Classification Assignment

**Problem Statement**

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) Problem statement

Stage 1: Machine Learning

Stage 2: Supervised

Stage 3: Classification

1. Total number of rows, columns

Num of Rows – 399

Num of columns – 25

1. Mention the pre-processing method:

Categorical data – Nominal data (Convert String to number)

By using One Hot Encoding Algorithm

1. Develop a good model with Evaluation Metrics:

Good Model = SVM - F1 Score 0.98

|  |  |  |
| --- | --- | --- |
| S. No | Machine Learning – Classification Algorithms | Evaluation metrics-f1 Score |
| 1. | Support Vector Machine | 0.98 |
| 2. | Decision Tree | 0.92 |
| 3. | Random Forest | 0.97 |
| 4. | Logistic Regression | 0.93 |
| 5. | KNN | 0.71 |
| 6. | Naïve Bayes – Bernoulli NB | 0.93 |

1. All the research values (Evaluation Metrics)
2. Final Model:

SVM - F1 Score 0.98