# Classification Assignment

**Problem Statement or Requirement:-**

A requirement from the Hospital, Management asked us to create a predictive model which will predict the Chronic Kidney Disease (CKD) based on the several parameters. The Client has provided the dataset of the same.

1.) Identify your problem statement

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

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

4.) Develop a good model with good evaluation metric. You can use any machine learning algorithm; you can create many models. Finally, you have to come up with final model.

5.) All the research values of each algorithm should be documented. (You can make tabulation or screenshot of the results.

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

## 1.) Identify your problem statement :-

**Stage-1- Machine Learning**
**Why its ML?**
**In this problem statement having numberical data then I use Ml**
**Stage-2- ML under Supervised Learning**
**Why its Sl?**
**In this problem statement cleared input and output then I decide to Supervised learning**
**Stage-3-SL- under Classification**
**Why its Classification?**

**In this problem statement having categorical data then I use Classification.**

**2.) Tell basic info about the dataset (Total number of rows, columns) - 399 Rows and 25 Colunms**

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

**dataset=pd.get_dummies(dataset,drop_first=True)**

In above line using pandas library to convert categorical data give dummies using One hot encoding algorithms and then use drop function drop the one colunm from the table **.**

**4.) Develop a good model with good evaluation metric. You can use any machine learning algorithm; you can create many models. Finally, you have to come up with final model. And**

**5.) All the research values of each algorithm should be documented. (You can make tabulation or screenshot of the results**.

**1.DecisonTree with Classification:-Report:**

```
the confusion matrix is
 [[50  1]
 [ 1 81]]
the classification_Report is
              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
```

**Accuracy Score=0.98**

**1.F1Score:-**

```
The f1_macro value for best parameter {'criterion': 'gini', 'max_features': 'auto', 'splitter': 'random'}: 0.9849624060150376
```

**F1Score=0.98**

**2.Roc_auc_score:-**

**The roc_auc_score= 0.98**

**2.SupportVectorMachine with Classification:-**

```
the confusion metrix is
 [[51  0]
 [ 1 81]]
the classification report is
              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
```

**Accuracy Value=0.99**

**2.F1_Score:-**

```
The f1_macro value for best parameter {'C': 10.0, 'decision_function_shape': 'ovo', 'gamma': 'auto', 'kernel': 'sigmoid'}: 0.99
24946382275899
```

**F1_Score=0.99**

**3.Roc_auc_score:-**

**The roc_auc_score= 1.0**

## 3.Randomfiorest_Classification:-

```
the confusion metrix is
 [[51  0]
 [ 1 81]]
the classification report is
              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
```

## Accuracy_value=0.99

## 2.F1_Score:-

```
The f1_macro value for best parameter {'class_weight': 'balanced', 'criterion': 'entropy', 'max_features': 'log2', 'n_estimator
s': 100}: 0.9924946382275899
```

## F1_Score=0.99

## 3.Roc_auc_score:-

**The roc_auc_score= 1.0**

## 4.Logistic_Classification:-

```
the confusion metrix is
 [[51  0]
 [ 1 81]]
the classification report is
              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
```

**Accuracy=0.99**

**2.F1_Score:-**

```
The f1_macro value for best parameter {'multi_class': 'multinomial', 'penalty': 'l2', 'solver': 'newton-cg'}: 0.9924946382275899
```

**F1_Score=0.99**

**3.Roc_auc_score:-**

**The roc_auc_score= 1.0**

## 5.K-Nearest_Neighbors_Classification:-

```
the confusion metrix is
 [[51  0]
 [ 8 74]]
the classification report is
              precision    recall  f1-score   support

           0       0.86      1.00      0.93        51
           1       1.00      0.90      0.95        82

    accuracy                           0.94       133
   macro avg       0.93      0.95      0.94       133
weighted avg       0.95      0.94      0.94       133
```

**Accuracy value=0.94**

**2.F1_Score:-**

```
The f1_macro value for best parameter {'algorithm': 'auto', 'metric': 'minkowski', 'n_neighbors': 7, 'p': 2, 'weights': 'uniform'}: 0.9404945931261721
```

**F1_Score=0.94**

**3.Roc_auc_score:-**

**The roc_auc_score= 0.99**

## 6.Navies_Bayes_Classification:-

### 1.Naive_bayes multinomiaLNB:-

```
the confusion metrix is
 [[50  1]
 [23 59]]
the classification report is
              precision    recall  f1-score   support

           0       0.68      0.98      0.81        51
           1       0.98      0.72      0.83        82

    accuracy                           0.82       133
   macro avg       0.83      0.85      0.82       133
weighted avg       0.87      0.82      0.82       133
```

**Accuracy=0.82**

### 2.F1_Score:-

```
The f1_macro value for best parameter {'alpha': 1.0, 'class_prior': None, 'fit_prior': True}: 0.8215780250262184
```

**F1_score=0.82**

### 3.Roc_auc_score:-

**The roc_auc_score= 0.91**

## 2. Bernoulli_Navies_Bayes_Classification:-

```
the confusion metrix is
 [[51  0]
 [ 8 74]]
the classification report is
              precision    recall  f1-score   support

           0       0.86      1.00      0.93        51
           1       1.00      0.90      0.95        82

    accuracy                           0.94       133
   macro avg       0.93      0.95      0.94       133
weighted avg       0.95      0.94      0.94       133
```

**Accuracy=0.94**

**2.F1_Score:-**

```
The f1_macro value for best parameter {'alpha': 1.0, 'class_prior': None, 'fit_prior': 'True'}: 0.9404945931261721
```

**F1_score=0.94**

**3.Roc_auc_score:-**

**The roc_auc_score= 0.99**

**3.Categorical_navies_bayes_Classification:-**

-

**4.Complement_Naviesbayes_Classification:-**

```
the confusion metrix is
 [[50  1]
 [23 59]]
the classification report is
              precision    recall  f1-score   support

           0       0.68      0.98      0.81        51
           1       0.98      0.72      0.83        82

    accuracy                           0.82       133
   macro avg       0.83      0.85      0.82       133
weighted avg       0.87      0.82      0.82       133
```

**Accuracy=0.82**

**1.F1_Score:-**

```
The f1_macro value for best parameter {'alpha': 1.0, 'class_prior': None, 'fit_prior': 'True'}: 0.8215780250262184
```

**F1_score=0.82**

**2. Roc_auc_score:-**

**roc_auc_score=0.91**

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

**In comparing to all of the algorithm are build a model but accuracy and f1_score and roc_auc_score is better in Support Vector Machine Classification ,**

- **Accuracy_value=0.99**
- **F1_Score=0.99**
- **Roc_auc_score=1.0**

**In above all the values are near to 1. Then this model is best model.**