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
   We have to predict whether the patient will have Kidney disease.
   Stage1: Machine Learning
   Stage 2: Supervised
   Stage 3: Classification

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

   Column:25

   Row:399

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

   Yes , we do One hot encoding for columns Sg, rbc, pc, pcc, ba, htn, dm, cad, appet, pe, ane, classification.

4. Models

Classifier	F1 score
SVM	0.99
Decision tree	0.96
Random forest	0.99
Logistics	0.99
KNN	0.94
Navie Bayes-Gaussian	0.98
Navie Bayes-Categorical	100
Navie Bayes-Complement	0.82
Navie Bayes-Multinominal	0.82
Navie Bayes-Bernoulli	0.94

5. Model we have chosen.

Navie Bayes Categorical shows 100 % however we haven't splitted the Train_test_split instead we have given direct independent and dependent.