CLASSIFICATION ASSIGNMENT

CHRONIC KIDNEY DISEASE PREDICTION

MUGILAN A

23 AUGUST 2024

Problem Statemnt

Hospital management needs to predict chronic kidney disease (CKD) using several data collected from past.

Dataset

By Analyzing the dataset, we came to know that it has total record of 399 patients and various parameters.

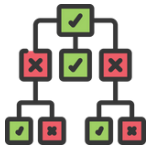
Machine Learning

↓

Supervised

↓

Classification



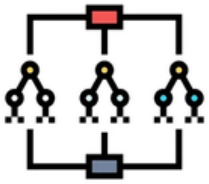
Decision Tree Classifier

Best Hyper Tuning Parameters

criterion : entropy
max_features : sqrt
splitter : random

Classification Report - Decision Tree

	precision	recall	f1-score	support
0	0.88	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



Random Forest Classifier

Best Hyper Tuning Parameters

criterion : entropy
max_features : log2
n_estimators : 100

Classification Report - Random Forest

	precision	recall	f1-score	support
0	0.96	0.98	0.97	45
1	0.99	0.97	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

Support Vector Classifier

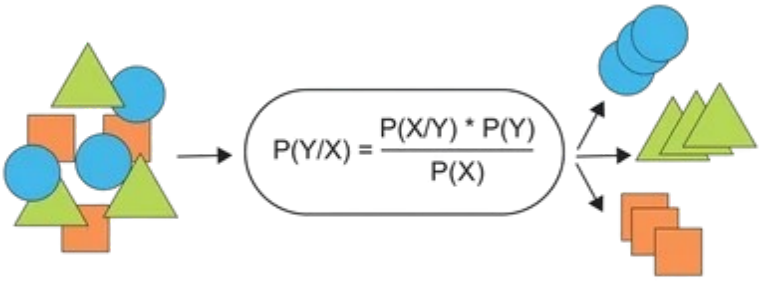
Best Hyper Tuning Parameters

kernel : rbf
c : 1000

Classification Report - Support Vector

	precision	recall	f1-score	support
0	0.79	0.84	0.82	45
1	0.90	0.87	0.88	75
accuracy			0.86	120
macro avg	0.85	0.86	0.85	120
weighted avg	0.86	0.86	0.86	120

GaussianNB Classifier



Classification Report - GaussianNB

	precision	recall	f1-score	support
0	0.98	0.98	0.98	45
1	0.99	0.99	0.99	75
accuracy			0.98	120
macro avg	0.98	0.98	0.98	120
weighted avg	0.98	0.98	0.98	120

KNeighbors Classifier

Classification Report - KNeighbors

	precision	recall	f1-score	support
0	0.98	0.98	0.98	45
1	0.99	0.99	0.99	75
accuracy			0.98	120
macro avg	0.98	0.98	0.98	120
weighted avg	0.98	0.98	0.98	120

Algorithm Used	Accuracy
Decision Tree Classifier	0.95
Random Forest Classifier	0.97
Support Vector Classifier	0.86
GaussianNB Classifier	0.98
KNeighbors Classifier	0.98

By Analyzing the classification report of various ML Models, the Accuracy for GaussianNB and Kneighbours Algorithm is 98%

So the best Model is GaussianNB and Kneighbours