ASSIGNMENT II

Load the DATASET:

```
#load the data from google.colab import files
# only use for google colab
Uploaded = files.upload()
#Only use for google colab
df = pd.read_csv("kidney_disease.csv")
#Print the first 5 rows
df.head()
```

Split the Data:

```
#Split the data
X= df.drop(["classification"],axis=1)
Y=df["classification"]
```

We will Split the data sets into 80% Training (x_train and y_train) and 20% Testing (x_test and y_test).

```
#Split the data into 80%training and 20% testing
X_train, x_test, y_train,y_test =train_test_Split( x, y, test_size = 0.2, shuffle = True)
```

Missing value:

```
#Create a list of columns to retain
Columns_to_retain = ["sg", "al", " sc ", "hemo", "pcv", "wbcc", " rbcc", "htn", " classification" ]
```

#Columns_to_retain = df.colums, Drop the columns that are not in columns_to_retain = df.drop([col for col in df.columns_if not col in columns_to_retain], axis =1)

#drop the rows with na or missing values

df = df.dropna(axis=0)

Scale the independent variable:

#feature scaling

X_scaler = MinMaxScaler()

X_scaler.fit(x)

Column_names = x.columns

X[column_names] =x_scaler.transform(x)

BI-VARIATE :

#compile the model

Model.compile(loss='binary_crossentropy', optimizer= 'adam',metrics=['accuracy'])

Submitted by,

Krishna veni.B

Harini Krishnaveni.M

Lavanya.S

Pavithra. V