Chronic kidney disease-Prediction

As a data scientist, the goal is to develop a predictive model that uses these parameters to accurately predict the Chronic kidney disease

Data set contain **399 rows × 28 columns** data.
It have categorical data in multiple column so we need to convert it into **INT** by using **get_dummies()** function from pandas.

With the use of GridSearchCV both Decision tree and Random forest predict with good accuracy.

SVM Sreenshot:

```
[[45  2]
 [ 1 85]]
              precision    recall  f1-score   support

           0       0.98      0.96      0.97        47
           1       0.98      0.99      0.98        86

    accuracy                           0.98       133
   macro avg       0.98      0.97      0.98       133
weighted avg       0.98      0.98      0.98       133
```

```
from sklearn.metrics import f1_score
fs=f1_score(y_test,y_pred,average='weighted')
print("The f1_macro value for the best parameters {}:".format(grid.best_params_),fs)

from sklearn.metrics import roc_auc_score
roc_auc_score=(y_test,grid.predict_proba(x_test)[:1])
```

```
The f1_macro value for the best parameters {'C': 10, 'gamma': 'auto', 'kernel': 'linear'}: 0.9773875299847604
```

Decision tree Screenshot:

```
[[47  0]
 [ 0 86]]
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        47
           1       1.00      1.00      1.00        86

    accuracy                           1.00       133
   macro avg       1.00      1.00      1.00       133
weighted avg       1.00      1.00      1.00       133

The f1_macro value for the best parameters {'criterion': 'gini', 'max_depth': 10, 'max_features': 'sqrt', 'min_samples_split': 2}: 1.0
```

Random Forest Screenshot:

```
[[47  0]
 [ 0 86]]
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        47
           1       1.00      1.00      1.00        86

    accuracy                           1.00       133
   macro avg       1.00      1.00      1.00       133
weighted avg       1.00      1.00      1.00       133

The f1_macro value for the best parameters {'criterion': 'log_loss', 'max_features': 'log2', 'n_estimators': 100}: 1.0
```

Both the **DecisionTree** and **RandomForest** predict Chronic kidney disease with **100** percent accuracy

DecisionTree :`{'criterion': 'gini', 'max_depth': 10, 'max_features': 'sqrt', 'min_samples_split': 2}`

RandomForest :`{'criterion': 'log_loss', 'max_features': 'log2', 'n_estimators': 100}`