**Classification Assignment**

**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.

1. Identify your problem statement

Need to predict CKD – Yes / No

1. Data table given – so we can choose ML
2. Both Input and Output are clearly given – Supervised Learning
3. Output is categorical – Classification
4. Tell basic info about the dataset (Total number of rows, columns)

399 rows × 25 columns

Input Contains both numerical and categorical dataset

Output contains categorical dataset

Index(['age', 'bp', 'sg', 'al', 'su', 'rbc', 'pc', 'pcc', 'ba', 'bgr', 'bu',

'sc', 'sod', 'pot', 'hrmo', 'pcv', 'wc', 'rc', 'htn', 'dm', 'cad',

'appet', 'pe', 'ane', 'classification']

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

Using get dummies function convert the str to flt