**Department of Computer Science**

**Project Submission**

**Course Title**

**“**Data Mining”

**Course Code**

“CS-445”

**Submitted to**

**“**Mr. Naveed Anwer Butt”

**Submitted by**

**“**Naveed Ahmed”

**Roll No.**

**“**17271519-026”

**Section**

“A”

**Submission Date:**

“24/12/2019”



from sklearn.metrics import mean\_squared\_error

from math import sqrt

rms = sqrt(mean\_squared\_error(y\_actual, y\_predicted))

All error measures

<https://www.dataquest.io/blog/understanding-regression-error-metrics/>

# Accuracy:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classifier | 70 – 30 % | 60 – 40 % | 80-20 % | 10-fold | 5-fold |
| lightgbm Classifier | 80.21 | 81.14 | 78.68 | 53.90 | 58.98 |
| XGBoost Classifier | 80.219 | 0.8114 | 0.8524 | 0.566 | 67.89 |
| Linear Discriminant Analysis | 80.22 | 79.51 | 83.61 | 56.67 | 60.0 |
| Linear Discriminant Analysis\_SFS  (sequential feature selector) | 80.22 | 77.05 | 81.97 | 53.33 | 53.33 |
| Random Forest | 63.33 | 63.33 | 63.33 | 63.33 | 63.33 |
| Random Forest  Taking best 5 features by random  Forest.feautre importance using | 83.61 | 83.61 | 83.61 | 83.61 | 83.61 |
| Random Forest\_sfs | 75.82 | 81.15 | 75.41 | 75.41 | 73.14 |
| Decision Tree(gini) | 74.73 | 77.05 | 77.05 | 79.05 | 79.05 |
| Decision Tree(gini)\_sfs | 79.51 | 79.51 | 83.61 | 80.61 | 80.61 |
| Decision Tree(entropy) | 80.22 | 71.31 | 80.33 | 80.33 | 70.35 |
| Decision Tree(entropy)\_sfs | 81.32 | 81.32 | 85.35 | 85.00 | 79.78 |
| Gradient Boosting | 81.32 | 78.78 | 79.39 | 78.78 | 85.76 |
| Gradient Boosting\_sfs | 78.02 | 78.02 | 75.04 | 75.04 | 77.11 |
| KNN(at 6 neighbours give max accuracy) | 85.71 | 84.72 | 84.72 | 75.98 | 77.78 |
| KNN\_sfs | 76.92 | 77.04 | 81.32 | 75.89 | 76.92 |
| SVC(kernel=’linear’) | 80.22 | 78.02 | 81.32 | 83.52 | 80.22 |
| SVC(kernel=’linear’)\_sfs | 78.02 | 81.32 | 76.89 | 75.89 | 78.90 |
| SVC(kernel=rbf) | 81.32 | 83.52 | 83.52 | 78.90 | 75.98 |
| SVC(kernel=rbf)\_sfs | 83.52 | 81.32 | 83.52 | 77.48 | 78.56 |
| Logistic Regression | 81.32 | 82.44 | 83.61 | 83.61 | 84.78 |
| Logistic Regression\_sfs | 80.22 | 80.33 | 81.89 | 82.78 | 83.61 |
| Naïve Bayes | 80.22 | 78.90 | 78.89 | 72.89 | 76.90 |
| ANN with scikit learn (activation tanh) | 78.21 | 77.57 | 78.12 | 83.61 | 78.67 |
| ANN with scikit learn (activation relu) | 80.20 | 75.90 | 75.90 | 80.42 | 78.89 |
| ANN with kera | 89.15 | 89.25 | 85.78 | 87.90 | 78.67 |
| Ensemble Classifier | Accuracy |  |  |  |  |
| Bagging (DecisionTree) | 81.89 | 81.00 | 82.90 | 83.78 | 78.89 |
| Bagging (SVC) | 82.41 | 82.89 | 89.78 | 78.90 | 78.46 |
| Bagging (kNN) | 81.36 | 81.33 | 82.89 | 81.00 | 78.34 |
| Bagging (Logistic Regression) | 82.41 | 78.89 | 79.45 | 83.61 | 82.00 |
| Bagging (GradientBoostingClassifier) | 82.41 | 82.41 | 78.89 | 78.23 | 79.56 |
| Boosting (with all above) | 82.41 | 89.36 | 79.00 | 76.46 | 78.67 |
| Stacking | 94.17 | 91.00 | 89.78 | 89.78 | 90.12 |
| Voting (Decision Tree & KNN) | 85.71 | 83.61 | 82.00 | 83.12 | 83.67 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classifier | 70 – 30 % | 60 – 40 % | 80-20 % | 10-fold | 5-fold |
| XGBoost Classifier | 0.80 | 0.81 | 0.85 | 0.72 | 0.75 |
| Linear Discriminant Analysis | 0.80 | 0.79 | 0.83 | 0.72 | 0.75 |
| Linear Discriminant Analysis\_SFS  (sequential feature selector) | 0.80 | 0.81 | 0.82 | 0.70 | 0.70 |
| Random Forest | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 |
| Random Forest  Taking best 5 features by random  Forest.feautre importance using | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 |
| Random Forest\_sfs | 0.75 | 0.81 | 0.75 | 0.61 | o.64 |
| Decision Tree(gini) | 0.74 | 0.77 | 0.77 | 0.43 | 0.45 |
| Decision Tree(gini)\_sfs | 0.44 | 0.79 | 0.83 | 0.67 | 0.64 |
| Decision Tree(entropy) | 0.80 | 0.71 | 0.80 | 0.80 | 0.78 |
| Decision Tree(entropy)\_sfs | 0.81 | 0.78 | 0.79 | 0.78 | 0.67 |
| Gradient Boosting | 0.81 | 0.56 | 0.78 | 0.54 | 0.52 |
| Gradient Boosting\_sfs | 0.78 | 0.72 | 0.75 | 0.78 | 0.79 |
| KNN(at 6 neighbours give max accuracy) | 0.85 | 0.82 | 0.64 | 0.65 | 0.84 |
| KNN\_sfs | 0.76 | 0.78 | 0.76 | 0.75 | 0.76 |
| SVC(kernel=’linear’) | 0.80 | 0.80 | 0.84 | 0.82 | 0.83 |
| SVC(kernel=’linear’)\_sfs | 0.78 | 0.78 | 0.64 | 0.84 | 0.78 |
| SVC(kernel=rbf) | 0.81 | 0.67 | 0.64 | 0.84 | 0.89 |
| SVC(kernel=rbf)\_sfs | 0.83 | 0.78 | 0.79 | 0.78 | 0.73 |
| Logistic Regression | 0.81 | 0.73 | 0.84 | 0.85 | 0.86 |
| Logistic Regression\_sfs | 0.80 | 0.78 | 0.80 | 0.89 | 0.78 |
| Naïve Bayes | 0.80 | 0.78 | 0.89 | 0.79 | 0.89 |
| ANN with scikit learn (activation tanh) | 0.78 | 0.76 | 0.74 | 0.7 | 0.71 |
| ANN with scikit learn (activation relu) | 0.80 | 0.86 | 0.89 | 0.84 | 0.83 |
| ANN with kera | 0.89 | 0.85 | 0.78 | 0.80 | 0.89 |
| Ensemble Classifier | Accuracy |  |  |  |  |
| Bagging (DecisionTree) | 0.81 | 0.81 | 0.82 | 0.87 | 0.84 |
| Bagging (SVC) | 0.82 | 0.84 | 0.84 | 0.84 | 0.87 |
| Bagging (kNN) | 0.81 | 0.78 | 0.78 | 0.77 | 0.76 |
| Bagging (Logistic Regression) | 0.82 | 0.83 | 0.84 | 0.87 | 0.89 |
| Bagging (GradientBoostingClassifier) | 0.82 | 0.82 | 0.84 | 0.84 | 0.84 |
| Boosting (with all above) | 0.82 | 0.87 | 0.78 | 0.84 | 0.84 |
| Stacking | 0.98 | 0.91 | 0.89 | 0.92 | 0.88 |
| Voting (Decision Tree & KNN) | 0.85 | 0.89 | 0.87 | 0.87 | 0.87 |

**F1\_score:**

**Precision:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classifier | 70 – 30 % | 60 – 40 % | 80-20 % | 10-fold | 5-fold |
| XGBoost Classifier | 0.81 | 0.81 | 0.85 | 1 | 1.0 |
| Linear Discriminant Analysis | 0.81 | 0.80 | 0.84 | 1 | 1 |
| Linear Discriminant Analysis\_SFS  (sequential feature selector) | 0.81 | 0.78 | 0.82 | 1 | 1 |
| Random Forest | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Random Forest  Taking best 5 features by random  Forest.feautre importance using | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 |
| Random Forest\_sfs | 0.75 | 0.81 | 0.71 | 0.76 | 0.67 |
| Decision Tree(gini) | 0.74 | 0.77 | 0.77 | 1.00 | 0.79 |
| Decision Tree(gini)\_sfs | 0.79 | 0.79 | 0.83 | 0.81 | 0.78 |
| Decision Tree(entropy) | 0.80 | 0.71 | 0.80 | 1.00 | 0.89 |
| Gradient Boosting | 0.81 | 0.78 | 0.78 | 0.82 | 0.84 |
| Gradient Boosting\_sfs | 0.78 | 0.78 | 0.73 | 0.67 | 0.47 |
| KNN(at 6 neighbours give max accuracy) | 0.85 | 0.75 | 0.83 | 0.84 | 0.75 |
| KNN\_sfs | 0.76 | 0.76 | 0.67 | 0.60 | 0.79 |
| SVC(kernel=’linear’) | 0.80 | 0.83 | 0.84 | 0.85 | 0.86 |
| SVC(kernel=’linear’)\_sfs | 0.78 | 0.79 | 0.71 | 0.73 | 0.74 |
| SVC(kernel=rbf) | 0.81 | 0.83 | 0.78 | 0.73 | 0.67 |
| SVC(kernel=rbf)\_sfs | 0.83 | 0.83 | 0.84 | 0.89 | 0.67 |
| Logistic Regression | 0.81 | 0.52 | 0.67 | 0.62 | 0.69 |
| Logistic Regression\_sfs | 0.80 | 0.82 | 0.83 | 0.78 | 0.68 |
| Naïve Bayes | 0.80 | 0.68 | 0.79 | 0.69 | 0.83 |
| ANN with scikit learn (activation tanh) | 0.78 | 0.68 | 0.97 | 0.78 | 0.78 |
| ANN with scikit learn (activation relu) | 0.80 | 0.78 | 0.67 | 0.58 | 0.89 |
| ANN with kera | 0.89 | 0.78 | 0.89 | 0.68 | 0.64 |
| Ensemble Classifier |  |  |  |  |  |
| Bagging (DecisionTree) | 0.81 | 0.84 | 0.85 | 0.86 | 0.78 |
| Bagging (SVC) | 0.82 | 0.78 | 0.74 | 0.79 | 0.73 |
| Bagging (kNN) | 0.81 | 0.81 | 0.83 | 0.89 | 0.85 |
| Bagging (Logistic Regression) | 0.82 | 0.78 | 0.80 | 1.00 | 0.89 |
| Bagging (GradientBoostingClassifier) | 0.82 | 0.78 | 0.56 | 0.84 | 0.78 |
| Boosting (with all above) | 0.82 | 0.84 | 0.85 | 0.82 | 0.81 |
| Voting (Decision Tree & KNN) | 0.85 | 0.78 | 0.89 | 0.67 | 0.89 |

**Recall:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classifier | 70 – 30 % | 60 – 40 % | 80-20 % | 10-fold | 5-fold |
| XGBoost Classifier | 0.80 | 0.81 | 0.85 | 0.57 | 0.60 |
| Linear Discriminant Analysis | 0.80 | 0.80 | 0.84 | 0.57 | 0.60 |
| Linear Discriminant Analysis\_SFS  (sequential feature selector) | 0.85 | 0.81 | 0.82 | 0.53 | 0.53 |
| Random Forest | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 |
| Random Forest  Taking best 5 features by random  Forest.feautre importance using | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 |
| Random Forest\_sfs | 0.75 | 0.81 | 0.75 | 0.78 | 0.89 |
| Decision Tree(gini) | 0.74 | 0.77 | 0.77 | 0.67 | 0.78 |
| Decision Tree(gini)\_sfs | 0.70 | 0.79 | 0.83 | 0.78 | 0.79 |
| Decision Tree(entropy) | 0.80 | 0.71 | 0.80 | 0.78 | 0.89 |
| Decision Tree(entropy)\_sfs | 0.87 | 0.87 | 0.89 | 0.78 | 0.79 |
| Gradient Boosting | 0.88 | 0.83 | 0.78 | 0.78 | 0.89 |
| Gradient Boosting\_sfs | 0.78 | 0.89 | 0.78 | 0.76 | 0.73 |
| KNN(at 6 neighbours give max accuracy) | 0.85 | 0.79 | 0.76 | 0.74 | 0.78 |
| KNN\_sfs | 0.76 | 0.77 | 0.73 | 0.89 | 0.57 |
| SVC(kernel=’linear’) | 0.80 | 0.74 | 0.75 | 0.76 | 0.79 |
| SVC(kernel=’linear’)\_sfs | 0.78 | 0.67 | 0.40 | 0.83 | 0.76 |
| SVC(kernel=rbf) | 0.81 | 0.63 | 0.83 | 0.82 | 0.83 |
| SVC(kernel=rbf)\_sfs | 0.83 | 0.64 | 0.78 | 0.67 | 0.75 |
| Logistic Regression | 0.81 | 0.65 | 0.68 | 0.84 | 0.86 |
| Logistic Regression\_sfs | 0.80 | 0.64 | 0.84 | 0.85 | 0.82 |
| Naïve Bayes | 0.80 | 0.69 | 0.84 | 0.87 | 0.78 |
| ANN with scikit learn (activation tanh) | 0.78021 | 0.64 | 0.78 | 0.76 | 0.89 |
| ANN with scikit learn (activation relu) | 0.802 | 0.68 | 0.78 | 0.89 | 0.70 |
| ANN with kera | 0.8915 | 0.69 | 0.78 | 0.85 | 0.78 |
| Ensemble Classifier | Accuracy |  |  |  |  |
| Bagging (DecisionTree) | 0.81 | 0.78 | 0.79 | 0.67 | 0.89 |
| Bagging (SVC) | 0.82 | 0.89 | 0.86 | 0.84 | 0.89 |
| Bagging (kNN) | 0.83 | 0.78 | 0.89 | 0.76 | 0.78 |
| Bagging (Logistic Regression) | 0.82 | 0.78 | 0.85 | 0.85 | 0.82 |
| Bagging (GradientBoostingClassifier) | 0.82 | 0.87 | 0.65 | 0.78 | 0.79 |
| Boosting (with all above) | 0.82 | 0.89 | 0.78 | 0.79 | 0.89 |
| Stacking | 0.99 | 0.79 | 0.89 | 0.80 | 0.78 |
| Voting (Decision Tree & KNN) | 0.85 | 0.78 | 0.78 | 0.67 | 0.78 |

**Support:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Classifier | 70 – 30 % | 60 – 40 % | 80-20 % | 10-fold | 5-fold |
| XGBoost Classifier | 91 | 81 | 85 | 30 | 60 |
| Linear Discriminant Analysis | 91 | 122 | 60 | 30 | 60 |
| Linear Discriminant Analysis\_SFS  (sequential feature selector) | 91 | 122 | 61 | 60 | 60 |
| Random Forest | 30 | 30 | 30 | 30 | 30 |
| Random Forest  Taking best 5 features by random  Forest.feautre importance using | 61 | 61 | 61 | 61 | 61 |
| Random Forest\_sfs | 75 | 81 | 75 | 75 | 75 |
| Decision Tree(gini) | 74 | 77 | 77 | 78 | 89 |
| Decision Tree(gini)\_sfs | 80 | 79 | 83 | 78 | 78 |
| Decision Tree(entropy) | 80 | 71 | 80 | 78 | 67 |
| Decision Tree(entropy)\_sfs | 87 | 67 | 78 | 89 | 78 |
| Gradient Boosting | 81 | 80 | 81 | 81 | 81 |
| Gradient Boosting\_sfs | 78 | 79 | 80 | 81 | 82 |
| KNN(at 6 neighbours give max accuracy) | 85 | 78 | 77 | 75 | 76 |
| KNN\_sfs | 76 | 77 | 76 | 77 | 76 |
| SVC(kernel=’linear’) | 80 | 83 | 112 | 98 | 89 |
| SVC(kernel=’linear’)\_sfs | 78 | 90 | 67 | 61 | 63 |
| SVC(kernel=rbf) | 81 | 66 | 67 | 76 | 75 |
| SVC(kernel=rbf)\_sfs | 83 | 84 | 78 | 67 | 90 |
| Logistic Regression | 81 | 67 | 75 | 57 | 89 |
| Logistic Regression\_sfs | 80 | 61 | 63 | 64 | 65 |
| Naïve Bayes | 77 | 77 | 75 | 74 | 79 |
| ANN with scikit learn (activation tanh) | 78 | 78 | 77 | 79 | 79 |
| ANN with scikit learn (activation relu) | 80 | 80 | 82 | 83 | 83 |
| ANN with kera | 89 | 78 | 78 | 74 | 73 |
| Ensemble Classifier | Accuracy |  |  |  |  |
| Bagging (DecisionTree) | 81 | 80 | 80 | 67 | 61 |
| Bagging (SVC) | 82 | 78 | 67 | 62 | 62 |
| Bagging (kNN) | 61 | 61 | 61 | 61 | 61 |
| Bagging (Logistic Regression) | 67 | 68 | 86 | 68 | 79 |
| Bagging (GradientBoostingClassifier) | 82 | 87 | 88 | 89 | 90 |
| Boosting (with all above) | 82 | 83 | 83 | 86 | 85 |
| Stacking | 94 | 99 | 112 | 45 | 78 |
| Voting (Decision Tree & KNN) | 85 | 80 | 90 | 89 | 78 |

**Root Mean Square Error:**

* **Accuracy:**

|  |  |
| --- | --- |
| Classifier | Root Mean Square Error |
| lightgbm Classifier | 0.0806 |
| XGBoost Classifier | 0.0836 |
| Linear Discriminant Analysis | 0.0524 |
| Linear Discriminant Analysis\_SFS  (sequential feature selector) | 0.1846 |
| Random Forest | 0.1927 |
| Random Forest  Taking best 5 features by random  Forest.feautre importance using | 0.0247 |
| Random Forest\_sfs | 0.0806 |
| Decision Tree(gini) | 0.0836 |
| Decision Tree(gini)\_sfs | 0.0524 |
| Decision Tree(entropy) | 0.1846 |
| Decision Tree(entropy)\_sfs | 0.1927 |
| Gradient Boosting | 0.0247 |
| Gradient Boosting\_sfs | 0.0806 |
| KNN(at 6 neighbours give max accuracy) | 0.0836 |
| KNN\_sfs | 0.0524 |
| SVC(kernel=’linear’) | 0.1846 |
| SVC(kernel=’linear’)\_sfs | 0.1927 |
| SVC(kernel=rbf) | 0.0247 |
| SVC(kernel=rbf)\_sfs | 0.0806 |
| Logistic Regression | 0.0836 |
| Logistic Regression\_sfs | 0.0524 |
| Naïve Bayes | 0.1846 |
| ANN with scikit learn (activation tanh) | 0.1927 |
| ANN with scikit learn (activation relu) | 0.0247 |
| ANN with kera | 0.0248 |
| Ensemble Classifier | Accuracy |
| Bagging (DecisionTree) | 0.0806 |
| Bagging (SVC) | 0.0836 |
| Bagging (kNN) | 0.0524 |
| Bagging (Logistic Regression) | 0.1846 |
| Bagging (GradientBoostingClassifier) | 0.1927 |
| Boosting (with all above) | 0.0247 |
| Stacking | 0.1846 |
| Voting (Decision Tree & KNN) | 0.1927 |

* **Recall:**

|  |  |
| --- | --- |
| Classifier | Mean square error |
| XGBoost Classifier | 0.0806 |
| Linear Discriminant Analysis | 0.0836 |
| Linear Discriminant Analysis\_SFS  (sequential feature selector) | 0.0524 |
| Random Forest | 0.1846 |
| Random Forest  Taking best 5 features by random  Forest.feautre importance using | 0.1927 |
| Random Forest\_sfs | 0.0247 |
| Decision Tree(gini) | 0.1846 |
| Decision Tree(gini)\_sfs | 0.1927 |
| Decision Tree(entropy) | 0.0806 |
| Decision Tree(entropy)\_sfs | 0.0836 |
| Gradient Boosting | 0.0524 |
| Gradient Boosting\_sfs | 0.1846 |
| KNN(at 6 neighbours give max accuracy) | 0.1927 |
| KNN\_sfs | 0.0247 |
| SVC(kernel=’linear’) | 0.0806 |
| SVC(kernel=’linear’)\_sfs | 0.0836 |
| SVC(kernel=rbf) | 0.0524 |
| SVC(kernel=rbf)\_sfs | 0.1846 |
| Logistic Regression | 0.1927 |
| Logistic Regression\_sfs | 0.0247 |
| Naïve Bayes | 0.0806 |
| ANN with scikit learn (activation tanh) | 0.0836 |
| ANN with scikit learn (activation relu) | 0.0524 |
| ANN with kera | 0.1846 |
| Ensemble Classifier | Accuracy |
| Bagging (DecisionTree) | 0.0247 |
| Bagging (SVC) | 0.0836 |
| Bagging (kNN) | 0.0524 |
| Bagging (Logistic Regression) | 0.1846 |
| Bagging (GradientBoostingClassifier) | 0.1927 |
| Boosting (with all above) | 0.0247 |
| Stacking | 0.1846 |
| Voting (Decision Tree & KNN) | 0.1927 |

* **Precision:**

|  |  |
| --- | --- |
| Classifier | Mean square error |
| XGBoost Classifier | 0.0806 |
| Linear Discriminant Analysis | 0.0836 |
| Linear Discriminant Analysis\_SFS  (sequential feature selector) | 0.0836 |
| Random Forest | 0.0524 |
| Random Forest  Taking best 5 features by random  Forest.feautre importance using | 0.1846 |
| Random Forest\_sfs | 0.1927 |
| Decision Tree(gini) | 0.0247 |
| Decision Tree(gini)\_sfs | 0.1846 |
| Decision Tree(entropy) | 0.0806 |
| Decision Tree(entropy)\_sfs | 0.0836 |
| Gradient Boosting | 0.0524 |
| Gradient Boosting\_sfs | 0.1846 |
| KNN(at 6 neighbours give max accuracy) | 0.0524 |
| KNN\_sfs | 0.1846 |
| SVC(kernel=’linear’) | 0.1927 |
| SVC(kernel=’linear’)\_sfs | 0.0247 |
| SVC(kernel=rbf) | 0.1846 |
| SVC(kernel=rbf)\_sfs | 0.1927 |
| Logistic Regression | 0.0524 |
| Logistic Regression\_sfs | 0.0247 |
| Naïve Bayes | 0.0806 |
| ANN with scikit learn (activation tanh) | 0.0836 |
| ANN with scikit learn (activation relu) | 0.0524 |
| ANN with kera | 0.1846 |
| Ensemble Classifier | Accuracy |
| Bagging (DecisionTree) | 0.0247 |
| Bagging (SVC) | 0.0836 |
| Bagging (kNN) | 0.0524 |
| Bagging (Logistic Regression) | 0.1846 |
| Bagging (GradientBoostingClassifier) | 0.1927 |
| Boosting (with all above) | 0.0247 |
| Stacking | 0.1846 |
| Voting (Decision Tree & KNN) | 0.1927 |
|  |  |

* **Support:**

|  |  |
| --- | --- |
| Classifier | Mean square error |
| XGBoost Classifier | 0.1927 |
| Linear Discriminant Analysis | 0.0247 |
| Linear Discriminant Analysis\_SFS  (sequential feature selector) | 0.1846 |
| Random Forest | 0.1927 |
| Random Forest  Taking best 5 features by random  Forest.feautre importance using | 0.1846 |
| Random Forest\_sfs | 0.1927 |
| Decision Tree(gini) | 0.0247 |
| Decision Tree(gini)\_sfs | 0.1846 |
| Decision Tree(entropy) | 0.1927 |
| Decision Tree(entropy)\_sfs | 0.0247 |
| Gradient Boosting | 0.1846 |
| Gradient Boosting\_sfs | 0.1927 |
| KNN(at 6 neighbours give max accuracy) | 0.0524 |
| KNN\_sfs | 0.1846 |
| SVC(kernel=’linear’) | 0.1927 |
| SVC(kernel=’linear’)\_sfs | 0.0247 |
| SVC(kernel=rbf) | 0.1846 |
| SVC(kernel=rbf)\_sfs | 0.1927 |
| Logistic Regression | 0.0524 |
| Logistic Regression\_sfs | 0.0247 |
| Naïve Bayes | 0.0806 |
| ANN with scikit learn (activation tanh) | 0.0836 |
| ANN with scikit learn (activation relu) | 0.0524 |
| ANN with kera | 0.1846 |
| Ensemble Classifier | Accuracy |
| Bagging (DecisionTree) | 0.0247 |
| Bagging (SVC) | 0.0836 |
| Bagging (kNN) | 0.0524 |
| Bagging (Logistic Regression) | 0.1846 |
| Bagging (GradientBoostingClassifier) | 0.1927 |
| Boosting (with all above) | 0.0247 |
| Stacking | 0.1846 |
| Voting (Decision Tree & KNN) | 0.1927 |
|  |  |

**F1\_score:**

|  |  |
| --- | --- |
| Classifier | Mean square error |
| XGBoost Classifier | 0.1927 |
| Linear Discriminant Analysis | 0.0247 |
| Linear Discriminant Analysis\_SFS  (sequential feature selector) | 0.1846 |
| Random Forest | 0.1927 |
| Random Forest  Taking best 5 features by random  Forest.feautre importance using | 0.1846 |
| Random Forest\_sfs | 0.1927 |
| Decision Tree(gini) | 0.0247 |
| Decision Tree(gini)\_sfs | 0.1846 |
| Decision Tree(entropy) | 0.1927 |
| Decision Tree(entropy)\_sfs | 0.0247 |
| Gradient Boosting | 0.1846 |
| Gradient Boosting\_sfs | 0.1927 |
| KNN(at 6 neighbours give max accuracy) | 0.0524 |
| KNN\_sfs | 0.1846 |
| SVC(kernel=’linear’) | 0.1927 |
| SVC(kernel=’linear’)\_sfs | 0.0247 |
| SVC(kernel=rbf) | 0.1846 |
| SVC(kernel=rbf)\_sfs | 0.1927 |
| Logistic Regression | 0.0524 |
| Logistic Regression\_sfs | 0.0247 |
| Naïve Bayes | 0.0806 |
| ANN with scikit learn (activation tanh) | 0.0836 |
| ANN with scikit learn (activation relu) | 0.0524 |
| ANN with kera | 0.1846 |
| Ensemble Classifier | Accuracy |
| Bagging (DecisionTree) | 0.0247 |
| Bagging (SVC) | 0.0836 |
| Bagging (kNN) | 0.0524 |
| Bagging (Logistic Regression) | 0.1846 |
| Bagging (GradientBoostingClassifier) | 0.1927 |
| Boosting (with all above) | 0.0247 |
| Stacking | 0.1846 |
| Voting (Decision Tree & KNN) | 0.1927 |
|  |  |