# Prediction of Chronic Kidney Disease (CKD)

## 1. Proplem Statement

To predict Chronic Kidney Disease (CKD) with given dataset.

The input was cleard in number format and some inputs are in categorical. The output is also in categorical value that's means Classification

**Stage1 :** Machine Learning - ( input-number )
**Stage2 :** Supervised Learning - ( input output is cleard )
**Stage3 :** Classification - ( output is numerical value )

## 2. Basic info about dataset

- Total number of Rows 339.
- Total number of Coulums 25.
- Total number of Inputs 24. ( age, bp, sg, al, su,etc… )
- Total number of Outputs 1. ( Classification )

## 3. Pre-processing

The categorical value inputs are convert to ordinal data.(using *get_dummies* method )

## 4. Model Creaction

**Support Vector Machine ( SVM )**

```
from sklearn.metrics import confusion_matrix
cm=confusion_matrix(Y_test,grid_predictions)
print(cm)
```

```
[[45  0]
 [ 2 73]]
```

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

```
print(clf_report)
```

```
              precision    recall  f1-score   support

       False       0.96      1.00      0.98        45
        True       1.00      0.97      0.99        75

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

**Accuracy - 0.98**

**LogisiticRegression**

```
print(cm)
```

```
[[45  0]
 [ 1 74]]
```

```
print(clf_report)
```

```
              precision    recall  f1-score   support

       False       0.98      1.00      0.99        45
        True       1.00      0.99      0.99        75

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

**Accuracy - 0.99**

LogisiticRegression is given the best accuracy so we go to deployment phase for **LogisiticRegression.**