Hope Artificial Intelligence

Classification Assignment

**Problem Statement or Requirement:**

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. The Client has provided the dataset of the same.

1. Identify your problem statement

Develop a predictive model to identify Chronic Kidney Disease (CKD) based on various parameters provided in the dataset.

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

399 rows × 25 columns

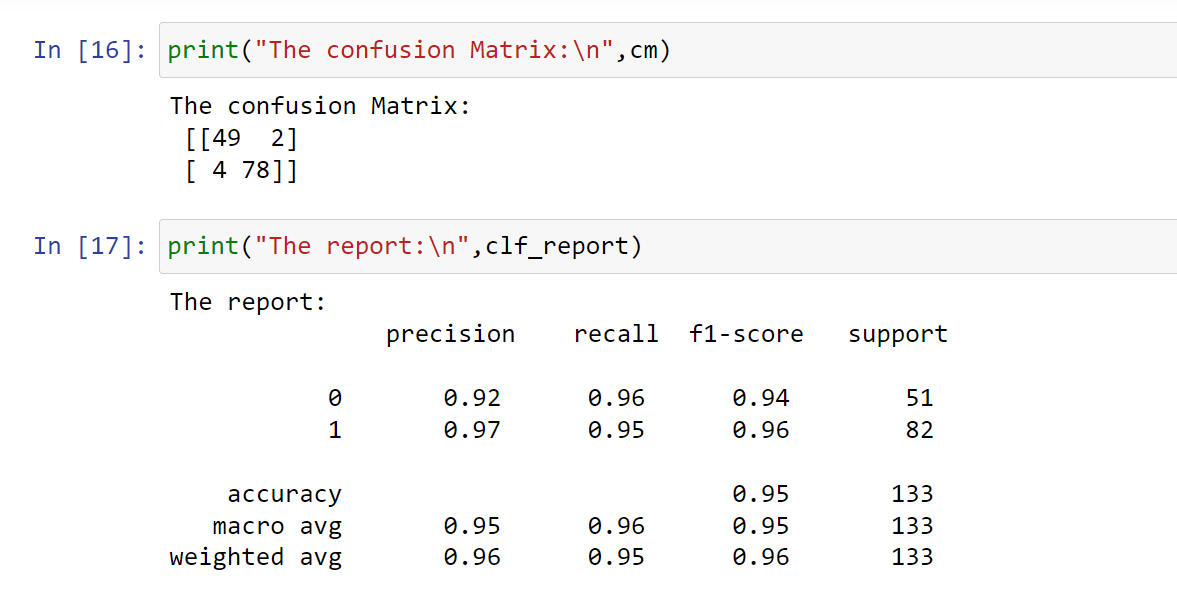
1. Mention the pre-processing method if you’re doing any (like converting string to number – nominal data)

**get\_dummies** -To convert string to number

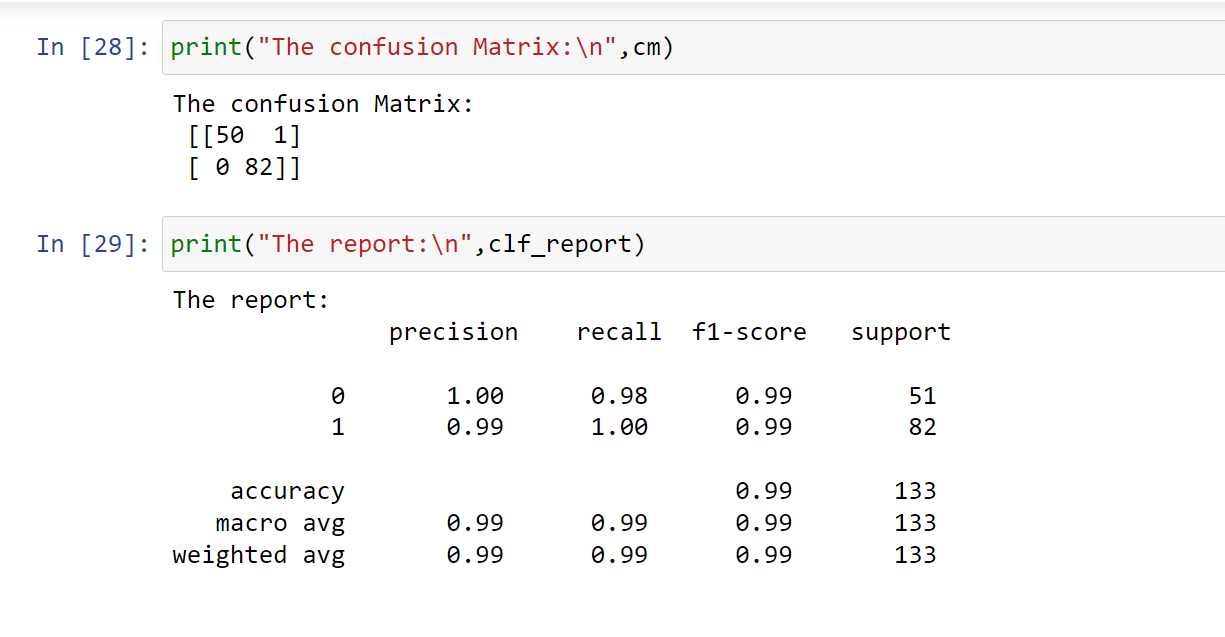
**StandardScaler**-Standardization rescales features to have a mean of 0 and a standard deviation of 1,

1. 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.
2. All the research values of each algorithm should be documented. (You can make tabulation or screenshot of the results.)

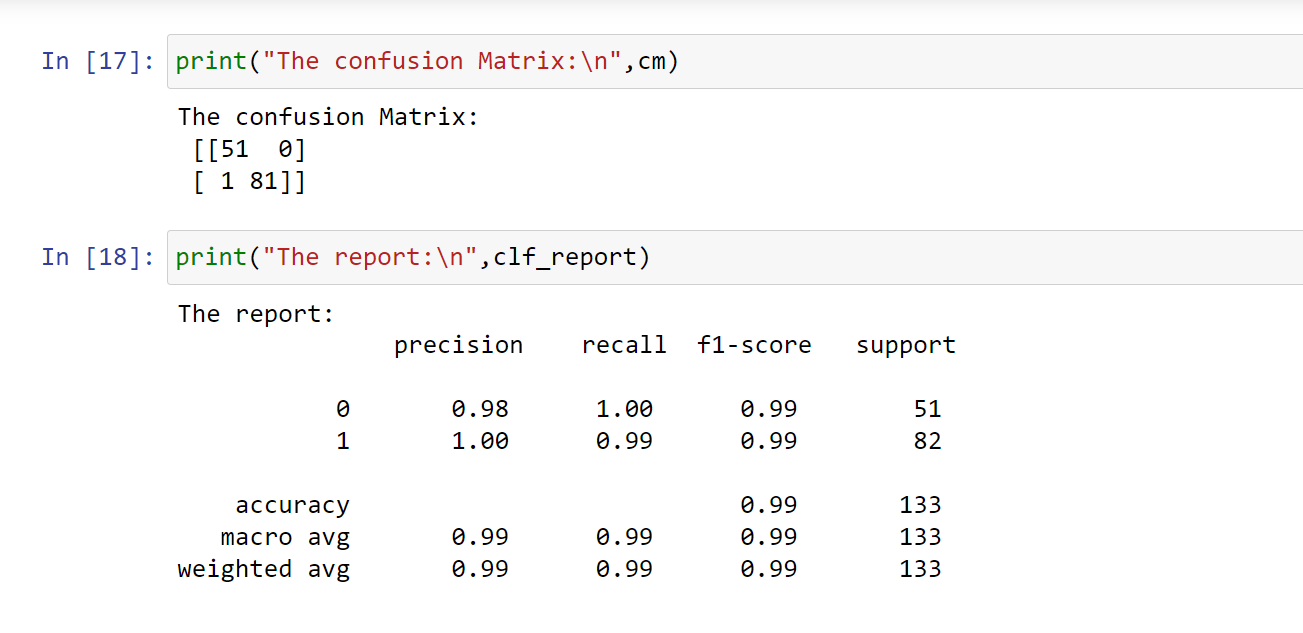
Decision tree results



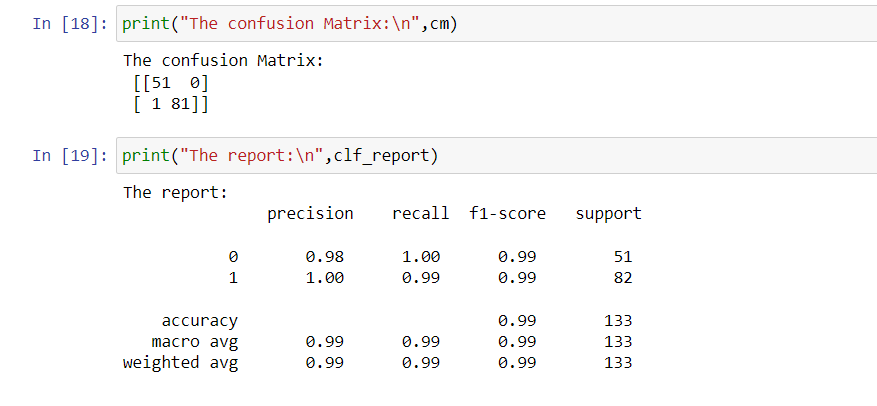
Random Forest Results



SVM Results



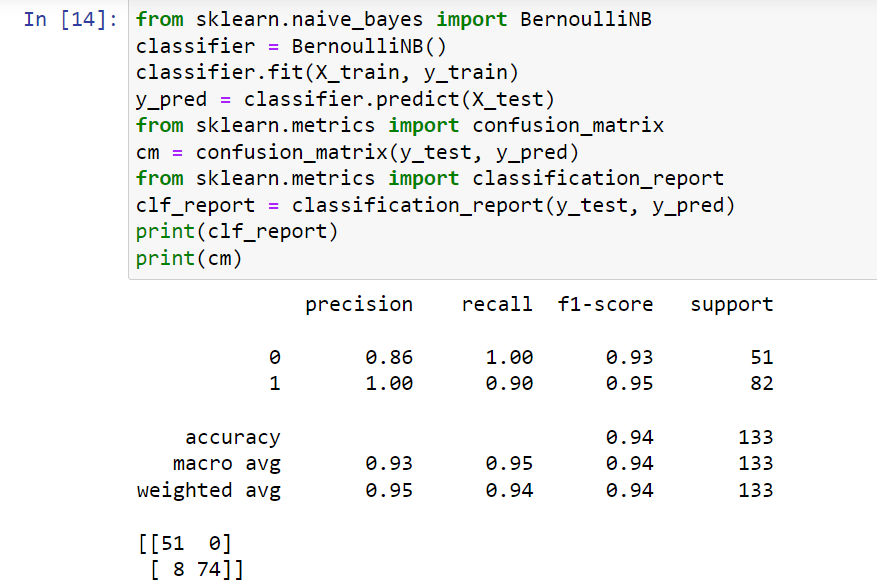
Logistic Regression\_results



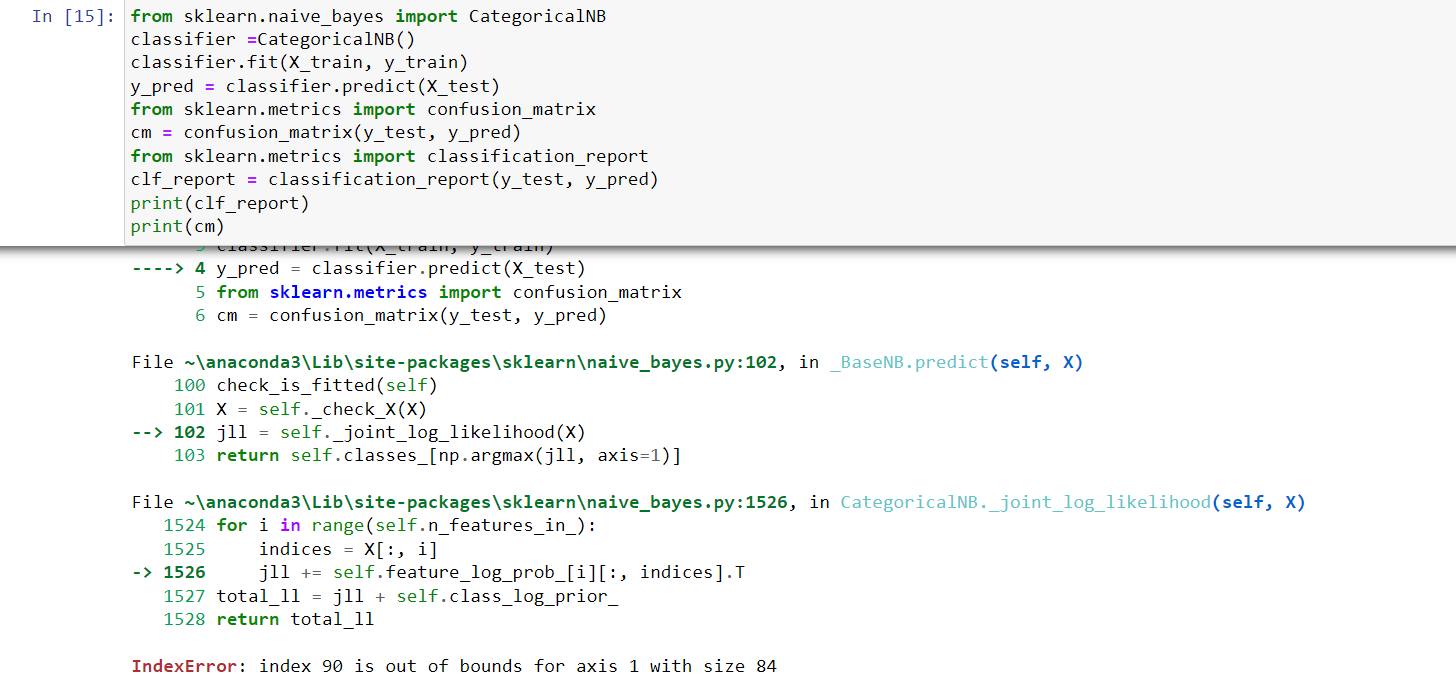
KNN Results

Error

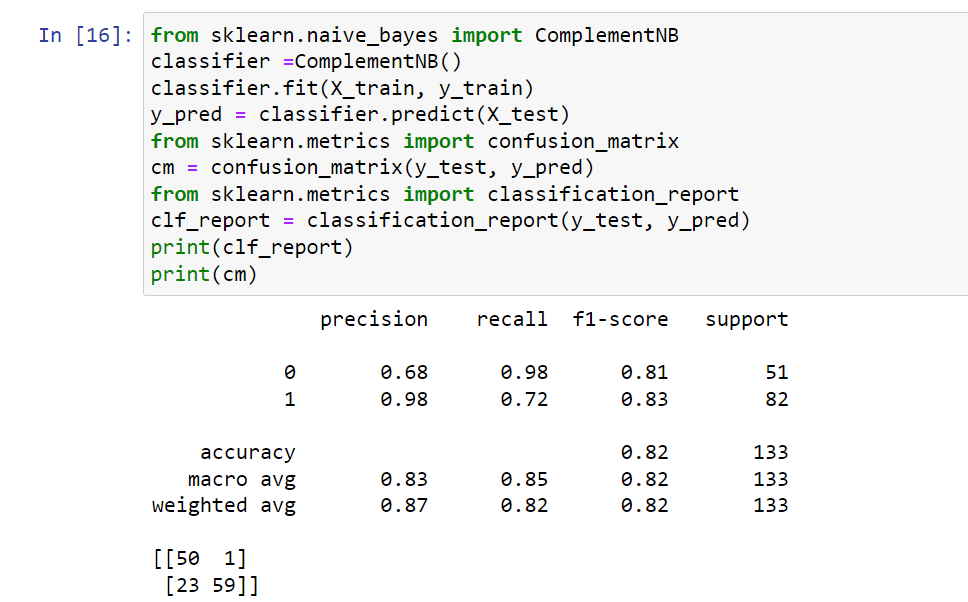
BernoulliNB\_results



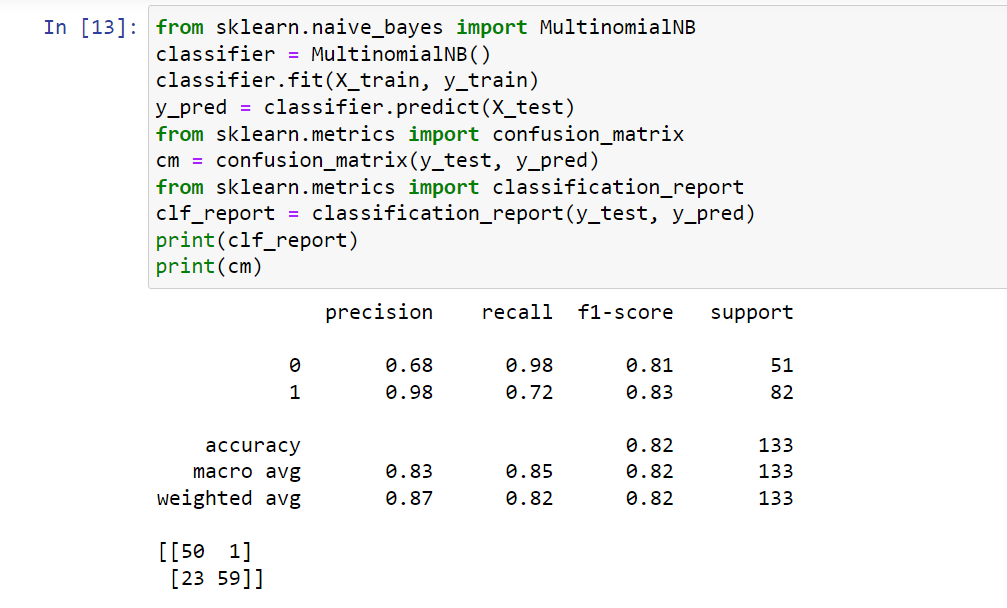
categorical NB\_results



ComplementNB\_Results



MultinomialNB\_results



1. Mention your final model, justify why u have chosen the same.

SVM or Logistic Regression are suitable options due to their low Type I error rates and higher accuracy compared to other models. Additionally, these models demonstrate higher precision and recall values, attributed to their training on unbalanced data.

Best parameter for SVM

The f1\_macro value for best parameter {'C': 10, 'gamma': 'auto', 'kernel': 'sigmoid'}: 0.9924946382275899

Best parameter of Logistic Regression

The f1\_macro value for best parameter {'penalty': 'l2', 'solver': 'newton-cg'}: 0.9924946382275899

Note: Mentioned points are necessary, kindly mail your document as well as .ipynb (code file) with respective name.



* + Sub file name also should be properly named for Example (**SVM\_Ramisha\_Assi-5.ipynb)**

Communication is important (How you are representing the document.)

Kindly uploaded in the Github and Share it with us