

CLASSIFICATION ASSIGNMENT

[CLICK HERE TO SEE THE DATASET](#)

PROBLEM IDENTIFICATION

STAGE 1:

DOMAIN SELECTION : Machine Learning

STAGE 2:

LEARNING SELECTION : Supervised Learning

STAGE 3:

CATEGORICAL SELECTION : Classification

BASIC INFORMATION ABOUT DATASET

PROJECT NAME : Chronic Kidney Disease Prediction

TOTAL NUMBER OF COLUMNS : 25

TOTAL NUMBER OF ROWS : 399

NOMINAL COLUMN : sg, rbc, pc, pcc, ba, htn, dm, cad, pe, ane, appet

SVM CLASSIFICATION REPORT

	Precision	recall	F1_score	support
0	0.94	1.00	0.97	45
1	1.00	0.96	0.98	75
Accuracy			0.97	120
macro avg	0.97	0.98	0.97	120
Weighted avg	0.98	0.97	0.98	120

roc_auc_score: 0.99

DECISION TREE CLASSIFICATION REPORT

	Precision	recall	F1_score	support
0	0.94	0.98	0.96	45
1	0.99	0.96	0.97	75
accuracy			0.97	120
macro avg	0.96	0.97	0.96	120
Weighted avg	0.97	0.97	0.97	120

roc_auc_score: 0.96

RANDOM FOREST CLASSIFICATION REPORT

	Precision	recall	F1_score	support
0	0.98	1.00	0.99	45
1	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

roc_auc_score: 0.99

LOGISTIC_GRID CLASSIFICATION REPORT

	Precision	recall	F1_score	support
0	0.98	1.00	0.99	45
1	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

roc_auc_score: 1.0

KNN CLASSIFICATION REPORT

	Precision	recall	F1_score	support
0	0.98	1.00	0.94	45
1	1.00	0.92	0.96	75
accuracy			0.95	120
macro avg	0.94	0.96	0.95	120
Weighted avg	0.96	0.95	0.95	120

roc_auc_score: 0.99

NAÏVE_BAYES CLASSIFICATION REPORT

	Precision	recall	F1_score	support
0	0.96	1.00	0.98	45
1	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

roc_auc_score: 1.0

BEST MODEL: LOGISTIC_GRID ALGORITHM

REASON: Logistic algorithm gives the best model because the f1 score and roc_auc_score is very high to compare with other algorithm

ACCURACY: 0.99

ROC_AUC_SCORE:1.0