1. Identify your problem statement
    - To develop a ML model that can predict whether a patient has **Chronic Kidney Disease (CKD)** based on various clinical and physiological attributes.
2. Tell basic info about the dataset (Total number of rows, columns)
    - Total columns: 25
    - Total rows: 399
3. Mention the pre-processing method if you're doing any (like converting string to number – nominal data)
    - All categorical (nominal) columns such as rbc, pc, pcc, ba, htn, dm, cad, appet, pe, ane, and classification were converted to numerical format
    - Post conversions total rows are 399 and total columns are 28
4. 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 a final model.

Parameters	SVM	KNN	Gradient Boosting	Random Forest	Logistic Regression
Accuracy	0.9667	0.9417	0.9750	0.9750	0.9750
Precision	0.9853	1.0000	1.0000	0.9855	0.9855
Recall	0.9571	0.9000	0.9571	0.9714	0.9714
F1 Score	0.9710	0.9474	0.9781	0.9784	0.9784
ROC AUC Score	0.9980	0.9836	0.9986	0.9990	0.9989

Random Forest is the best model as most of the scores are consistently above 0.97

5. All the research values of each algorithm should be documented. (You can make tabulation or screenshot of the results.)
    - Refer to the above table
6. Mention your final model, justify why u have chosen the same.
    - Refer to the above justification