*A project report on*

**Chronic Kidney Disease Prediction**

*Submitted in partial fulfillment for the award of the degree of*

**Integrated MTech Software Engineering**

*by*

**Sai Sheshank Gaddam (17MIS7099)**



**AMARAVATI**

MAY,2020

**CERTIFICATE**

This is to certify that the thesis entitled “[**Chronic**](https://eu-gb.dataplatform.cloud.ibm.com/projects/cd3c10a5-6d1c-43e1-b2fb-7cce743b7731?context=analytics) **Kidney Disease Prediction**” submitted by G Sai Sheshank (17MIS7099) VITAP, for the award of the Summer Internship for the bonafide work carried out by him under my supervision.

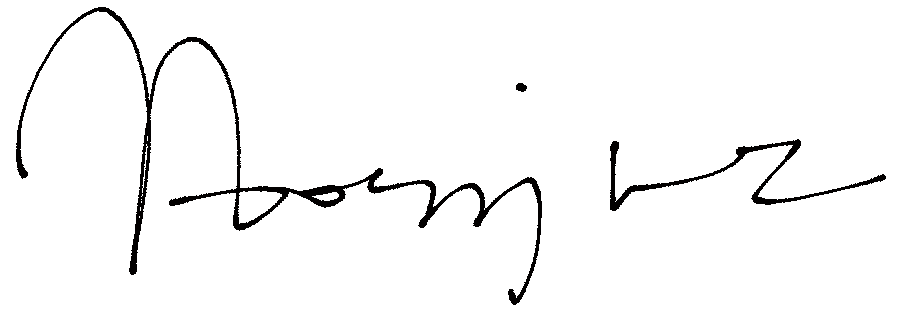
The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma in this institute or  
any other institute or university. The Project report fulfils the requirements and regulations of  
VIT-AP and in my opinion meets the necessary standards for submission.

**Signature of the Guide Signature of the HOD**



**CERTIFICATE BY THE EXTERNAL GUIDE**

This is to certify that the project report entitled “**Chronic Kidney Disease Prediction**” submitted by G Sai Sheshank (**17MIS7099**) to VIT-AP in partial fulfilment of the requirement for the award of the degree of Integrated MTech Software Engineering for the bonafide work carried out by him/her under my guidance. The project fulfil the requirements as per the regulations of this Institute and in my opinion meets the necessary standards for submission. The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.



(Nagarjuna Madluri)

Signature of the External Supervisor

**G. Sai Sheshank**

**ABSTRACT**

The Internship in the The SmartBridge company(In Collaboration with IBM) as Artificial Intelligence Intern. The process of the Internship is that first we will be getting trained on the Python for Data Science and then Deep Learning algorithms, where we have come across how AI will be leading the future and the types of Learnings in ML, Deep Learning algorithms such as ANN, CNN, NLP with which we worked on number of datasets and Pre-process the data later apply suitable AI algorithms to find the best accuracy of the model and predict the application necessary. Later, complete overview of IBM Watson Studio where we will be deploying and Integrate apps through Node-Red for the and then the final project development.

This project is aimed at developing a model which predicts whether a person is affected by kidney disease based on certain parameters, demonstrates a promising capability of reducing the uncertainties in the prediction model. Accurate and reliable disease predictions are necessary to the people in need.

**ACKNOWLEDGEMENT**

It is my pleasure to express with deep sense of gratitude to Professor BKSP. Kumar Raju, VIT-AP, for his constant guidance, continual encouragement, understanding; more than all, he taught me patience in my Endeavor. My association with him is not confined to academics only, but it is a great opportunity on my part of work with an intellectual and expert in the field of coding.

I would like to express my gratitude Dr. G. Viswanathan, Dr. Sankar Viswanathan, Dr. D. Subhakar, and Prof. Jagadish Mudiganti, Dean of School of Computer Science & Engineering, for providing with an environment to work in and for his inspiration during the tenure of the course.

In jubilant mood I express ingeniously my whole-hearted thanks to Dr. Pradeep Reddy CH, Associate Professor, all teaching staff and members working as limbs of our university for their not-self-centered enthusiasm coupled with timely encouragements showered on me with zeal, which prompted the acquirement of the requisite knowledge to finalize my course study successfully. I would like to thank my parents for their support.

It is indeed a pleasure to thank my friends who persuaded and encouraged me to take up and complete this task. At last but not least, I express my gratitude and appreciation to all those who have helped me directly or indirectly toward the successful completion of this project.

Place: AMARAVATHI  **G. Sai Sheshank**

Date: 05.07.2020  **17MIS7099**

**CONTENTS**

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

**CHAPTER 1**

**INTRODUCTION**

* 1. INTRODUCTION 11
  2. PROBLEM STATEMENT .11

1.3 STUDY OF THE SYSTEM ………………………. .11

1.4 MODULE DESCRIPTION ………12-14

1.5 SCOPE OF THE PROJECT 14

**CHAPTER 2**

2.1 DATA PREPROCESSING………………………………………………………….15

2.2 METHODOLOGY……………………………………………………………..……16

2.2.1 IMPORTING DATA.………………………………………………………….16

2.2.2 DATA CLEANING……………………….……………………………….…..16-17

2.2.3 EXPLORING DATA ANALYSIS …..………………………………………...18

2.2.4 VISUALIZING THE DATA ………………………………………………….19-22

2.3 MODELLING…………………………………………………………………….…23-24

2.4 BUILDING USER INTERFACE……………………………………………………25

2.5 PREDICTION……………………………………………………………………….26

**CHAPTER 3**

CODE…………………………………………………………………….27-41

**CHAPTER 4**

CONCLUSION AND FUTURE WORK…………….……………………44

**CHAPTER 5**

REFERENCES….………………………………………………………….45

**LIST OF FIGURES**

ACCURACY 42

**LIST OF TABLES**

DATASET……………………………………………………………………….43

**Chapter1**

**1.1 INTRODUCTION**

A global health problem which is steadily growing is Chronic kidney disease (CKD). It is a chronic condition associated with increased morbidity and mortality, a high risk of many other diseases including cardiovascular disease, and high health care costs.

**1.2 PROBLEM STATEMENT**

* There is significant importance in the early detection, controlling, and managing of the disease. It is necessary to predict the progression of CKD with reasonable accuracy because of its dynamic and covert nature in the early stages, and patient heterogeneity. CKD is often described by severity stages.
* Clinical decisions are influenced by the stage, whether a patient is progressing, and the rate of progression. Also, defining the disease stage is quite crucial as it gives several indications that support the determination of required intervention and treatments.

# 1.3 STUDY OF THE SYSTEM

Over two million people worldwide receive dialysis or kidney transplant treatment to stay alive, yet this number may represent only 10% of people who need treatment to live. The majority of the 2 million people who receive treatment for kidney failure are in only five relatively wealthy countries, which represent 12% of the global population. By comparison, only 20% of the world's population is treated in about 100 developing countries, and they represent almost half the global population. Annually, more than one million people in 112 lower-income countries die from untreated kidney failure, due to the huge financial burden of dialysis or kidney transplantation treatment.

**1.4 MODULE DESCRIPTION**

The system after careful analysis has been identified to be presented with the following modules:

**MODULES INVOLVED**

* Loading of Data
* Data Pre-Processing
* Data Visualization
* ANN Classification
* Model Creation
* Integrating through Flask
* Designing UI

### **LOADING OF DATA**

The given data set will be loaded into the Jupyter notebook for data analyzation and further processes will be made to predict the correct output in the note book. Later, the data given is pre-processed and visualised to understand the exact view of the given data even to the illiterate person. Finally, this will we generating a model using ANN and integrated the model to flask using python where it interacts with HTML and CSS to make the final web App.

### **DATA PRE-PROCESSING**

The raw data set will be given to us and then we need to clean the data in which we find out any missing data and encode the data in a labelled format so as to supervise learning. The given data consists of various unlabelled data in which every row and every column need to be understandable by the system. Since, human can read the data with his naked eye but machine understands only the binary data which is 1 and 0. So, even categorical data need to be changed to the numerical data such that to make machine understandable and scalable for training and testing the model. In this process the given raw data will be classified and labelled.

### **DATA VISUALIZATION**

In this process the cleaned data is further Regressed and classified. But to make the data understandable by each and every person the given health care data is being visualized with each attribute to each other attribute. It makes the Data Analyst the work easy to understand the correlation and convolution among the attributes.

### **ANN CLASSIFICATION**

### In this process a fixed Artificial Neural Network algorithm is applied to act on the data which is deeply classified and labeled which is known as the supervised learning. Artificial Neural Networks (ANN) are multi-layer fully-connected neural nets that look like the figure below. They consist of an input layer, multiple hidden layers, and an output layer.  Every node in one layer is connected to every other node in the next layer. We make the network deeper by increasing the number of hidden layers.

### A given node takes the weighted sum of its inputs, and passes it through a non-linear activation function. This is the output of the node, which then becomes the input of another node in the next layer. The signal flows from left to right, and the final output is calculated by performing this procedure for all the nodes. Training this deep neural network means learning the weights associated with all the edges.

**MODEL CREATION**

Once the pre-processed data is classified and trained / classified with ANN, there will be an accuracy check based on certain parameters like number of modes in the input layer, number nodes in the hidden layer, the output node, activation parameter for both input and output layers, optimizer used and all. Finally after the model we got the most accurate then it saved through .H5 file so as to integrate with web page using Flask.

**INTEGRATION THROUGH FLASK**

After the model is saved through .H5 file, we need to create a flask file in python where the .H5 file model is integrated here in Flask so as to connect with the web files to produce the final UI and get the correct prediction same as the accuracy we got in the model saving.

**DESIGNING UI**

The UI we created is nothing but the simple HTML file which includes the background colour, all the parameters as input required some as text for the numerical inputs and some as radio buttons for the categorical inputs and the prediction result which displays whether the person is affected by the kidney disease or not i.e, 0 – Yes, 1- No.

* 1. **SCOPE OF THE PROJECT**

AI can plays a major role in extracting hidden data from the large patient medical and clinical dataset that physicians frequently collect from patients to obtain insights about the diagnostic information, and to implement precise treatment plans. Artificial Intelligence can be defined as the process of extracting hidden data from a large dataset. AI techniques are applied and used widely in various contexts and fields. With ANN technique we could predict, classify, filter and predict data.

Pros:

With the best accuracy we predict the Kidney Disease. You don’t have to go to doctor for the confirmation. After you confirm with this prediction model you can go to the doctor and get the medication from them. Instead if you don’t have the disease then you don’t have to go to doctor all through the way. Doctors can also recheck the patients and their analysis through this model.

Cons:

Few people do negligence having these kind of model. Because they think that they are having a predictor so we can check at some time and we don’t have to go to doctor. So, people through their negligence they might get harm themselves. This might cause danger to their life.

**Chapter-2**

**2.1 DATA PREPROCESSING**

Firstly, in Artificial Intelligence we have to analyse whether the model is Supervised or Unsupervised as we have a target variable which is “**classification”** and also other predicting variables. So, it is a Supervised Learning as we have an output to generate.



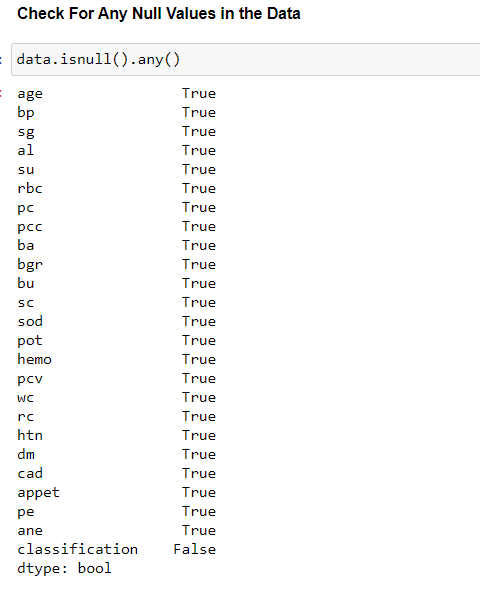
**2.2 METHODOLOGY**

**2.2.1 IMPORTING DATA**

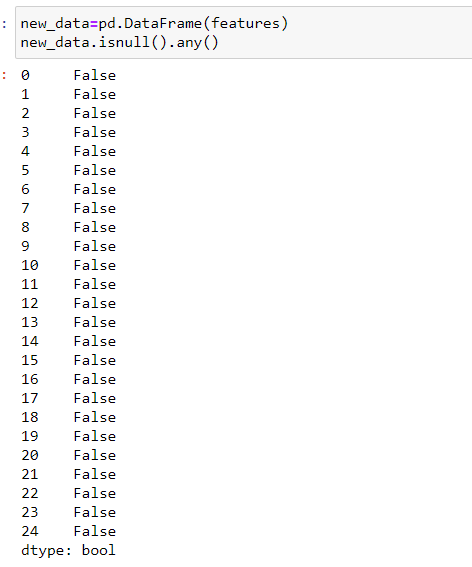
Data can be imported to a data frame in python using pandas.

**2.2.2 DATA CLEANING**

Data should be consistent for better prediction. So we need to check the data for null values. If we have null ,we need to either replace or delete the rows.

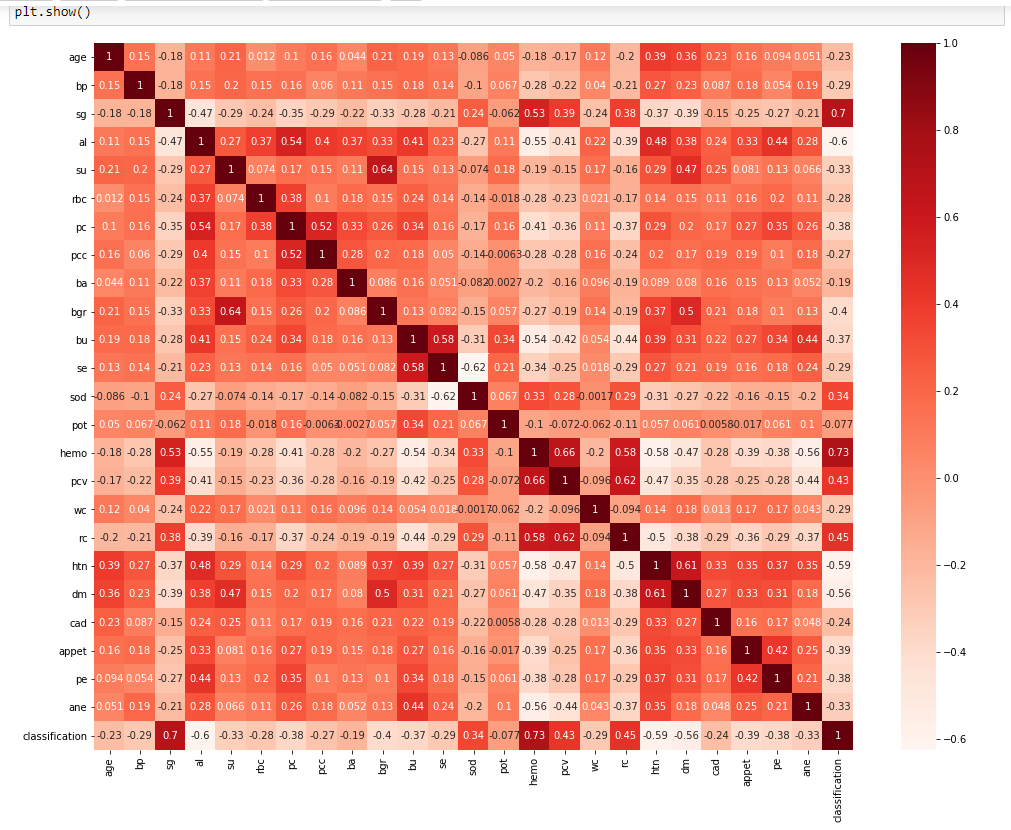


After checking for null values, we can clearly see that the dataset does not contain any null values.



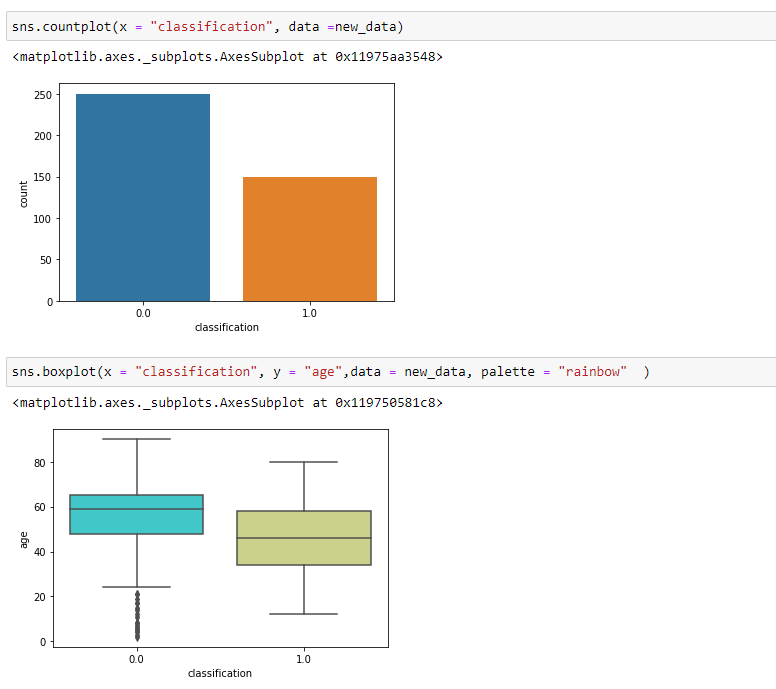
**2.2.3 EXPLORATORY DATA ANALYSIS**

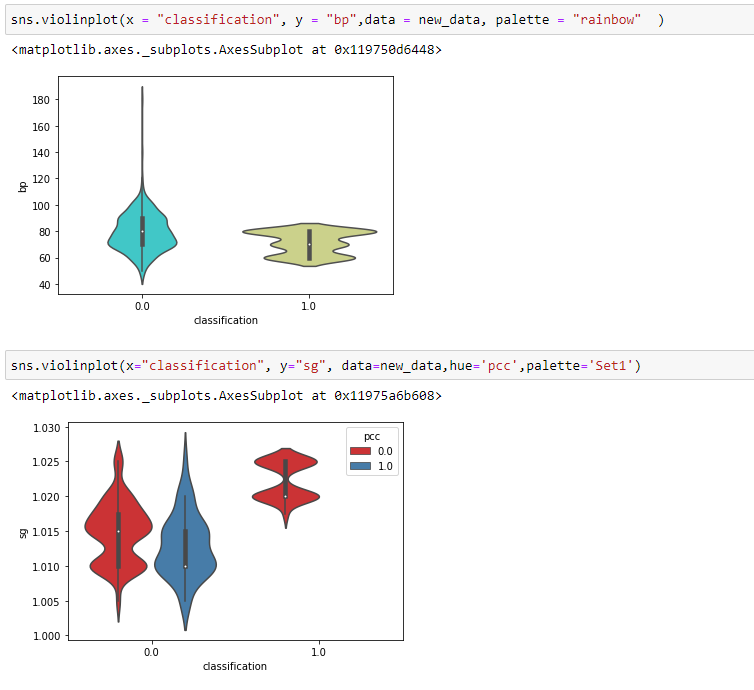
Correlation: It is the term which shows how efficiently every attribute is correlated to the every-other attribute. As we know the target variable, we’ll just have a look on how all the attributes are correlated to the Target variable.



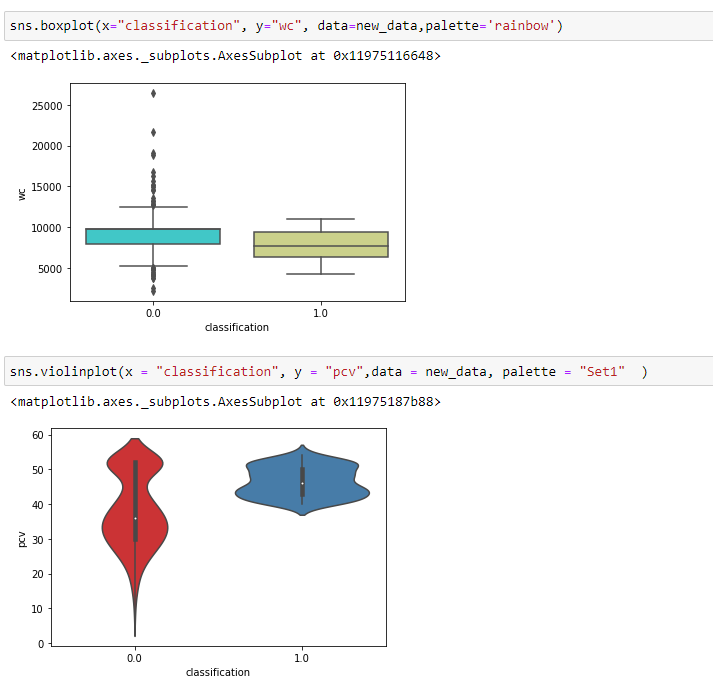
This visualization is same as the correlation tables which shows the magnitude/ relationship between two attributes.

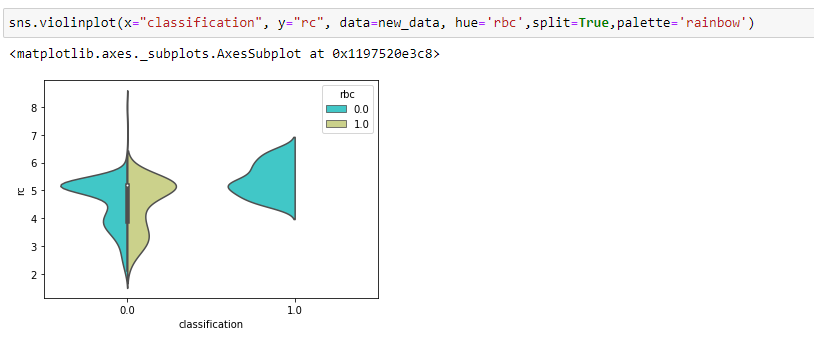
**2.2.4 VISUALIZING THE DATA**









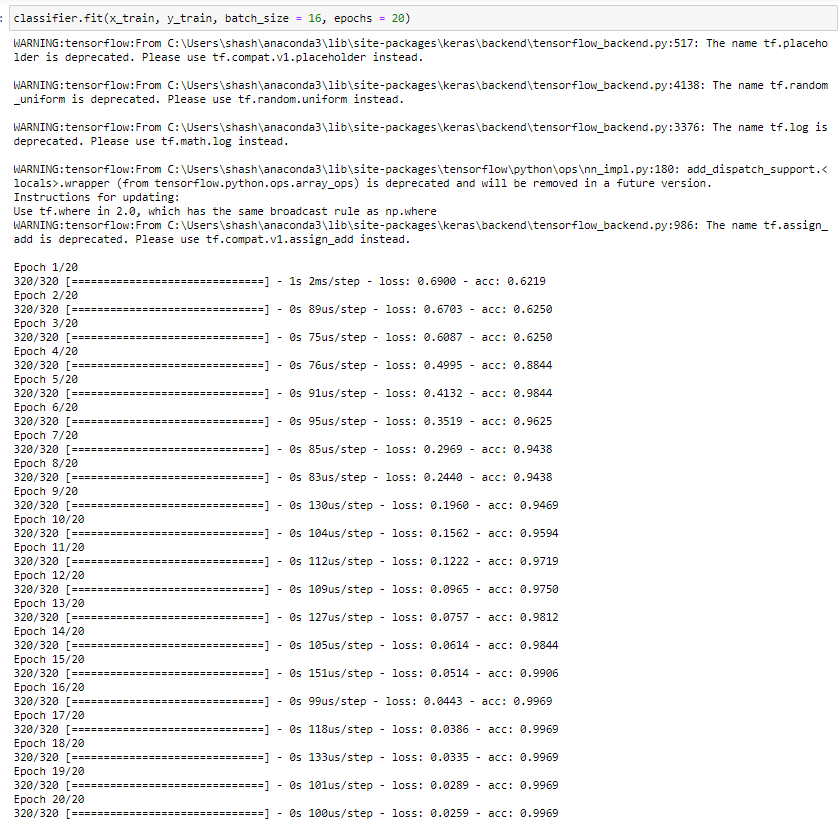


**2.3 MODELLING**

* In Artificial Intelligence we have Deep Learning algorithms such as ANN, CNN, NLP and so on. ANN is the correct algorithm to be used for our project.
* RNN is used when we have a categorical output and Numerical along with some other categorical outputs.
* For this project we used Artificial Neural Networks method since we have categorical output variable satisfaction have continuous / categorical values and metrics as accuracy.

**Artificial Neural Network-** Artificial Neural Network algorithm is applied to act on the data which is deeply classified and labeled which is known as the supervised learning. Artificial Neural Networks (ANN) are multi-layer fully-connected neural nets that look like the figure below. They consist of an input layer, multiple hidden layers, and an output layer.  Every node in one layer is connected to every other node in the next layer.

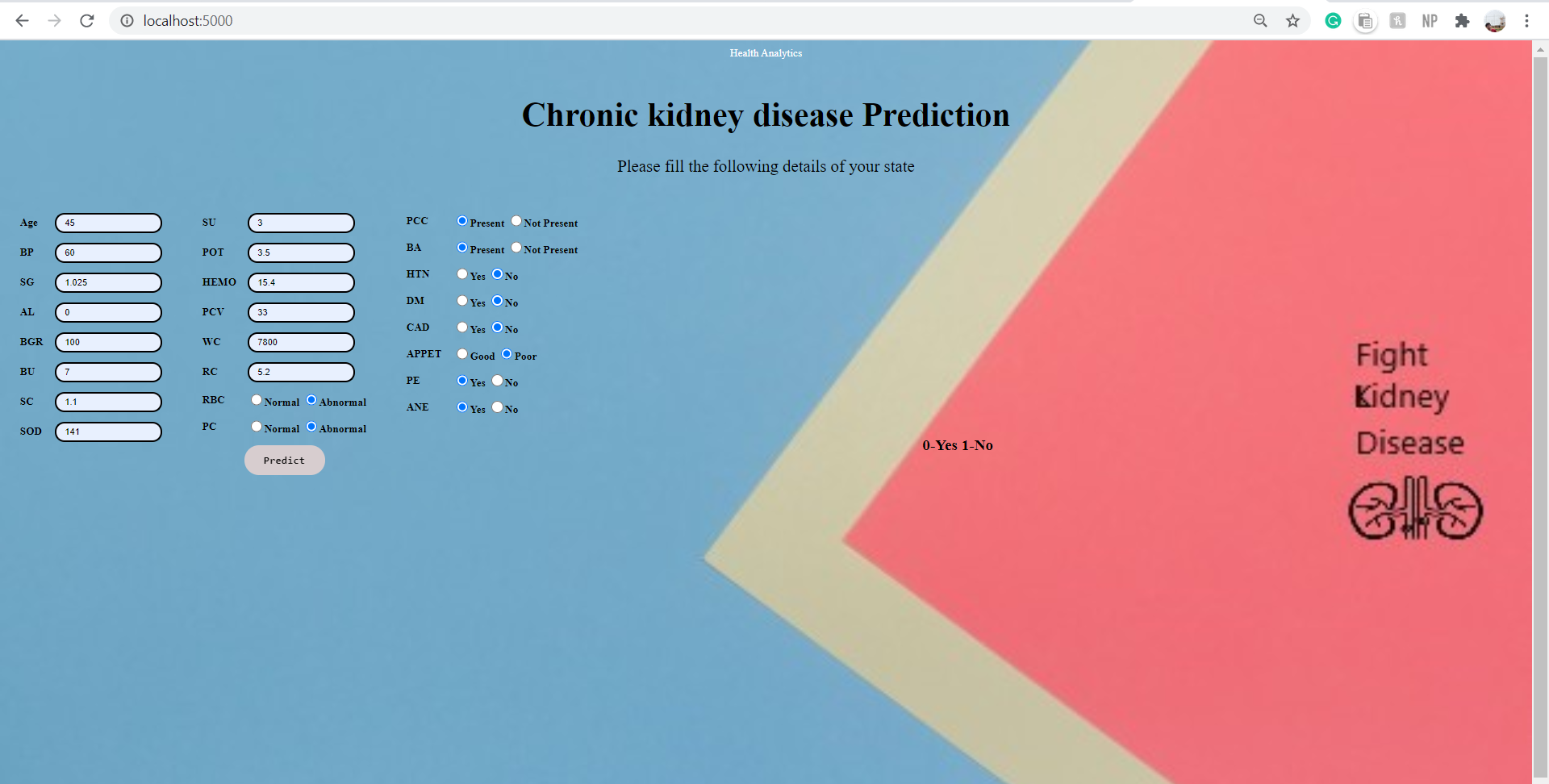




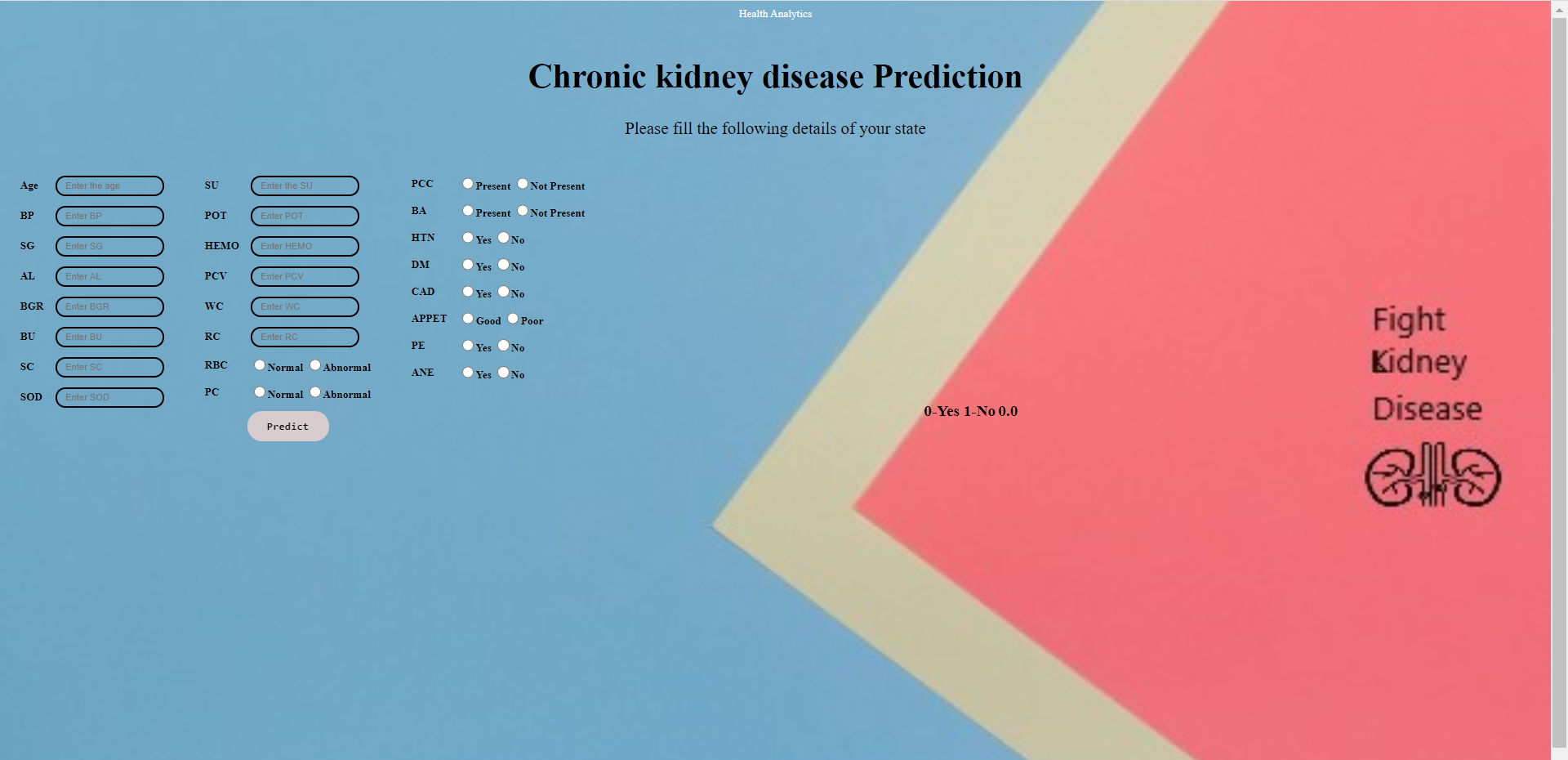
**2.4 BUILDING USER INTERFACE FOR THE MODEL:**

Flask is a lightweight WSGI web application framework. It is designed to make getting started quick and easy, with the ability to scale up to complex applications. It began as a simple wrapper around Werkzeug and Jinja and has become one of the most popular Python web application frameworks.

It is an API of Python that allows us to build up web-applications. It was developed by Armin Ronacher. Flask's framework is more explicit than Django's framework and is also easier to learn because it has less base code to implement a simple web-Application.



**2.5 PREDICTION(UI):**



The predicted value is displayed beside.

**Chapter – 3**

**CODE**

ANN Model (IPNBY):

I Almost mentioned this codes as snippets in the earlier descriptions of the project.

Flask File (PY):

# -\*- coding: utf-8 -\*-

"""

Created on Sun Apr 26 14:27:17 2020

@author: shash

"""

from flask import Flask, render\_template, request

from keras.models import load\_model

import numpy as np

global model,graph

import tensorflow as tf

# load the pre-trained Keras model

model = load\_model('kidney\_disease.h5')

graph = tf.get\_default\_graph()

app = Flask(\_\_name\_\_)

@app.route('/')

def index():

return render\_template('final.html')

@app.route('/age')

@app.route('/', methods=['GET','POST'])

def form\_post():

a = request.form['age']

b= request.form['bp']

c = request.form['sg']

d = request.form['al']

e = request.form['bgr']

f= request.form['bu']

g = request.form['sc']

h = request.form['sod']

i = request.form['su']

j= request.form['pot']

k = request.form['hemo']

l = request.form['pcv']

m = request.form['wc']

n= request.form['rc']

o = request.form['rbc']

if (o == "abnormal"):

o1 = 0

else:

o1 = 1

p = request.form['pc']

if (p == "abnormal"):

p1 = 0

else:

p1 = 1

q = request.form['pcc']

if (q == "notpresent"):

q1 = 0

else:

q1 = 1

r= request.form['ba']

if (r == "notpresent"):

r1 = 0

else:

r1 = 1

s = request.form['htn']

if (s == "no"):

s1 = 0

else:

s1 = 1

t = request.form['dm']

if (t == "no"):

t1 = 0

else:

t1 = 1

u = request.form['cad']

if (u == "no"):

u1 = 0

else:

u1 = 1

v= request.form['appet']

if (v == "good"):

v1 = 0

else:

v1 = 1

w = request.form['pe']

if (w == "no"):

w1 = 0

else:

w1 = 1

x = request.form['ane']

if (x == "no"):

x1 = 0

else:

x1 = 1

total = [[a,b,c,d,i,o1,p1,q1,r1,e,f,g,h,j,k,l,m,n,s1,t1,u1,v1,w1,x1]]

print(total)

with graph.as\_default():

y\_pred = model.predict(np.array(total))

return render\_template('final.html',ypred = str(y\_pred[0][0]))

if \_\_name\_\_ == '\_\_main\_\_':

app.run(host='localhost', debug=True, threaded=False)

HTML File:

<html>

<title >Chronic kidney disease Prediction</title>

<head>

</head>

<style>

body, html {

height: 100%;

margin: 0;

}

.body1 {

background-image: url("final1.jpg");

height: 100%;

background-position: center;

background-repeat: no-repeat;

background-size: cover;

}

#rcorners1 {

border-radius: 15px;

background: transparent;

border-color: black;

padding: 13px;

width: 160px;

height: 30px;

}

.button4 {

border-radius: 100%;

border-color: #d57d7a;

width:0px;

font-size: 17px;

backface-visibility: hidden;

position: relative;

width: 0120px;

height: 044px;

cursor: pointer;

display: inline-block;

white-space: nowrap;

background: #d7cdcf;

border-radius: 100px;

border: 0px solid #444;

border-width: 0px 0px 0px 0px;

padding: 10px 13px 10px 13px;

color: #0a0909;

font-size: 16px;

font-family: consolas;

font-weight: 900;

font-style: normal

}

.radion {

border:2px;

border-color:black;

height:17px;

width:17px;

}

</style>

<body style="background-image: url('../static/img/i3.jpg'); background-repeat: no-repeat; background-attachment: fixed;background-size: cover;" class="serif">

<div class="container" style="margin-top: 10px">

<div class="row justify-content-md-center">

<div class="col-md-29">

<div class="card">

<div style="background-color: voilet;" class="card-header">

<center style="color:white">Health Analytics</center>

</div>

<div class="col-sm-10 bd" >

<br>

</div>

<form method="POST">

<div class="container">

<h1 align = 'center' style="color:black;font-family: Times New Roman; font-size:50px" >Chronic kidney disease Prediction</h1>

<p align = 'center'style="font-size:25px">Please fill the following details of your state</p>

<form action='/age' method="post">

<table align='left' style='padding:15px; border-spacing: 15px; font-size:20 px'>

<tr>

<td><label for="age"><b>Age</b></label></td>

<td><input type="text" placeholder="Enter the age" name="age" id="rcorners1" required></td>

</tr>

<tr>

<td><label for="Bp"><b>BP</b></label></td>

<td><input type="text" placeholder="Enter BP" name="bp" id="rcorners1" required></td>

</tr>

<tr>

<td><label for="sg"><b>SG</b></label></td>

<td><input type="text" placeholder="Enter SG" name="sg" id="rcorners1" required></td>

</tr>

<tr>

<td><label for="al"><b>AL</b></label></td>

<td><input type="text" placeholder="Enter AL" name="al" id="rcorners1" required></td>

</tr>

<tr>

<td><label for="bgr"><b>BGR</b></label></td>

<td><input type="text" placeholder="Enter BGR" name="bgr" id="rcorners1" required></td>

</tr>

<tr>

<td><label for="bu"><b>BU</b></label></td>

<td><input type="text" placeholder="Enter BU" name="bu" id="rcorners1" required></td>

</tr>

<tr>

<td><label for="sc"><b>SC</b></label></td>

<td><input type="text" placeholder="Enter SC" name="sc" id="rcorners1" required></td>

</tr>

<tr>

<td><label for="sod"><b>SOD</b></label></td>

<td><input type="text" placeholder="Enter SOD" name="sod" id="rcorners1" required></td>

</tr>

</table>

</form>

</table>

<table align='left' style='padding:15px; border-spacing: 15px; font-size:20 px'>

<tr>

<td><label for="su"><b>SU</b></label></td>

<td><input type="text" placeholder="Enter the SU" name="su" id="rcorners1" required></td>

</tr>

<tr>

<td><label for="pot"><b>POT</b></label></td>

<td><input type="text" placeholder="Enter POT" name="pot" id="rcorners1" required></td>

</tr>

<tr>

<td><label for="hemo"><b>HEMO</b></label></td>

<td><input type="text" placeholder="Enter HEMO" name="hemo" id="rcorners1" required></td>

</tr>

<tr>

<td><label for="pcv"><b>PCV</b></label></td>

<td><input type="text" placeholder="Enter PCV" name="pcv" id="rcorners1" required></td>

</tr>

<tr>

<td><label for="wc"><b>WC</b></label></td>

<td><input type="text" placeholder="Enter WC" name="wc" id="rcorners1" required></td>

</tr>

<tr>

<td><label for="rc"><b>RC</b></label></td>

<td><input type="text" placeholder="Enter RC" name="rc" id="rcorners1" required></td>

</tr>

<tr>

<td><label for="rbc"><b>RBC</b></label></td>

<td style = "vertical-align:middle">

<input type="radio" class="radion" name="rbc" value="normal" required><b>Normal</b>

<input type="radio" class="radion" name="rbc" value="abnormal" ><b>Abnormal</b>

</td>

</tr>

<tr>

<td><label for="pc"><b>PC</b></label></td>

<td style = "vertical-align:middle">

<input type="radio" class="radion" name="pc" value="normal" required><b>Normal</b>

<input type="radio" class="radion" name="pc" value="abnormal" ><b>Abnormal</b>

</td>

</tr>

<tr>

<td colspan="2" align="center" ><button class="button4" type="submit" name="predict" align='center' value="Predict" >Predict</button></td>

</tr>

</table>

<table align='left center' style='padding:15px; border-spacing: 15px; font-size:20 px'>

<tr>

<td><label for="pcc"><b>PCC</b></label></td>

<td style = "vertical-align:middle">

<input type="radio" class="radion" name="pcc" value="present" required><b>Present</b>

<input type="radio" class="radion" name="pcc" value="notpresent" ><b>Not Present</b>

</td>

</tr>

<tr>

<td><label for="ba"><b>BA</b></label></td>

<td style = "vertical-align:middle">

<input type="radio" class="radion" name="ba" value="present" required><b>Present</b>

<input type="radio" class="radion" name="ba" value="notpresent" ><b>Not Present</b>

</td>

</tr>

<tr>

<td><label for="htn"><b>HTN</b></label></td>

<td style = "vertical-align:middle">

<input type="radio" class="radion" name="htn" value="yes" required><b>Yes</b>

<input type="radio" class="radion" name="htn" value="no" ><b>No</b>

</td>

</tr>

<tr>

<td><label for="dm"><b>DM</b></label></td>

<td style = "vertical-align:middle">

<input type="radio" class="radion" name="dm" value="yes" required><b>Yes</b>

<input type="radio" class="radion" name="dm" value="no" ><b>No</b>

</td>

</tr>

<tr>

<td><label for="cad"><b>CAD</b></label></td>

<td style = "vertical-align:middle">

<input type="radio" class="radion" name="cad" value="yes" required><b>Yes</b>

<input type="radio" class="radion" name="cad" value="no" ><b>No</b>

</td>

</tr>

<tr>

<td><label for="appet"><b>APPET</b></label></td>

<td style = "vertical-align:middle">

<input type="radio" class="radion" name="appet" value="good" required><b>Good</b>

<input type="radio" class="radion" name="appet" value="poor" ><b>Poor</b>

</td>

</tr>

<tr>

<td><label for="pe"><b>PE</b></label></td>

<td style = "vertical-align:middle">

<input type="radio" class="radion" name="pe" value="yes" required><b>Yes</b>

<input type="radio" class="radion" name="pe" value="no" ><b>No</b>

</td>

</tr>

<tr>

<td><label for="ane"><b>ANE</b></label></td>

<td style = "vertical-align:middle">

<input type="radio" class="radion" name="ane" value="yes" required><b>Yes</b>

<input type="radio" class="radion" name="ane" value="no" ><b>No</b>

</td>

</tr>

</table>

<table align="center">

<tr>

<td><label for="age" style="font-size:23px"><b>0-Yes 1-No</b></label></td>

<td style="font-size:23px"><b>{{ypred}}</b></td>

</tr>

</table>

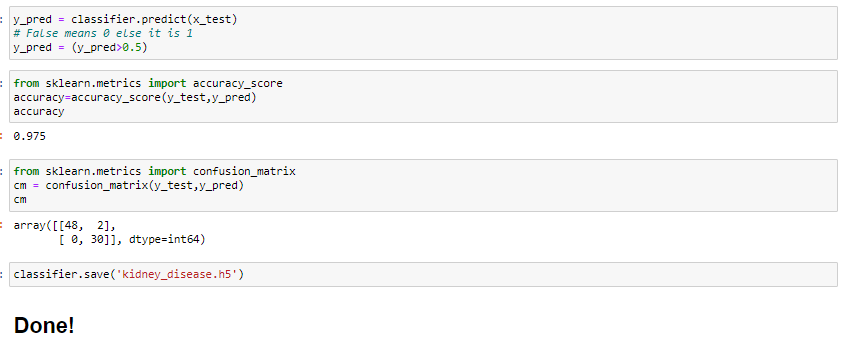
</form>

</div>

</form>

**LIST OF FIGURES**

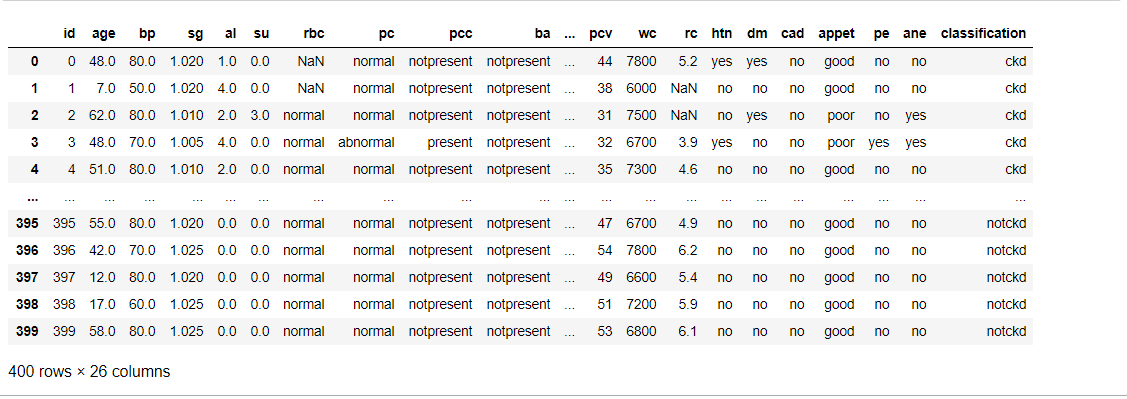
ACCURACY



Here, our model is giving an accuracy of 97.5%. There if we have a look on confusion matrix above we have only few False outcomes for the testing dataset.

**LIST OF TABLES**

DATASET



**Chapter-4**

**CONCLUSION AND FUTURE WORK**

* We developed this project using python3.
* By taking the dataset we evaluated with different models and checked the accuracy for which model we can get maximum percentage so that particular model can be sustained and taken as good model.
* Future work, we could take this project to higher levels after the good research and we’ll prepare a dataset so that we can even give the prescription for the patient according to his health condition and level of disease spread in his body.

**Chapter-5**

**REFERENCES:**

<https://www.kaggle.com/mansoordaku/ckdisease>

<https://towardsdatascience.com/data-visualization-for-machine-learning-and-data-sciencea45178970be7>

<https://thesmartbridge.com/documents/spsaimldocs/Datapreprocessing.pdf>

<https://thesmartbridge.com/documents/spsaimldocs/Datapreprocessing.pdf>

<https://thesmartbridge.com/documents/spsaimldocs/Datapreprocessing.pdf>

<https://thesmartbridge.com/documents/spsaimldocs/FlaskML.pdf>