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. **Supervised Classification**

2.) Tell basic info about the dataset (Total number of rows, columns). **399 rows × 28 columns**

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

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 final model. **Decision Tree**

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

6.) Mention your final model, justify why u have chosen the same. **Decision tree because it is a good model with 0.99 accuracy.**

5. EVALUATION METRICS O/P:

SVC:

```
[[44  1]
 [ 1 74]]
              precision    recall  f1-score   support

           0       0.98      0.98      0.98        45
           1       0.99      0.99      0.99        75

    accuracy                           0.98       120
   macro avg       0.98      0.98      0.98       120
weighted avg       0.98      0.98      0.98       120
```

Random Forest:

```
[[44  1]
 [ 1 74]]
              precision    recall  f1-score   support

           0       0.98      0.98      0.98        45
           1       0.99      0.99      0.99        75

    accuracy                           0.98       120
   macro avg       0.98      0.98      0.98       120
weighted avg       0.98      0.98      0.98       120
```

Decision Tree:

```
[[45  1]
 [ 0 74]]
              precision    recall  f1-score   support

           0       1.00      0.98      0.99        46
           1       0.99      1.00      0.99        74

    accuracy                           0.99       120
   macro avg       0.99      0.99      0.99       120
weighted avg       0.99      0.99      0.99       120
```

Logistic Regression:

```
[[43  0]
 [ 2 75]]
              precision    recall  f1-score   support

           0       0.96      1.00      0.98        43
           1       1.00      0.97      0.99        77

    accuracy                           0.98       120
   macro avg       0.98      0.99      0.98       120
weighted avg       0.98      0.98      0.98       120
```