To find the following Machine Learning Classification using Confusion Matrix

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.

Sample dataset: "CKD.csv"

1.) Identify your problem statement

- Machine Learning – (predict the insurance charge – numbers)
- Supervised – (we have input and output)
- Classification Type – (Confusion Matrix type)

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

- Columns – (25 columns and 1 column is output)
- No.of rows is 399 (exclude column header)
- Input column is 24 (age, bp, sg, rbc, etc.)
- Output column is 1 (classification)

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

- We are using Standardization(Pre-processing, StandardScaler) method

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.

- I'm going with Logistics Grid comparing with other model. Execution time is less and more productivity to other model.
- Accuracy of this model is 0.99 near to 1.0 compare to other model.
- Precision – 1.00, recall – 0.99, f1-score – 0.99 and roc_auc_score exactly 1.

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

- Refer the document "Grid Classification Assignment.pdf

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

- Logistics Grid – solver(newton-cg), penalty(12)
- Roc_auc_score is equal to 1