**Classification Assignment**

**Problem Statement**

A requirement from the Hospital, Management asked us to create a predictive

model which will predict the Chronic Kidney Disease (CKD) based on the

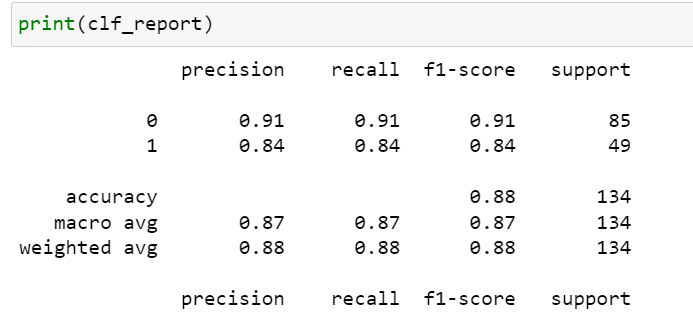
several parameters.

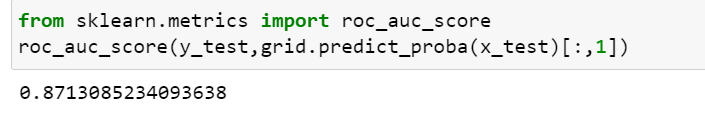
**Recognition Based On The Problem Statement**

* From the given problem statement it identified that to predict the Chronic Kidney Disease from the given problem statement
* For the provided Data set, the Machine Learning Classification can be used to predict the Chronic Kidney Disease.
* Dataset contains 399 rows and 25 columns
* Dataset contains nominal data so it is pre-processed

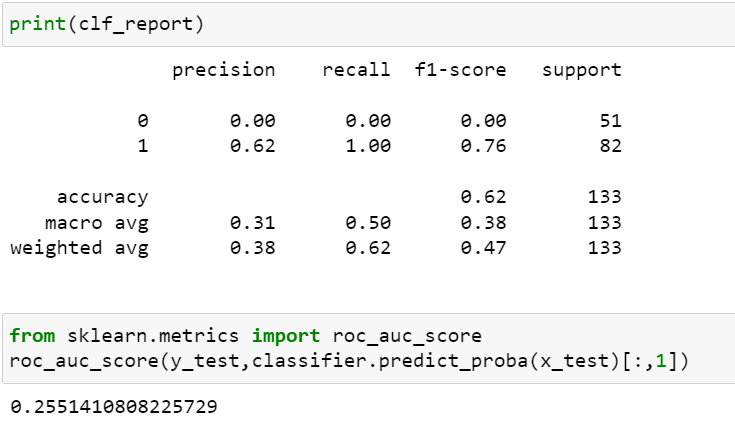
**Machine Learning Models**

**Decision Tree**

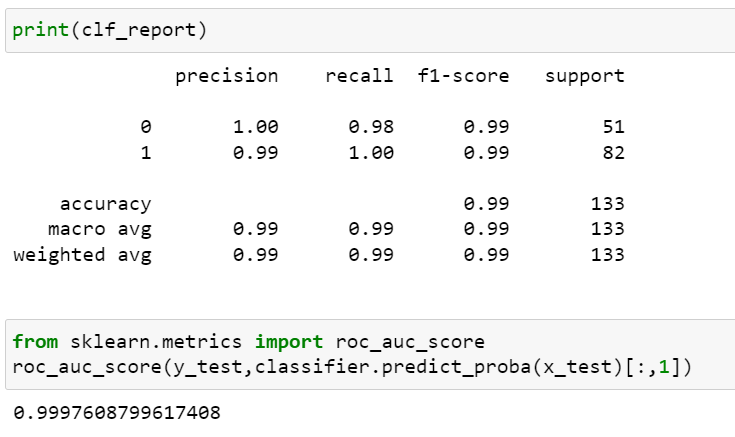




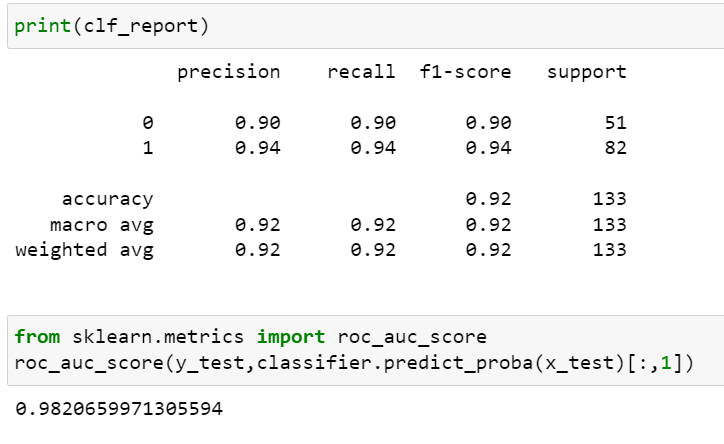
**Support Vector Classification**



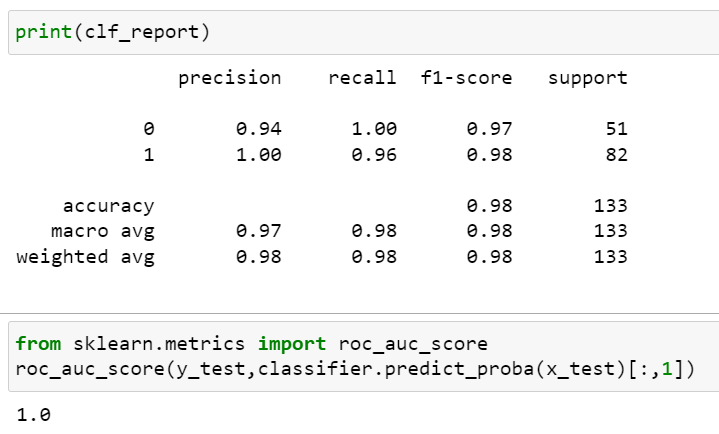
**Random Forest**



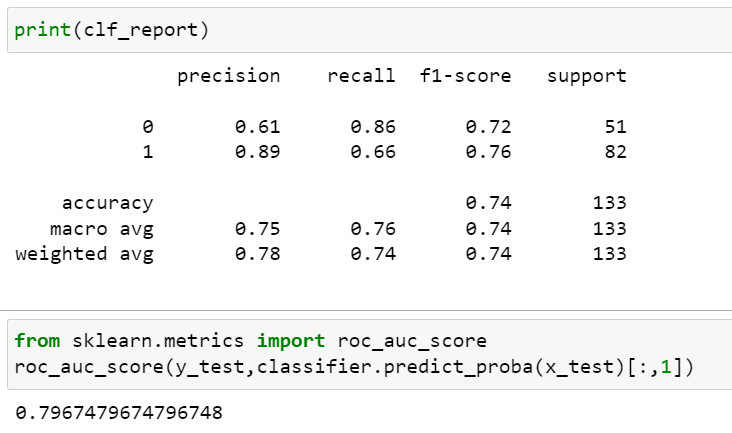
**Logistic Regression**



**Navie’s Baye’s**



**K-nearest Neighbor**



**CONCLUSION**

|  |  |  |
| --- | --- | --- |
| **Model** | **roc\_auc\_score** | **Over all accuracy** |
| **Support vector mechanics** | 0.2551410808225729 | 0.62 |
| **Decision Tree** | 0.929220468675275 | 0.92 |
| **Random Forest** | 0.9997608799617408 | 0.99 |
| **Logistic Regression** | 0.9820659971305594 | 0.92 |
| **K-nearest neighbor** | 0.7967479674796748 | 0.74 |
| **Navie’s Baye’s** | 1.0 | 0.98 |

**Result**

**Random forest machine model has a good level of accuracy and roc\_auc\_score for the provided dataset.**