LogisticRegression\_liblinear | 0.8282 | 0.6904 | 0.6993 | 0.8879 | 0.5970 |
| LogisticRegression\_lbfgs | 0.7698 | 0.5762 | 0.5943 | 0.8824 | 0.6485 |
| LogisticRegression\_newton-cg | 0.8728 | 0.7905 | 0.7907 | 0.9487 | 0.7739 |
| LogisticRegression\_sag | / | / | / | / | / |
| LogisticRegression\_saga | / | / | / | / | / |
| SVM\_SVC\_linear | 0.8635 | 0.7619 | 0.7691 | 0.9497 | 0.7643 |
| SVM\_SVC\_poly | 0.8680 | 0.7904 | 0.7806 | 0.9430 | 0.7945 |
| Naïve\_bayes\_GaussianNB | 0.8776 | 0.9381 | 0.7968 | 0.8914 | 0.7222 |
| Naïve\_bayes\_BernoulliNB | 0.9268 | 0.7405 | 0.8391 | 0.9345 | 0.8312 |
| XGBoost | 0.8775 | 0.8643 | 0.8765 | 0.9615 | 0.8056 |
| AdaBoost | 0.8860 | 0.7190 | 0.7640 | 0.9694 | 0.7692 |
| KNN | 0.8776 | 0.7643 | 0.7982 | 0.9514 | 0.6056 |

PCA

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Accuracy | Recall | F1 | Auc | Mcc |
| Decision Tree （CART） | 0.8503 | 0.7833 | 0.7603 | 0.8937 | 0.7460 |
| Random Forest | 0.8830 | 0.7413 | 0.7841 | 0.9648 | 0.8249 |
| LogisticRegression\_liblinear | 0.8959 | 0.8095 | 0.8011 | 0.9702 | 0.6573 |
| LogisticRegression\_lbfgs | 0.8688 | 0.8095 | 0.7851 | 0.9720 | 0.6892 |
| LogisticRegression\_newton-cg | 0.8869 | 0.8095 | 0.7851 | 0.9684 | 0.7410 |
| LogisticRegression\_sag | / | / | / | / | / |
| LogisticRegression\_saga | / | / | / | / | / |
| SVM\_SVC\_linear | 0.8860 | 0.8857 | 0.8342 | 0.9781 | 0.7855 |
| SVM\_SVC\_poly | 0.8774 | 0.8314 | 0.8022 | 0.9610 | 0.7946 |
| Naïve\_bayes\_GaussianNB | 0.7981 | 0.9429 | 0.7643 | 0.8921 | 0.7405 |
| Naïve\_bayes\_BernoulliNB | 0.8065 | 0.6405 | 0.6702 | 0.8619 | 0.7946 |
| XGBoost | 0.8997 | 0.8238 | 0.8335 | 0.9627 | 0.8326 |
| AdaBoost | 0.8964 | 0.8548 | 0.8313 | 0.9569 | 0.8175 |
| KNN | 0.8691 | 0.8119 | 0.7882 | 0.9396 | 0.7361 |

F16

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Accuracy | Recall | F1 | Auc | Mcc |
| Decision Tree （CART） | 0.9318 | 0.8547 | 0.8717 | 0.9550 | 0.8165 |
| Random Forest | 0.9275 | 0.8809 | 0.8809 | 0.9478 | 0.8276 |
| LogisticRegression\_liblinear | 0.9179 | 0.8524 | 0.8674 | 0.9686 | 0.7813 |
| LogisticRegression\_lbfgs | 0.8826 | 0.8690 | 0.8273 | 0.9440 | 0.7409 |
| LogisticRegression\_newton-cg | 0.9179 | 0.8667 | 0.8674 | 0.9711 | 0.8158 |
| LogisticRegression\_sag | / | / | / | / | / |
| LogisticRegression\_saga | / | / | / | / | / |
| SVM\_SVC\_linear | 0.8955 | 0.8271 | 0.8332 | 0.9616 | 0.8133 |
| SVM\_SVC\_poly | 0.8819 | 0.7985 | 0.8121 | 0.9696 | 0.8046 |
| Naïve\_bayes\_GaussianNB | 0.8549 | 0.9714 | 0.8183 | 0.9662 | 0.8258 |
| Naïve\_bayes\_BernoulliNB | 0.9144 | 0.7952 | 0.8474 | 0.9553 | 0.8474 |
| XGBoost | 0.9274 | 0.8810 | 0.8772 | 0.9795 | 0.8678 |
| AdaBoost | 0.9458 | 0.9000 | 0.9094 | 0.9793 | 0.8432 |
| KNN | 0.9131 | 0.9857 | 0.8402 | 0.9688 | 0.8052 |