Assignment-Classification Algorithm

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.

S.No	Question	Answer
1	Identify your problem statement	Stage1:Machine learning
		Stage2:Supervised learning
		Stage3:Classification
2	Tell basic info about the dataset (Total number of rows, columns)	399 Rows & 25 Columns
3	Mention the pre-processing method if you're doing any (like converting string to number – nominal data)	It is nominal data - because it is not comparable and column expandable
4	Develop a good model with r2_score. You can use any machine learning algorithm; you can create many models. Finally, you have to come up with final model	Good model ,final model attached in Github
5	All the research values (r2_score of the models) should be documented. (You can make tabulation or screenshot of the results.)	Yes,Documented in word
6	Mention your final model, justify why u have chosen the same.	Choosen Random forest In overall classification best model is Random Forest (entropy,log2) – accuracy:0.98 is a best model ,because its having error is less

1) **Logistic CKD-Classification**
   Accuracy:0.93

2) **Support Vector Machine CKD-Classification**
   Accuracy:0.97

3) **Decision Tree CKD-Classification**
   Accuracy:0.94

4) **Random Forest CKD-Classification**
   Accuracy:0.98

5) **KNN CKD-Classification**
   Accuracy:0.95

6) **Navie Baye's CKD-Classification**
   MultinomialNB Accuracy: 0.81
   BernouliNB Accuracy:0.93
   Complement Accuracy:0.81

In overall regression best model is Random Forest (entropy,log2) – accuracy:0.98