**Classification Assignment Problem**

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.

1. Identify your problem statement

**To predict the Chronic Kidney disease based on the inputs provided by the Hospital Management**

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

**399 rows × 25 columns**

**classification**

**yes 249**

**no 150**

**Name: count, dtype: int64**

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

**The One hot Encoding method is used to convert categorical columns into ordinal data**

**After 399 rows \* 40 columns**

**Standardization to improve the model as there are many rows**

1. 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.

**Random Forest makes best model with accuracy of 0.99**

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





