Assignment- Classification

Chronic Kidney Disease prediction

1. Input and output values are well defined. So Supervised learning.
2. The target of the project is categorical values.

So, it falls under Classification

1. The dataset has 399 rows and 25 columns.
2. We used preprocessing method to convert strings to numeric data using One Hot Encoding method. For this I have used python’s “get\_dummies” method.
3. We tried the below classification algorithms
4. Logistic Regression:

A screenshot of a computer

Description automatically generated

1. SVMC

A screenshot of a computer

Description automatically generated

3.Decision\_TreeA screenshot of a computer

Description automatically generated 4.Random Forest A screenshot of a computer

Description automatically generated

1. KNN

A screenshot of a computer

Description automatically generated

1. GausianNB

A screenshot of a computer

Description automatically generated

A screenshot of a computer code

Description automatically generated

1. BernoulliNB

A screenshot of a computer program

Description automatically generated

A screen shot of a computer code

Description automatically generated

1. Multinomial

A screenshot of a graph

Description automatically generated

A screenshot of a computer

Description automatically generated

1. ComplementNB

A screenshot of a computer

Description automatically generated

A screenshot of a computer code

Description automatically generated

1. Categorical NB

A screenshot of a computer

Description automatically generated

11.Ada boosting:

A screenshot of a computer

Description automatically generated

12.LG BoostingA screenshot of a computer program

Description automatically generated

13.XG Boosting:

A screenshot of a computer

Description automatically generated

The Final model will be “ Logistic Regression” since the Accuracy and F1-score are high more than other algorithms.

So we will save and deploy the “ Logistic Regression”