**Classification Assignment**

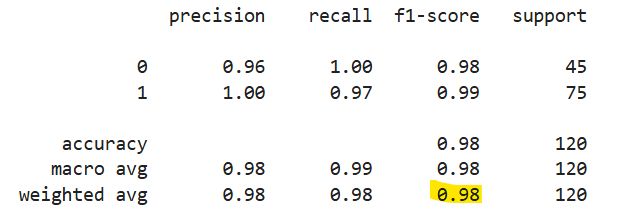
Chronic Kidney Disease prediction based on the given parameters:

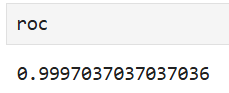
1. The client wants to predict the Chronic Kidney Disease (CKD) by analysing the given parameters.

The three stages involved in identification of the problem statement are

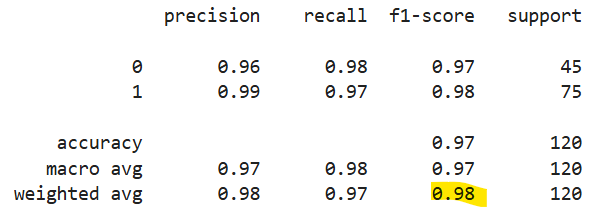
1. Machine Learning as the inputs involve numericals.
2. Supervised Learning as the requirement is clear and the input & output both are available for prediction.
3. Classification, as the output is a categorical data.
4. The given dataset contains 24 input columns and one output column with total train and test sets including 399 row datas.
5. The preprocessing includes 11 columns as they are categorical datas. Accordingly, One hot encoding is being used for preprocessing.
6. Tabulation:

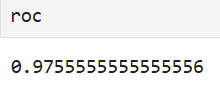
* Support Vector Machine:



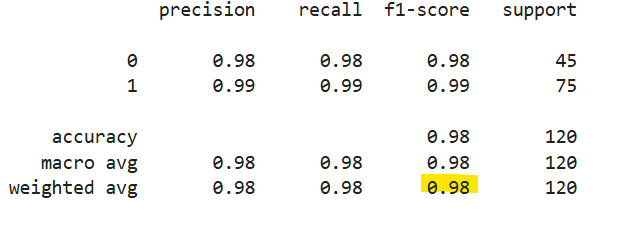


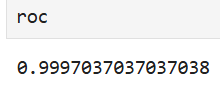
* Decision Tree



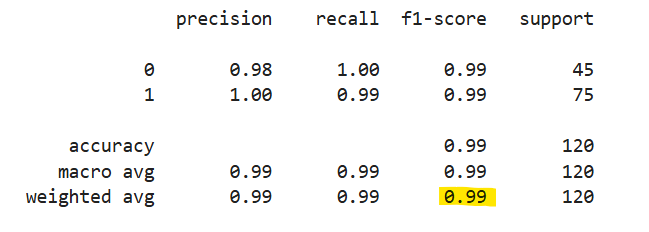


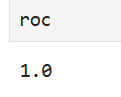
* Random Forest:



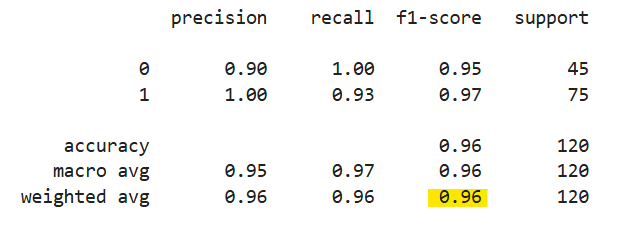


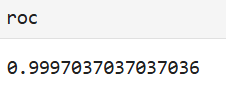
* Linear Regression:





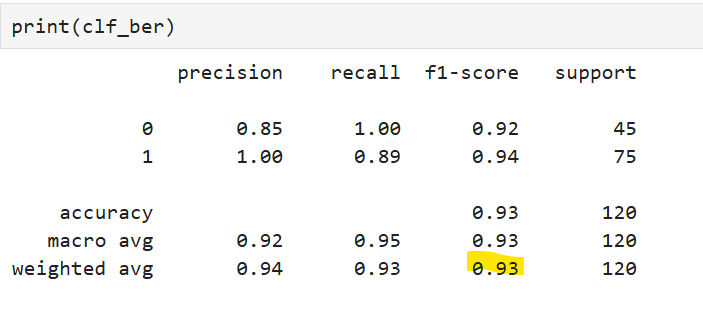
* K Nearest Neighbour:



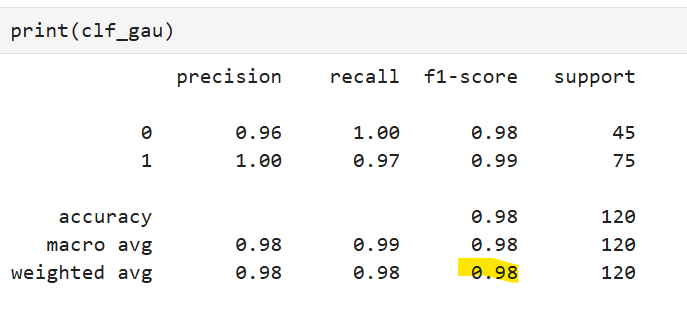


* Naïve Bayes:

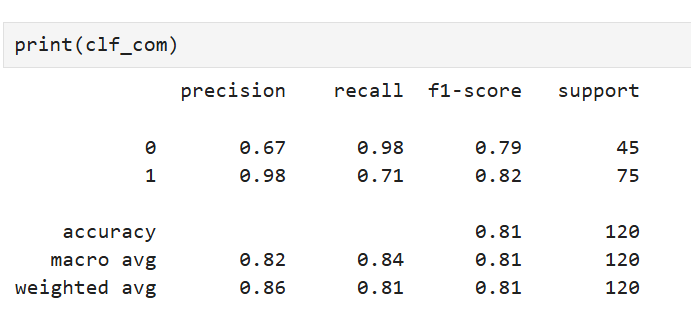
Bernouilli



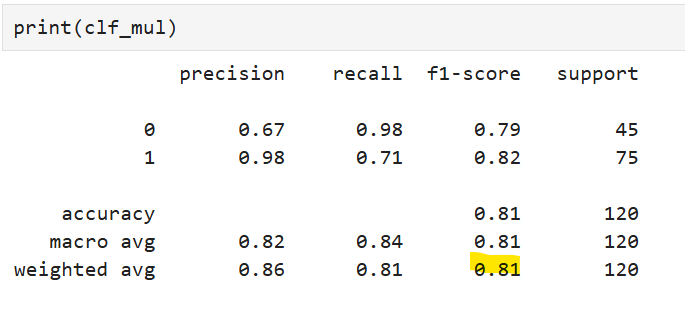
Gaussian

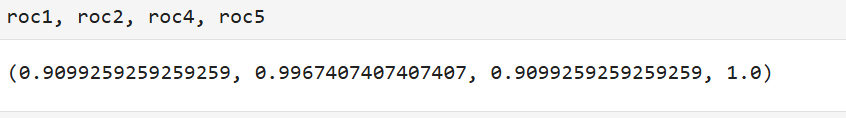


Complement



Multinomial





Roc: [Multinomial, Bernouilli, Complement, Guassian]

Conclusion:

Out of all the algorithms in Machine Learning, the best classification model for this problem statement which is having the best weighted f1 score is : Linear Regression:0.99 and Roc:1.0