CKD – Diseased Prediction

1. **Problem Statement:** To Predict the Chronic Kidney Disease (CKD)
2. **About Dataset:** Has 399 Rows and 28 Columns of Data
3. **Pre-processing method:** StandardScaler
4. **Algorithm:** Machine Learning Algorithms (Logistics, Support Vector Machine, Random Forest, Decision Tree, KNN, Naive Baye's)
5. **Models:** Results as below.
6. **Best Model: Logistic Grid Classification AND SVM Grid Classification.**

   These models provide the Prediction Accuracy as 99%, which is the highest

The report:

	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

Details:

Logistic Grid Classification: (This is one of the best model for this Data, since the Accuracy, Macro & Weighted Avg is 99% which is the highest)

#1.What is the % of Correct classification of both Diseased and Not-Diseased to the Total input in the test set = 99% (Accuracy)

#2.What is the % of Correct classification of Diseased people to the Correctly and Wrongly Diseased in the test set = 100% (Precision)

#3.What is the % of Correct classification of Diseased people to the Correctly Diseased in the test set = 99% (Recall)

#4.Overall Performance of Diseased people = 99% (F1-Score)

#5.What is the % of Correct classification of Not-Diseased people to the Correctly and Wrongly Not-Diseased in the test set = 98% (Precision)

#6.What is the % of Correct classification of Not-Diseased people to the Correctly Not-Diseased in the test set = 100% (Recall)

#7.Overall Performance of Not-Diseased people = 99% (F1-Score)

#8.What is the average performance of Correctly and Wrongly classified = 99% (Macro Avg- Precision)

#9.What is the average performance of Correctly classified = 99% (Macro Avg- Recall)

#10.What is the average performance of Overall Performance = 99% (Macro Avg- F1-Score)

#11.What is the sum of the product of the proportion rate and the Correctly & Wrongly Diseased in the test set = 99% (Weighted Avg- Precision)

#12.What is the sum of the product of the proportion rate and the Correctly Diseased in the test set = 99% (Weighted Avg- Recall)

#13.What is the sum of the product of the proportion rate and the Overall Performance of the test set = 99% (Weighted Avg- F1-score)

SVM Grid Classification: (This is one of the best model for this Data, since the Accuracy, Macro & Weighted Avg is 99% which is the highest)

#1.What is the % of Correct classification of both Diseased and Not-Diseased to the Total input in the test set = 99% (Accuracy)

#2.What is the % of Correct classification of Diseased people to the Correctly and Wrongly Diseased in the test set = 100% (Precision)

#3.What is the % of Correct classification of Diseased people to the Correctly Diseased in the test set = 99% (Recall)

#4.Overall Performance of Diseased people = 99% (F1-Score)

#5.What is the % of Correct classification of Not-Diseased people to the Correctly and Wrongly Not-Diseased in the test set = 98% (Precision)

#6.What is the % of Correct classification of Not-Diseased people to the Correctly Not-Diseased in the test set = 100% (Recall)

#7.Overall Performance of Not-Diseased people = 99% (F1-Score)

#8.What is the average performance of Correctly and Wrongly classified = 99% (Macro Avg- Precision)

#9.What is the average performance of Correctly classified = 99% (Macro Avg- Recall)

#10.What is the average performance of Overall Performance = 99% (Macro Avg- F1-Score)

#11.What is the sum of the product of the proportion rate and the Correctly & Wrongly Diseased in the test set = 99% (Weighted Avg- Precision)

#12.What is the sum of the product of the proportion rate and the Correctly Diseased in the test set = 99% (Weighted Avg- Recall)

#13.What is the sum of the product of the proportion rate and the Overall Performance of the test set = 99% (Weighted Avg- F1-score)

RF Grid Classification:

#1.What is the % of Correct classification of both Diseased and Not-Diseased to the Total input in the test set = 99% (Accuracy)

#2.What is the % of Correct classification of Diseased people to the Correctly and Wrongly Diseased in the test set = 99% (Precision)

#3.What is the % of Correct classification of Diseased people to the Correctly Diseased in the test set = 99% (Recall)

#4.Overall Performance of Diseased people = 98% (F1-Score)

#5.What is the % of Correct classification of Not-Diseased people to the Correctly and Wrongly Not-Diseased in the test set = 98% (Precision)

#6.What is the % of Correct classification of Not-Diseased people to the Correctly Not-Diseased in the test set = 98% (Recall)

#7.Overall Performance of Not-Diseased people = 98% (F1-Score)

#8.What is the average performance of Correctly and Wrongly classified = 98% (Macro Avg- Precision)

#9.What is the average performance of Correctly classified = 98% (Macro Avg- Recall)

#10.What is the average performance of Overall Performance = 98% (Macro Avg- F1-Score)

#11.What is the sum of the product of the proportion rate and the Correctly & Wrongly Diseased in the test set = 98% (Weighted Avg- Precision)

#12.What is the sum of the product of the proportion rate and the Correctly Diseased in the test set = 98% (Weighted Avg- Recall)

#13.What is the sum of the product of the proportion rate and the Overall Performance of the test set = 98% (Weighted Avg- F1-score)

Decision Tree Grid Classification

#1.What is the % of Correct classification of both Diseased and Not-Diseased to the Total input in the test set = 95% (Accuracy)

#2.What is the % of Correct classification of Diseased people to the Correctly and Wrongly Diseased in the test set = 97% (Precision)

#3.What is the % of Correct classification of Diseased people to the Correctly Diseased in the test set = 95% (Recall)

#4.Overall Performance of Diseased people = 96% (F1-Score)

#5.What is the % of Correct classification of Not-Diseased people to the Correctly and Wrongly Not-Diseased in the test set = 92% (Precision)

#6.What is the % of Correct classification of Not-Diseased people to the Correctly Not-Diseased in the test set = 96% (Recall)

#7.Overall Performance of Not-Diseased people = 94% (F1-Score)

#8.What is the average performance of Correctly and Wrongly classified = 95% (Macro Avg- Precision)

#9.What is the average performance of Correctly classified = 96% (Macro Avg- Recall)

#10.What is the average performance of Overall Performance = 95% (Macro Avg- F1-Score)

#11.What is the sum of the product of the proportion rate and the Correctly & Wrongly Diseased in the test set = 96% (Weighted Avg- Precision)

#12.What is the sum of the product of the proportion rate and the Correctly Diseased in the test set = 95% (Weighted Avg- Recall)

#13.What is the sum of the product of the proportion rate and the Overall Performance of the test set = 96% (Weighted Avg- F1-score)
