**Assignment-Classification**

**1.Problem statement –**

stage1-Supervised learning

stage2-Machine learning

stage3-Classification

**2.Total number of rows-399**

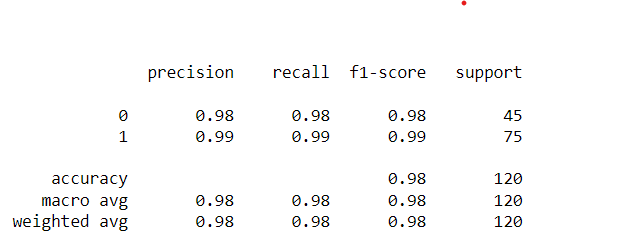
**Total number of columns-25**

**3.Changed String to number – rbc,pc,pcc,ba,htn,dm,cad,appet,ane,classification**

**4. Best model-** Random forest &GaussianNB=0.98

**To find Machine learning Classification method using confusion matrix & classificationReport**

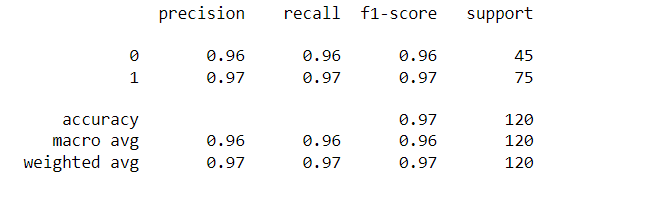
1.**Random Forest**

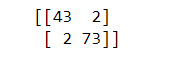




Roc\_auc score=0.999

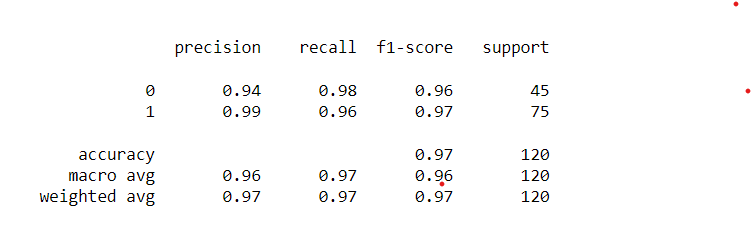
**2.Logistic Regression**

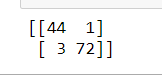




Roc\_auc score=0.997

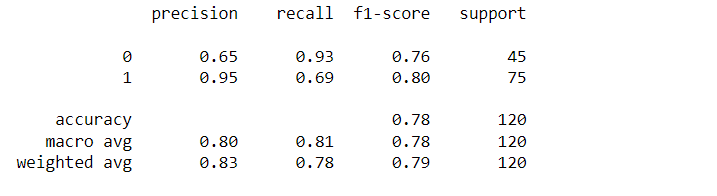
**3.Decision Tree**

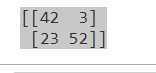




Roc\_auc score=0.968

4.KNN

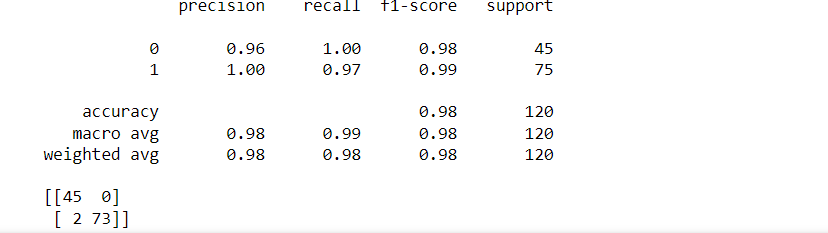




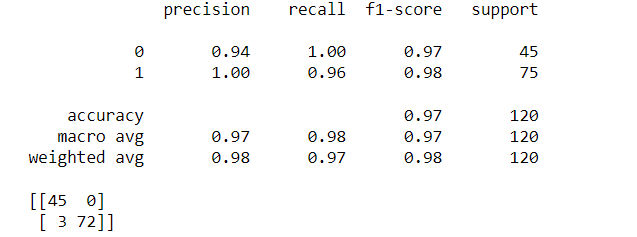
Roc\_auc score=0.864

5. Naive bayes

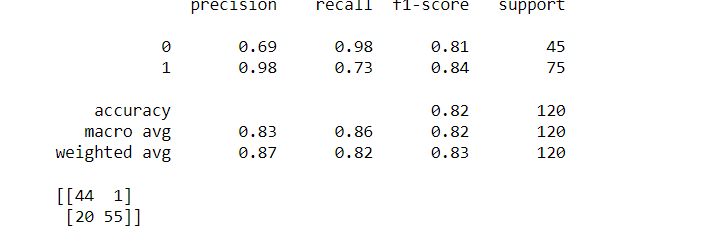
i.GaussianNB



ii. MultinomialNB

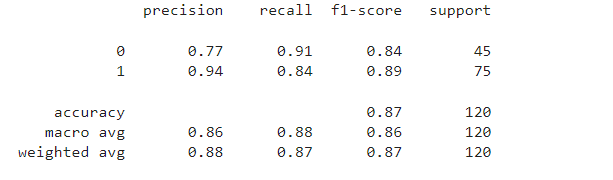


iii.BernoulliNB



Roc\_auc score=0.996

6.SVM





Roc\_auc score=0.93

As a result of comparison between multiple Classification algorthim ,We selected random forest&GaussianNB algorthim as best model overall performance Accuracy=0.98

Random forest&GaussianNB algorthim =0.98