ASSIGNMENT\_CLASSIFICATION

1.Machine Learning

Supervised Learning- stage 2

Classfication- stage 3

2.399rows X25 columns and pre-processing 399 rows X 28 columns

3.Rbc,Pc,Pcc,Pa,Htn,Dm,Cad,Appet,Pc,Ane,Classification is nominal data(one hot encoding) so coverting to numerical data.

4. The final models is SVM (The f1\_macro value for best parameter{'C': 10, 'gamma': 'auto', 'kernel': 'poly'}: 1.0)

5. The research values of each algorithm

Decision Tree

F1\_Macro

The f1\_macro value for best parameter{'criterion': 'gini', 'max\_features': 'auto', 'splitter': 'random'}: 0.9547078539722802

Cif\_report

The report:

precision recall f1-score support

0 0.96 0.92 0.94 51

1 0.95 0.98 0.96 82

accuracy 0.95 133

macro avg 0.96 0.95 0.95 133

weighted avg 0.95 0.95 0.95 133

Roc\_Auc\_Score

0.9485891917742707

SVM

F1\_Macro

The f1\_macro value for best parameter{'C': 10, 'gamma': 'auto', 'kernel': 'poly'}: 1.0

Cif\_report

The report:

precision recall f1-score support

0 1.00 1.00 1.00 51

1 1.00 1.00 1.00 82

accuracy 1.00 133

macro avg 1.00 1.00 1.00 133

weighted avg 1.00 1.00 1.00 133

Roc\_Auc\_Score

1.0

Random Forest

F1\_Macro

The f1\_macro value for best parameter{'criterion': 'gini', 'max\_features': 'log2', 'n\_estimators': 100}: 0.9924667654735397

Cif\_report

The report:

precision recall f1-score support

0 1.00 0.98 0.99 51

1 0.99 1.00 0.99 82

accuracy 0.99 133

macro avg 0.99 0.99 0.99 133

weighted avg 0.99 0.99 0.99 133

Roc\_Auc\_Score

0.9998804399808704

Logistic Regression

F1\_Macro

The f1\_macro value for best parameter{'penalty': 'l2', 'solver': 'newton-cg'}: 0.9924946382275899

Cif\_report

The report:

precision recall f1-score support

0 0.98 1.00 0.99 51

1 1.00 0.99 0.99 82

accuracy 0.99 133

macro avg 0.99 0.99 0.99 133

weighted avg 0.99 0.99 0.99 133

Roc\_Auc\_Score

1.0

K\_ nearest Neighbour

F1\_Macro

The f1\_macro value for best parameter{'algorithm': 'auto', 'weights': 'distance'}: 0.9404945931261721

Cif\_report

The report:

precision recall f1-score support

0 0.86 1.00 0.93 51

1 1.00 0.90 0.95 82

accuracy 0.94 133

macro avg 0.93 0.95 0.94 133

weighted avg 0.95 0.94 0.94 133

Roc\_Auc\_Score

1.0

6. . The final models is SVM. I think compare to Decision Tree, Random Forest, Logistic Regression, and N\_ nearest neighbor is F1\_Macro,Cif\_report,and Roc\_Auc\_Score is better than to another . So that dataset best model is SVM.