# 1) Identify your problem statement :

# 1.Name of the Project:

**Chronic Kidney Disease Prediction**

# Stage 1:

**Domain:** Machine Learning

# Stage 2:

**Learning:** Supervised Learning

# Stage 3:

Classification

# 2.) Tell basic info about the dataset (Total number of rows, columns) :

# Total number of rows and columns = 399 rows × 25 columns

# Input Columns=24

# Output Columns =1

# 1 – Classification\_yes = 249

# 0 \_ Classification\_no = 150

# 3.) Mention the pre-processing method if you’re doing any (like converting string to number – nominal data) :

# 1. converting string to number – nominal data

# 4.) Develop a good model with good evaluation metric :

**1.Decision Tree :**

**1.F1\_Macro :**

f1\_macro value for best parameter :

{'criterion': 'gini', 'max\_features': 'sqrt', 'splitter': 'random'}

F1\_macro : **0.9668037602820211**

# 

# 2.Clf\_report:

# 

# 

# 3.ROC-AUC :

# Roc\_auc\_score=0.9688888888888889

**2.Random Forest :**

**1.F1\_Macro :**

f1\_macro values for the best parameter:

{'class\_weight': 'balanced', 'criterion': 'gini', 'max\_features': 'log2', 'n\_estimators': 100}

F1 macro = 0.9833333333333335

# 2.Clf\_report:

# 

# 3.ROC\_AUC :

**Roc\_auc\_score=** 0.9997037037037036

**3.SVM :**

**1.F1\_Macro :**

f1\_macro values for the best parameter:

'C': 10, 'gamma': 'auto', 'kernel': 'sigmoid'}

F1 macro = 0.9834018801410106

# 2.Clf\_report:

# 

# 

# 3.ROC\_AUC :

**Roc\_auc\_score=** 0.9997037037037036

**4.Logistic :**

**1.F1\_Macro :**

f1\_macro values for the best parameter:

{'penalty': 'l2', 'random\_state': 0, 'solver': 'liblinear'}

F1 macro = 0.9832535885167464

# 2.Clf\_report:

# 

# 3.ROC\_AUC :

**Roc\_auc\_score=** 0.9973333333333334

**5.KNN :**

**1.F1\_Macro :**

f1\_macro values for the best parameter:

{'algorithm':'auto','metric':'minkowski','n\_neighbors': 10,'p': 2,'weights':'distance'}

F1 macro = 0.9505208333333334

# 2.Clf\_report:

# 

# 3.ROC\_AUC :

**Roc\_auc\_score=** 0.9997037037037038

**6.Naive Bayes :**

**1.GaussianNB:**

**1.F1\_Macro :**

F1 macro values for the best parameter GaussianNB() = 0.9834018801410106

# 2.Clf\_report:

# 

# 3.ROC\_AUC :

**Roc\_auc\_score = 1.0**

**2.MultinomialNB:**

**1.F1\_Macro :**

F1 macro values for the best parameter MultinomialNB() = 0.8108631887701655

# 2.Clf\_report:

# 

# 3.ROC\_AUC :

**Roc\_auc\_score =** 0.9099259259259259

**3.BernoulliNB:**

**1.F1\_Macro :**

F1 macro values for the best parameter BernoulliNB() = 0.934176487496407

# 2.Clf\_report:

# 

# 3.ROC\_AUC :

**Roc\_auc\_score =** 0.9967407407407407

**3.BernoulliNB:**

**1.F1\_Macro :**

F1 macro values for the best parameter BernoulliNB() = 0.934176487496407

# 2.Clf\_report:

# 

# 3.ROC\_AUC :

**Roc\_auc\_score =** 0.9967407407407407

**7.Ordinary Least Square :**

**1.F1\_Macro :**

F1 macro values for the best parameter LinearRegression(): 1.0

# 2.Clf\_report:

# 

# 3.ROC\_AUC :

# ROC\_AUC for OLS Model: 0.9946666666666666

**8.** **StochasticGradientDescent:**

**1.F1\_Macro :**

f1\_macro values for the best parameter:

{'loss': 'hinge', 'penalty': 'elasticnet'}

F1 macro = 0.9916844900066377

# 2.Clf\_report:

# 

# 

# 3.ROC\_AUC :

# ROC\_AUC for StochasticGD Model: 0.9733333333333334

**9.MultiTaskLasso:**

**1.F1\_Macro :**

f1\_macro values for the best parameter:

{alpha=0.1,max\_iter=1000, tol=0.0001,random\_state=None,selection='random'}

F1 macro = 0.96

# 2.Clf\_report:

# 

# 

**3.ROC\_AUC :**

ROC\_AUC for MultiTaskLasso Model: 0.9961481481481482

**10.QuantileRegressor:**

**1.F1\_Macro :**

f1\_macro values for the best parameter:

{alpha=0.1,max\_iter=1000, tol=0.0001,random\_state=None,selection='random'}

F1 macro = 0.87

# 2.Clf\_report:

# 

# 

# 3.ROC\_AUC :

ROC\_AUC for QuantileRegressor Model: 0.9703703703703703

**11.KernelRidge:**

**1.F1\_Macro :**

f1\_macro values for the best parameter:

{alpha=1, kernel='linear', gamma=None, degree=3, coef0=1, kernel\_params=None}

F1 macro = 0.93

# 2.Clf\_report:

# 

# 

# 3.ROC\_AUC :

**Roc\_auc\_score=** 0.6974814814814816

**12.RidgeClassifier:**

**1.F1\_Macro :**

f1\_macro values for the best parameter:

{'alpha': 1.0, 'random\_state': 0, 'solver': 'auto', 'tol': 0.0001}

F1 macro = 0.934176487496407

# 2.Clf\_report:

# 

# 

# 3.ROC\_AUC :

**Roc\_auc\_score=** 0.9952592592592593

**13.Lasso:**

**1.F1\_Macro :**

f1\_macro values for the best parameter:

{alpha=0.1,max\_iter=500, tol=0.0001,random\_state=None,selection='cyclic'}

F1 macro = 0.96

# 2.Clf\_report:

# 

# 

# 3.ROC\_AUC :

# ROC\_AUC for Lasso Model: 0.9961481481481482

**14.** **LinearDiscriminantAnalysis:**

**1.F1\_Macro :**

f1\_macro values for the best parameter:

{'solver': 'svd', 'tol': 0.0001}:

F1 macro = 0.934176487496407

# 2.Clf\_report:

# 

# 

# 3.ROC\_AUC :

**Roc\_auc\_score=**0.9946666666666666

**15.** **Perceptron:**

**1.F1\_Macro :**

f1\_macro values for the best parameter:

{'solver': 'svd', 'tol': 0.0001}:

F1 macro = 0.93

# 2.Clf\_report:

# 

# 

# 3.ROC\_AUC :

**Roc\_auc\_score=**0.9988148148148148

**6.) Mention your final model, justify why u have chosen the same.**

I have chosen final Model as **‘Ordinary Least Square’** in Linear model which has Accuracy, Precision,Recall and ROC\_AUC has value as equal to **‘1’,so which condider as Good Model**