CLASSIFICATION ASSIGNMENT

PROJECT NAME: CHRONIC KIDNEY DISEASE

1.PROBLEM STATEMENT

MACHINE LEARNING - SUPERVISED LEARNING -CLASSIFICATION

2.DATASET

NUMER OF ROWS: 25

NUMBER OF COLUMNS:400

3.PREPROCESSING METHOD

THE CATEGORICAL VALUE IS CONVERTED BY NUMERICAL VALUE BY USING NOMINAL DATA.THEN USING get_dummies IN PANDAS LIBRARY. INPUT/OUTPUT SPLITATION PROCESS

TEST/TRAINING SET CREATED.

4.MODEL

RANDOM FOREST CLASSIFICATION

```
print(cm)
```

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

```
from sklearn.metrics import classification_report
clf_report=classification_report(Y_test,y_pred)
```

```
print(clf_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
```

SVM CLASSIFICATION

```
print(cm)
```

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

```
from sklearn.metrics import classification_report
clf_report=classification_report(Y_test,y_pred)
```

```
print(clf_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
```

DECISION TREE CLASSIFICATION

```
print(cm)
```

interrupt the kernel

```
[[49  2]
 [ 8 74]]
```

```
from sklearn.metrics import classification_report
clf_report=classification_report(Y_test,y_pred)
```

```
print(clf_report)
```

```
              precision    recall  f1-score   support

           0       0.86      0.96      0.91        51
           1       0.97      0.90      0.94        82

    accuracy                           0.92       133
   macro avg       0.92      0.93      0.92       133
weighted avg       0.93      0.92      0.93       133
```

LOGISTIC TREE CLASSIFICATION

```
print(cm)
```

```
[[51  0]
 [ 2 80]]
```

```
from sklearn.metrics import classification_report
clf_report=classification_report(Y_test,y_pred)
print(clf_report)
```

```
              precision    recall  f1-score   support

           0       0.96      1.00      0.98        51
           1       1.00      0.98      0.99        82

    accuracy                           0.98       133
   macro avg       0.98      0.99      0.98       133
weighted avg       0.99      0.98      0.99       133
```

SUPPORT VECTOR MACHINE CLASSIFICATION GIVES BETTER MODEL:0.99