**Assignment – Classification Algorithm**

**1. Problem Statement:**

Todevelop a model which will predict the Chronic Kidney Disease (CKD) is the Problem Statement.

**2. Info about the dataset:**

Total number of rows - 399

Total number of columns – 25

**3. Pre-processing method:**

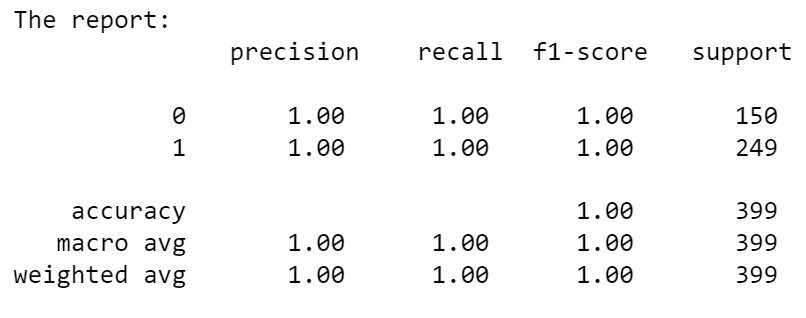
By using One Hot Encoding the categorical column (rbc, pc, pcc, ba, htn, dm, cad, appet, pe, ane, classification) is converted into nominal data as 0s and 1s.

**4. Good Model:**

After creating many models, Random Forest, Decision Tree, SVM, Logistic Regression these are the best model among them, which gives the highest roc auc score.

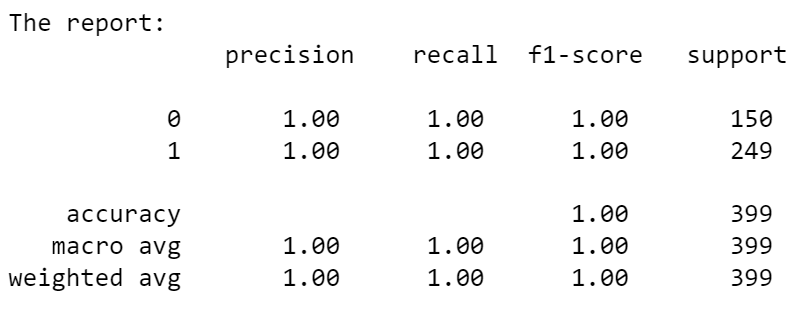
**5. To find the best model by using Classification Report:**

**1. Random Forest:**

****

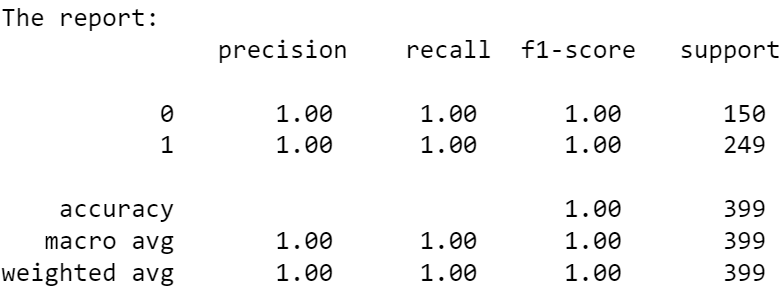


**2. Decision Tree:**

****

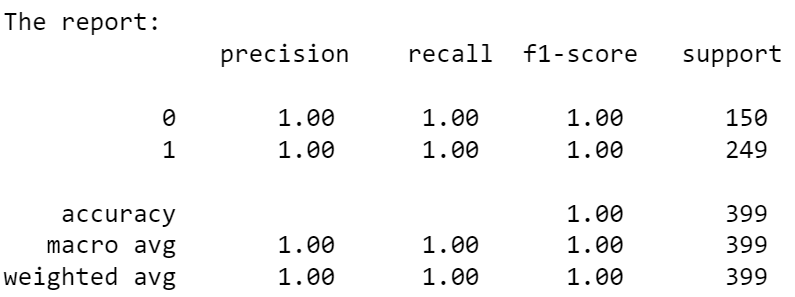
****

**3. Support Vector Machine(SVM):**

****

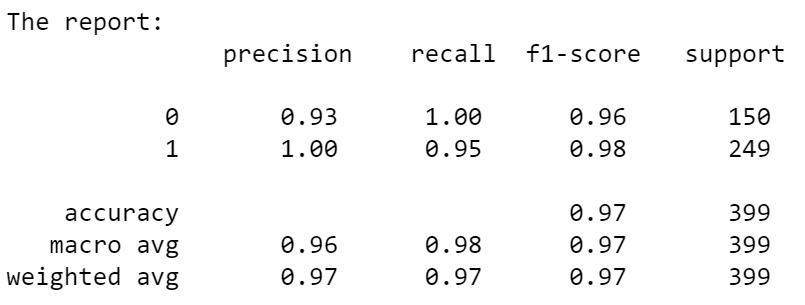
****

**4. Logistic Regression:**

****

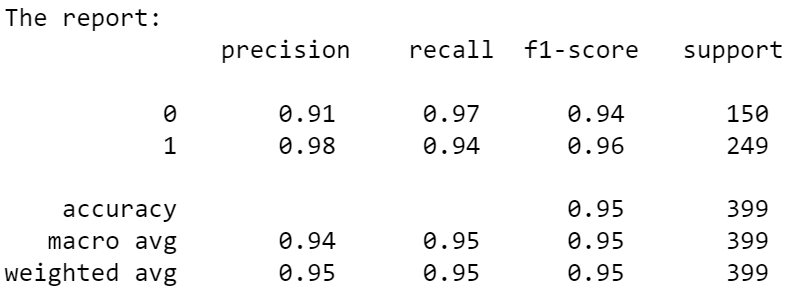
****

**5. KNN Classification:**

****

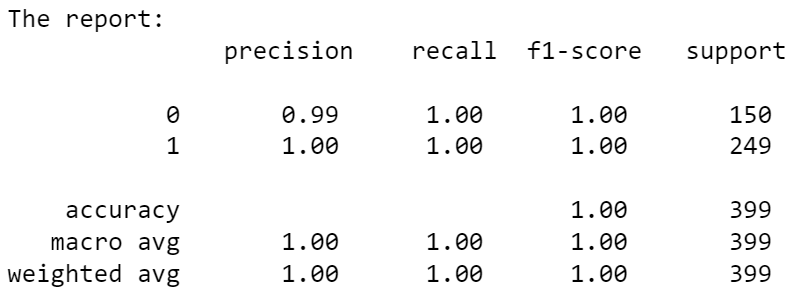
****

**6. Multinomial NB:**

****

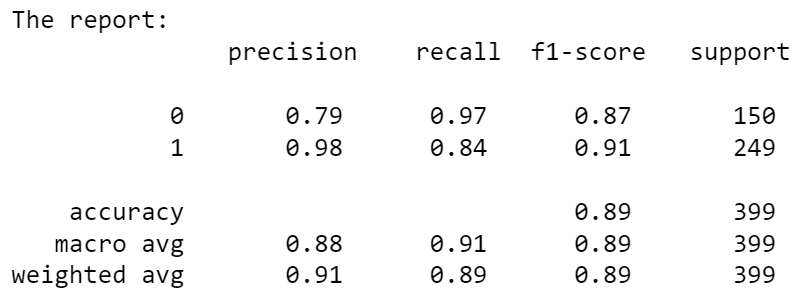
****

**7. Categorical NB:**

****

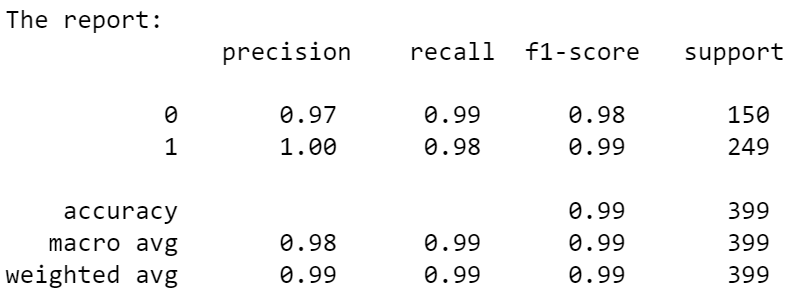
****

**8. Complement NB:**

****

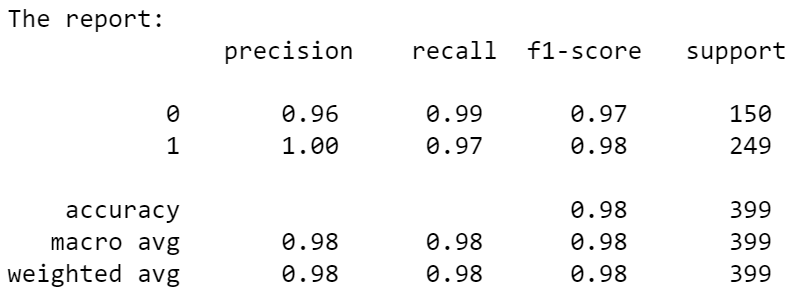
****

**9. Gaussian NB:**

****

****

**10. Bernaulli’s NB:**





**6. Final Model:**

The Random Forest, Decision Tree, SVM, Logistic Regression these are the final model, when compared with other models which is predicted by the classification report and roc auc score.