**Classification Assignment**

1. **Identify your problem statement**

*The Problem statement is to create a predictive*

*model which will predict the Chronic Kidney Disease (CKD) based on the several parameters.*

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

*There are totally 25 parameters. There are 399 rows × 25 columns*

1. **Mention the pre-processing method if you’re doing any (like converting**

**string to number – nominal data)**

*11 columns are pre-processed using get dummies, converting from string to number - nominal data*

**4.) Develop a good model with good evaluation metric. You can use any**

**machine learning algorithm; you can create many models. Finally, you**

**have to come up with final model.**

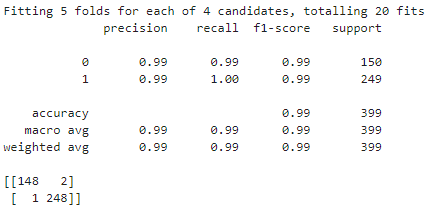
*Among all the classification algorithms, I got the best roc\_auc\_score 0.99 for Logistic GridSearch*

**5.) All the research values of each algorithm should be documented. (You**

**can make tabulation or screenshot of the results.)**

1. ***Logistic***

*Classification Report*

**

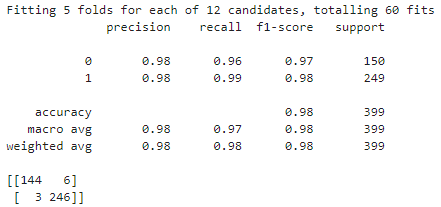
*roc\_auc score*

**

***Highest Accuracy – 0.99***

1. ***Support Vector Machine***

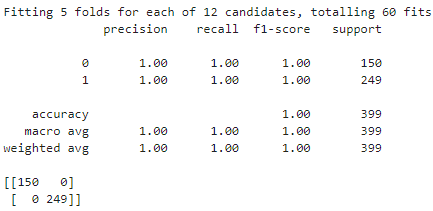
*Classification Report*

**

*roc\_auc score*

**

1. ***Decision Tree***

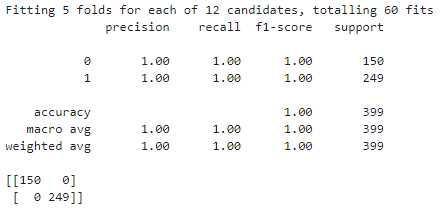
*Classification Report*

*roc\_auc score*

**

1. ***Random Forest***

*Classification Report*

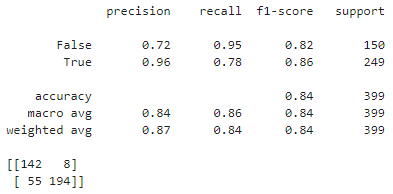
**

*roc\_auc score*

**

1. ***KNN***

*Classification Report*

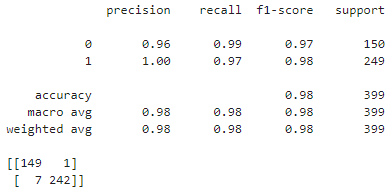
**

*roc\_auc score*

**

1. ***Naive Bayes - Bernaulli's***

*Classification Report*

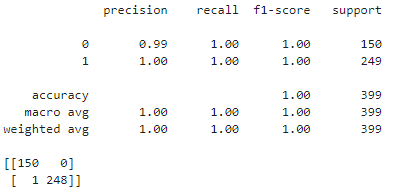
**

*roc\_auc score*

**

1. ***Naive Bayes – Categorical***

*Classification Report*

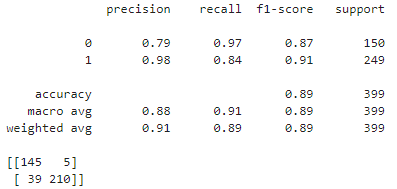
**

*roc\_auc score*

**

1. ***Naive Bayes – Complement***

*Classification Report*

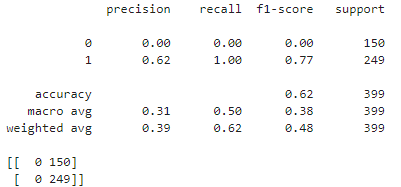
**

*roc\_auc score*

**

1. ***Naive Bayes – Gaussian***

*Classification Report*

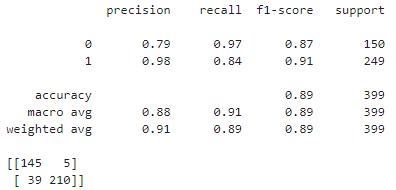
**

*roc\_auc score*

**

***10.* *Naive Bayes - Multinomial***

*Classification Report*

**

*roc\_auc score*

**

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

*Among all the classification algorithms, I got the best roc\_auc\_score 0.99 for Logistic GridSearch*

*So I have chosen Logistic GridSearch as final model and I have done the deployment for that classification algorithm*

*And I have got the highest roc\_auc when I pass these below hyperparameters*

param\_grid = {'solver':['newton-cg', 'lbfgs', 'liblinear', 'saga'],

'penalty':['l2']}