**CLASSIFICATION ASSIGNMENT**

**Problem Statement or Requirement:**

A requirement from the Hospital, Management asked us to create a predictive model which will predict the Chronic Kidney Disease (CKD) based on the several parameters. The Client has provided the dataset of the same.

1. Identify your problem statement

predict the Chronic Kidney Disease (CKD)

1. Tell basic info about the dataset (Total number of rows, columns)

Total number of rows = 399, Total number of rows Columns = 25.

1. Mention the pre-processing method if you’re doing any (like converting string to number – nominal data)

The Client Given Dataset, I convert the string to a number

I use the pre-processing method **Nominal – One Hot Encoding.**

1. Develop a good model with good evaluation metric. You can use any machine learning algorithm; you can create many models. All the research values of each algorithm should be documented. And take a screenshot of the results.

**DOMAIN SELECTION** = **Machine Learning**.

**LEARNING SELECTION** = **Supervised Learning.**

**SUPERVISED LEARNING** = **Classification**.

1. **SVM-****GridSearchCV\_Classification**

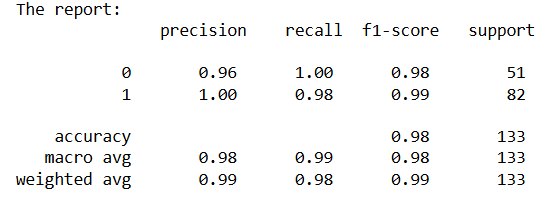
**roc\_auc\_score=1.0**

**The f1\_macro value=0.9850141736106648**

**The Confusion Matrix:**

**[[51 0]**

**[ 2 80]]**

****

1. **Decision Tree-GridSearchCV\_Classification**

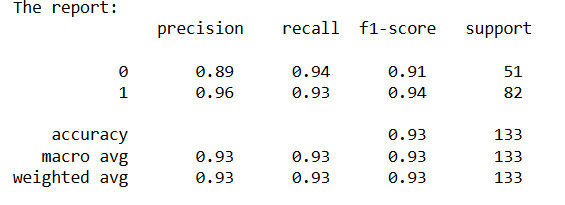
**roc\_auc\_score= 0.9340028694404591**

**The f1\_macro value= 0.9326670714052211**

**The Confusion Matrix:**

**[[48 3]**

**[ 6 76]]**

****

1. **Random Forest-GridSearchCV\_Classification**

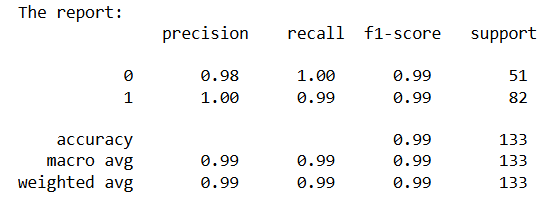
**roc\_auc\_score= 0.9503825920612148**

**The f1\_macro value= 0.9924946382275899**

**The Confusion Matrix:**

**[[51 0]**

**[ 1 81]]**

****

1. **Logistic-Grid-****Classification**

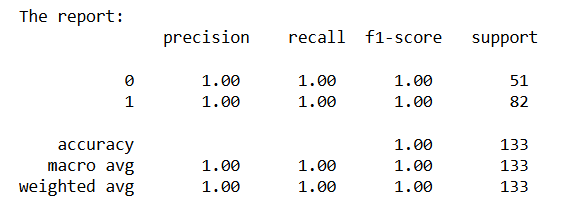
**roc\_auc\_score= 1.0**

**The f1\_macro value= 1.0**

**The Confusion Matrix:**

**[[51 0]**

**[ 0 82]]**

****

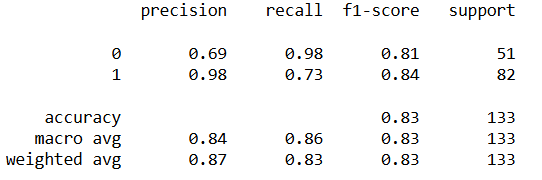
1. **Navie bayes-Classification**

**The Confusion Matrix:**

**[[50 1]**

**[22 60]]**

**The report:**

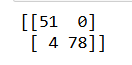
****

1. **KNN-Grid\_Search-Classification**

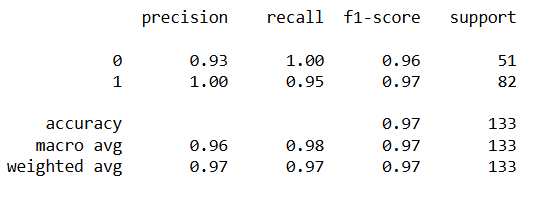
**roc\_auc\_score = 1.0**

**The f1\_macro value = 0.9701163285572423**

**The Confusion Matrix:**



**The report:**

****

1. **Mention your final model, justify why u have chosen the same.**

**The Final model is** **Logistic-Grid-Classification**

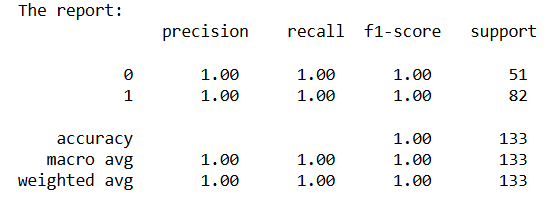
**roc\_auc\_score= 1.0**

**The f1\_macro value= 1.0**

**The Confusion Matrix:**

**[[51 0]**

**[ 0 82]]**

****

The **Logistic-Grid-Classification Roc\_Auc\_scor ,F1\_Macro Value, Confusion Matrix, and The Report**

Is **higher compared to SVM, DC, RF, KNN, and Navie Bayes. that’s why the finalized model comes to Logistic-Grid-Classification.**