CLASSIFICATION ASSIGNMENT

1.Problem statement:

The problem statement is to classify whether a person as chronic kidney disease or not.it is a classification problem under supervised learning.

2.Dataset Description

The Dataset as 399 rows and 27 columns

3.Preprocessing steps

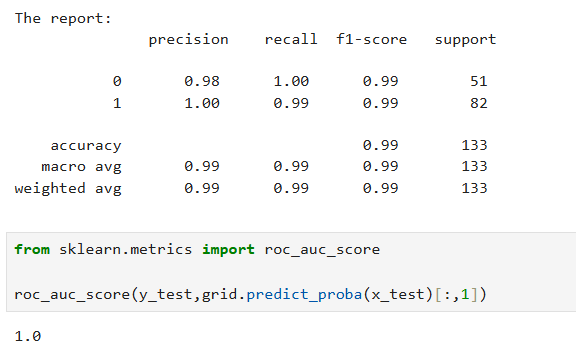
The dataset consists of strings(categorical data).They are converted to numbers using getdummies by one hot encoding.

Also the values are brought to a specified range using standardscaler.

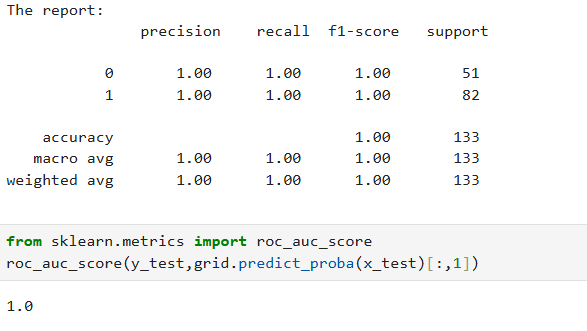
4.Training the dataset with different ML classification Algorithms(Logistic Regressor,SVM,Decision,RandomForest,KNN and Naïve Bayes)

5.Results

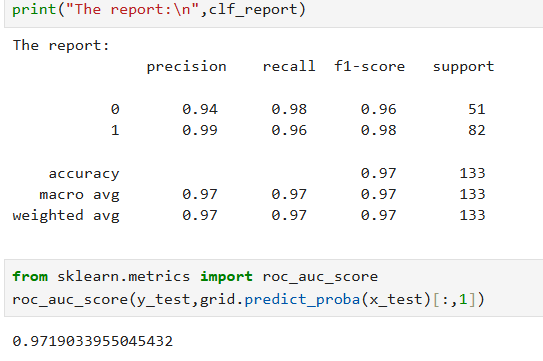
Logistics Regression



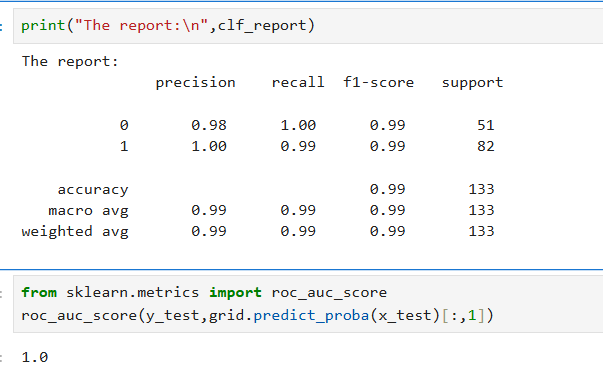
Random Forest



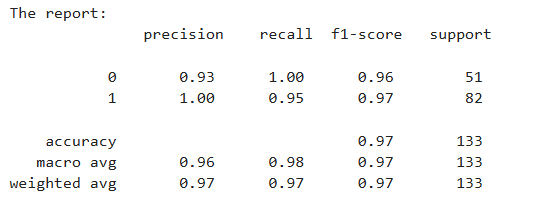
Decision tree



Support Vector Machine

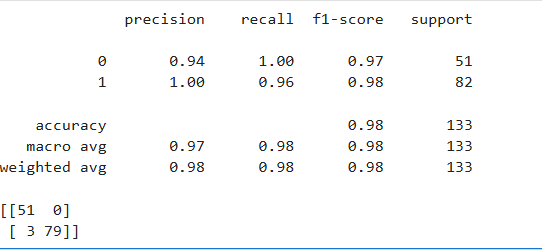


KNN-K NEAREST NEIGHBOUR(USED GRIDSEARCHCV )



Naïve Bayes

GaussianNB and BernoulliNB



6. Random Forest Classifier has given the best result

Acuracy:1.0

ROCAUC:1