CLASSIFICATION ASSIGNMENT

Database: Chronic Kidney Disease

Problem Statement

Stage1: Machine Learning

Stage2: Supervised Learning[Input & output is clear]

Stage3: Classification**[**Output is categorical value**]**

Information About Database

No of Rows:399

No of Columns:28

Preprocessing Method

Converting String to Number :One Hot Encoding Algorithm

[In our database we are converting bg,rbc,pc,pcc,ba,htn,dm,cad,appet pe, & ane columns into number]

Model Creation

1.Decision Tree:

Confusion Matrix

```
[[51  0]
 [ 1 81]]
```

Classification Report

	precision	recall	f1-score	support
0	0.98	1.00	0.99	51
1	1.00	0.99	0.99	82
accuracy			0.99	133
macro avg	0.99	0.99	0.99	133
weighted avg	0.99	0.99	0.99	133

Best Parameter (f1_score)

```
The best parameter{'criterion': 'log_loss', 'max_features': 'sqrt', 'splitter': 'random'}: 0.9924946382
```

Roc_Auc_Score

0.9939024390243902

2.Logistic Regression:

Confusion Matrix

```
[[51  0]
 [ 1 81]]
```

Classification Report

	precision	recall	f1-score	support
0	0.98	1.00	0.99	51
1	1.00	0.99	0.99	82
accuracy			0.99	133
macro avg	0.99	0.99	0.99	133
eighted avg	0.99	0.99	0.99	133

Best Parameter (f1_score)

```
The best parameter{'penalty': 'l2', 'solver': 'sag'}: 0.9924946382275899
```

Roc_Auc_Score

1.0

3.Random Forest:

Confusion Matrix

```
[[50  1]
 [ 1 81]]
```

Classification Report

	precision	recall	f1-score	support
0	0.98	0.98	0.98	51
1	0.99	0.99	0.99	82
accuracy			0.98	133
macro avg	0.98	0.98	0.98	133
weighted avg	0.98	0.98	0.98	133

Best Parameter f1_score)

```
The best parameter{'criterion': 'log_loss', 'max_features': 'sqrt', 'n_estimators': 100}: 0.9849624060
```

Roc_Auc_Score

0.9997608799617408

4.Support Vector Machine:

Confusion Matrix

```
[[51  0]
 [ 1 81]]
```

Classification Report

	precision	recall	f1-score	support
0	0.98	1.00	0.99	51
1	1.00	0.99	0.99	82
accuracy			0.99	133
macro avg	0.99	0.99	0.99	133
weighted avg	0.99	0.99	0.99	133

Best Parameter f1_score)

```
The best parameter{'C': 10, 'gamma': 'scale', 'kernel': 'sigmoid', 'probability': True}: 0.9924946382275899
```

Roc_Auc_Score

1.0

Best model of CKD database is

a) **Logistic Regression**

b) **Support Vector Machine**

Based on Classification Report & Roc_Auc_Score

I Choosed this model.