CHRONIC KIDNEY DISEASE PREDICTION USING MACHINE LEARNING

TEAM ID: NM2023TMID31952

	Name of the Student	NM ID
TEAM LEADER	V. Vaira Muthu	A86E48562203FFFBAE24F7942F05B03C
TEAM MEMBERS	S. Esakkiraju	711EFA651BB52A994744124042516CA0
	M. Prem Muthu Pandi	A63050DA4F3BB82A02E49769FFB0BE6A
	P. Alageshan	5AD4982331EFA70B7294B8E70EA20914

GOVERNMENT ARTS & SCIENCE COLLEGE
KADAYANALLUR

1.INTRODUCTION

OVER VIEW

Chronic Kidney Disease (CKD) is considered as an important threat for the society with respect to the health in the present era. Chronic kidney disease can be detected with regular laboratory tests, and some treatments are present which can prevent development, slow disease progression, reduce complications of decreased Glomerular Filtration Rate(GFR) and risk of cardiovascular disease, and improve survival and quality of life. CKD can be caused due to lack of water consumption, smoking, improper diet, loss of sleep and many other factors. This disease affected 753 million people globally in 2016 in which 417 million are females and 336 million are males. Majority of the time the disease is detected in its final stage and which sometimes leads to kidney failure. The existing system of diagnosis is based on the examination of urine with the help of serum creatinine level. Many medical methods are used for this purpose such as screening, ultrasound method. In screening, the patients with hypertension, history of cardiovascular disease, disease in the past, and the patients who have relatives who had kidney disease are screened. This paper focuses on machine learning techniques like ACO and SVM by minimizing the features and selecting best features to improve the accuracy of prediction.

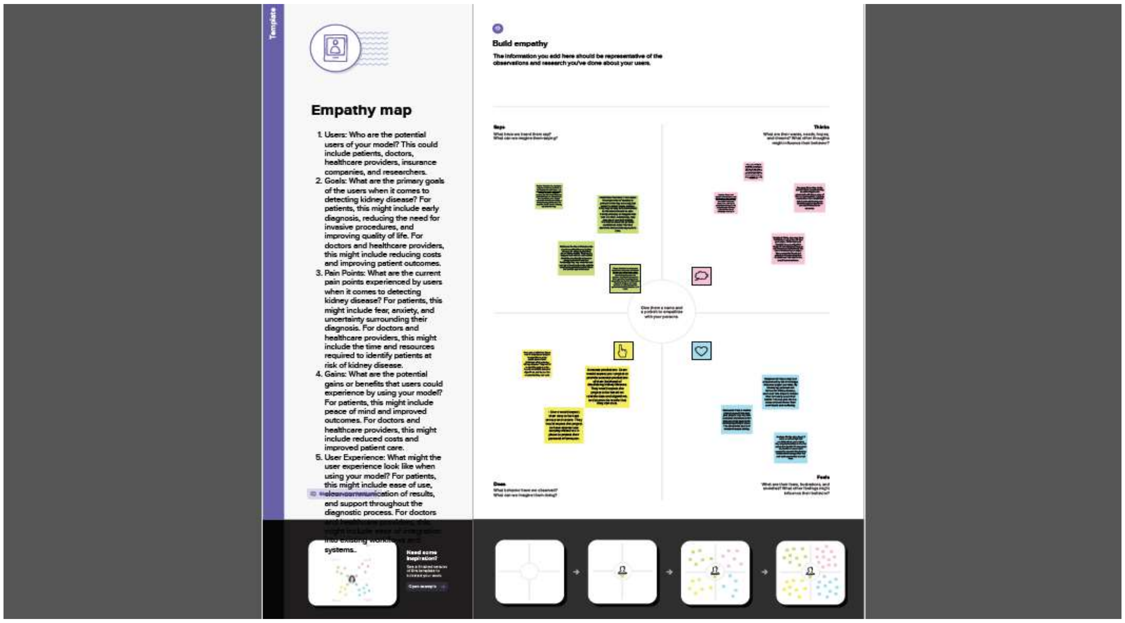
This work proposes a workflow to predict CKD status based on clinical data, incorporating data prepossessing, a missing value handling method with collaborative filtering and attributes selection. This study proposes the use of machine learning techniques for CKD such as Ant Colony Optimization (ACO) technique and Support Vector Machine (SVM) classifier. Final output predicts whether the person is having CKD or not by using minimum number of features. SVM works by mapping data to a high-dimensional feature space so that data points can be classified, even when the data are not otherwise linearly separable.

PURPOSE

The purpose of predicting chronic kidney disease is to identify individuals who are at risk of developing the condition so that appropriate interventions can be put in place to prevent or delay its progression. Early detection and management of chronic kidney disease can significantly improve outcomes and quality of life for affected individuals.Chronic kidney disease is a condition in which the kidneys gradually lose their ability to function properly over time. It is a common and serious condition that can lead to a range of complications, including cardiovascular disease, anemia, bone disease, and kidney failure. Risk factors for chronic kidney disease include diabetes, high blood pressure, smoking, obesity, family history, and certain medications.By predicting chronic kidney disease, healthcare providers can screen at-risk individuals for the condition and provide interventions to manage underlying risk factors, such as blood pressure and blood sugar control, smoking cessation, and medication management. This can help prevent or delay the progression of chronic kidney disease and improve outcomes for affected individuals

2.PROBLEM DEFINITION & DESIGN THINKING

EMPATHY MAP



Empathy map

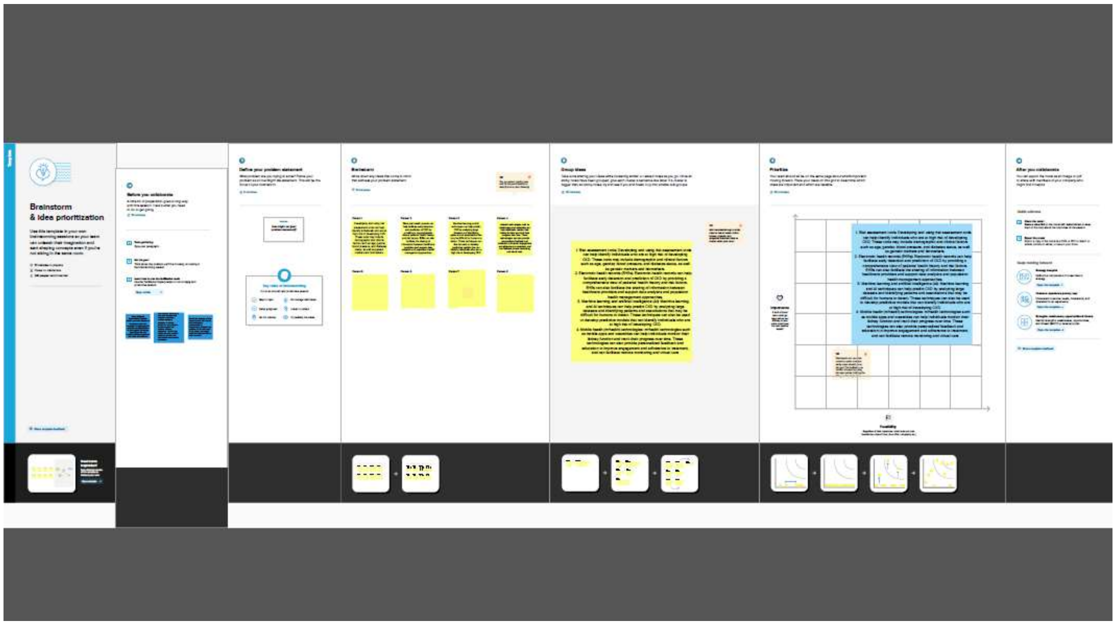
1. Users: Who are the potential users of your model? This could include patients, doctors, healthcare providers, insurance companies, and researchers.
2. Goals: What are the primary goals of the users when it comes to detecting kidney disease? For patients, this might include early diagnosis, reducing the need for invasive procedures, and improving quality of life. For doctors and healthcare providers, this might include reducing costs and improving patient outcomes.
3. Pain Points: What are the current pain points experienced by users when it comes to detecting kidney disease? For patients, this might include fear, anxiety, and uncertainty surrounding their diagnosis. For doctors and healthcare providers, this might include the time and resources required to identify patients at risk of kidney disease.
4. Gains: What are the potential gains or benefits that users could experience by using your model? For patients, this might include peace of mind and improved outcomes. For doctors and healthcare providers, this might include reduced costs and improved patient care.
5. User Experience: What might the user experience look like when using your model? For patients, this might include ease of use, clear communication of results, and support throughout the diagnostic process. For doctors and healthcare providers, this might include ease of integration into existing workflows and systems..

Build empathy

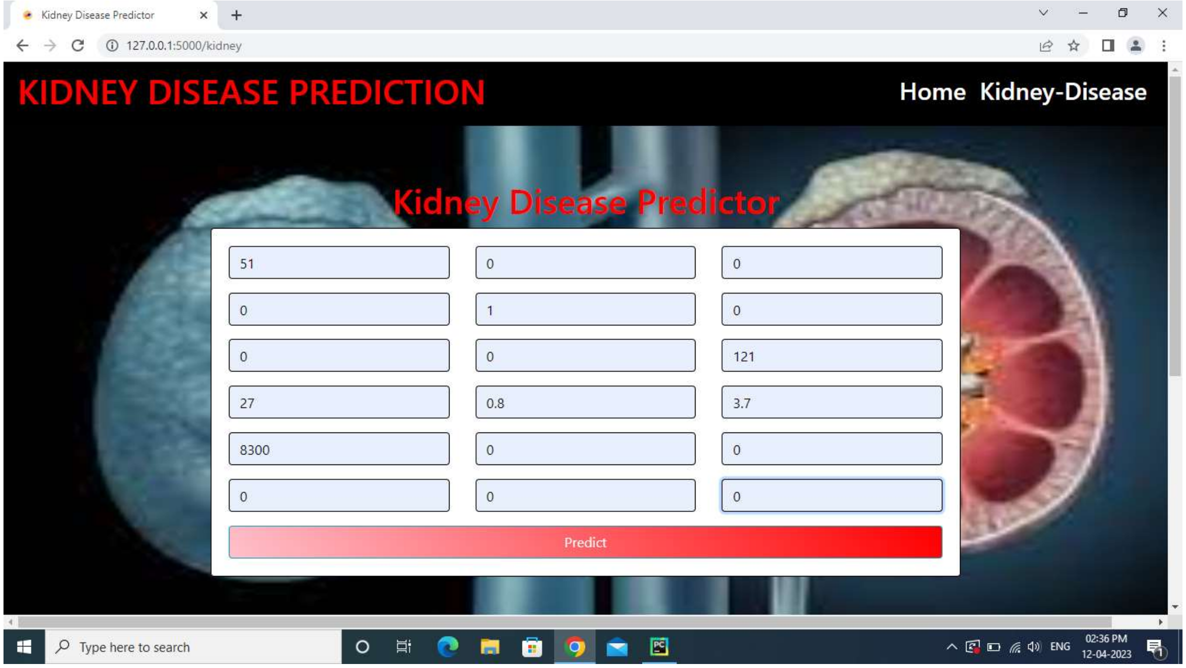
The information you add here should be representative of the observations and research you've done about your users.

Says
What have we heard them say?
What can we imagine them saying?
Thinks
What are their wants, needs, hopes, and dreams? What other thoughts might influence their behavior?
Give them a name and a portrait to empathize with your persona.
Does
What behavior have we observed?
What can we imagine them doing?
Feels
What are their fears, frustrations, and anxieties? What other feelings might influence their behavior?
Need some inspiration?
Open example

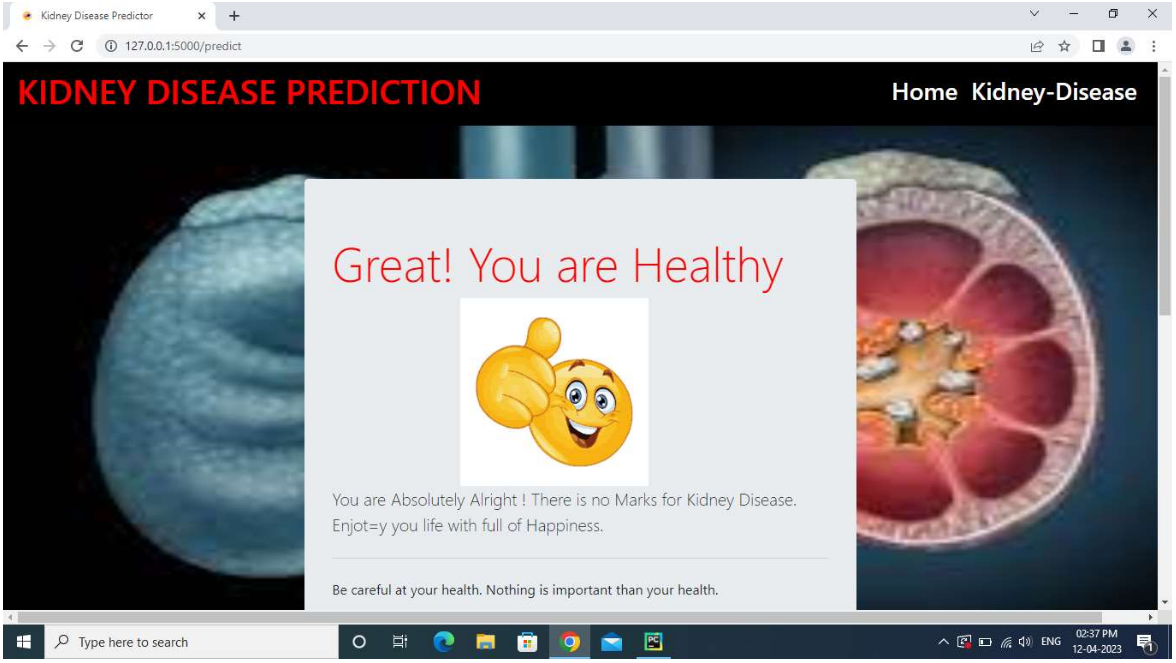
IDEATION & BRAINSTORMING MAP



3.RESULT



51	0	0
0	1	0
0	0	121
27	0.8	3.7
8300	0	0
0	0	0

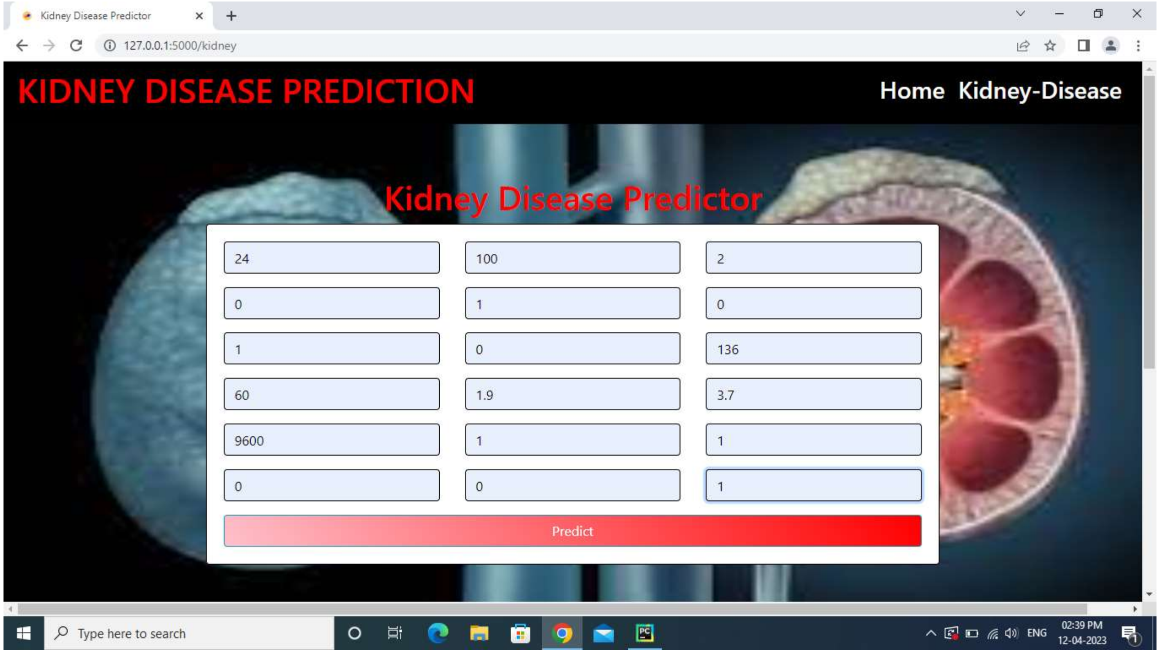


KIDNEY DISEASE PREDICTION

Great! You are Healthy

You are Absolutely Alright ! There is no Marks for Kidney Disease. Enjot=y you life with full of Happiness.

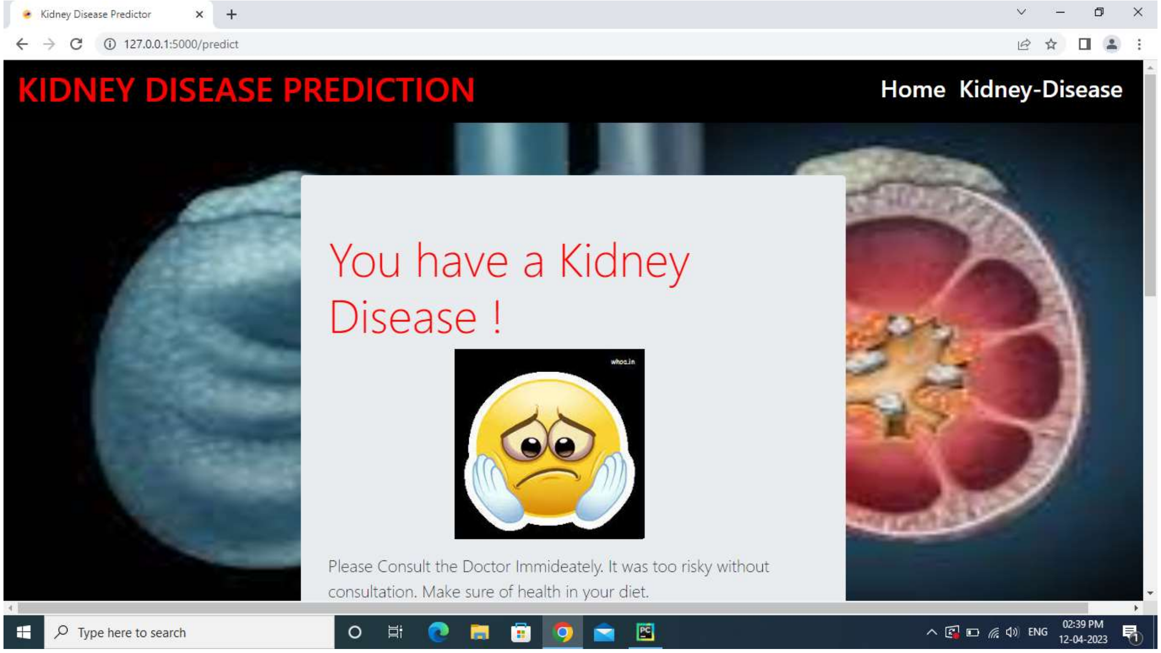
Be careful at your health. Nothing is important than your health.



KIDNEY DISEASE PREDICTION

Kidney Disease Predictor

24	100	2
0	1	0
1	0	136
60	1.9	3.7
9600	1	1
0	0	1



You have a Kidney Disease !

Please Consult the Doctor Immideately. It was too risky without consultation. Make sure of health in your diet.

4.ADVANTAGES AND DISADVANTAGES

ADVANTAGES

Early detection: Predictive models can help identify people at high risk of CKD, allowing for earlier interventions and treatments to prevent or slow disease progression.

Improved outcomes: Early detection and intervention can improve outcomes, including delaying the need for dialysis or kidney transplant and reducing mortality.

Cost-effective: Early detection and intervention can also be cost-effective by reducing the need for expensive treatments and hospitalizations.

Personalized care: Predictive models can help healthcare providers tailor care to individual patients based on their risk factors and medical history.

DISADVANTAGES

False positives: Predictive models may identify people as being at high risk of CKD when they do not actually have the disease, leading to unnecessary testing and treatments.

Privacy concerns: Collecting and analyzing patient data for predictive models raises concerns about privacy and the potential for misuse of sensitive health information.

Limited accuracy: Predictive models may not accurately predict all cases of CKD or may miss some cases, leading to missed opportunities for early intervention.

Limited generalizability: Predictive models may not be generalizable to all populations or settings, limiting their usefulness in some contexts.

5.APPLICATIONS

Chronic kidney disease (CKD) prediction has several applications in healthcare. Here are some of them:

Screening and diagnosis: Predictive models can be used to screen individuals for CKD and to aid in the diagnosis of the disease.

Risk stratification: Predictive models can help stratify patients into high- or low-risk categories based on their likelihood of developing CKD or experiencing disease progression. This can inform treatment decisions and interventions.

Disease management: Predictive models can be used to monitor patients with CKD over time and to adjust treatment plans as necessary.

Public health: Predictive models can be used to estimate the burden of CKD on a population level and to identify high-risk groups for targeted interventions.

Clinical trials: Predictive models can be used to identify patients who are most likely to benefit from a particular treatment in clinical trials, improving the efficiency of the trial and reducing the number of participants needed.

Overall, CKD prediction models have the potential to improve the management and outcomes of patients with CKD and to inform public health strategies to reduce the burden of the disease.

6.CONCLUSION

Chronic kidney disease (CKD) prediction models have the potential to improve healthcare outcomes for patients with CKD by identifying individuals at high risk of the disease, enabling early intervention and personalized care. CKD prediction models can be used for screening and diagnosis, risk stratification, disease management, public health, and clinical trials. However, there are also potential drawbacks to using predictive models, such as false positives, privacy concerns, limited accuracy, and limited generalizability. Healthcare providers and policymakers should consider these factors when implementing CKD prediction models to ensure their benefits outweigh any potential risks or drawbacks. Overall, CKD prediction models have the potential to improve the quality of care for patients with CKD and reduce the burden of the disease on public health.

7.FUTURE SCOPE

The future scope of chronic kidney disease (CKD) prediction is promising. Here are some potential areas of development and advancement:Integration with electronic health records (EHRs): As healthcare systems increasingly adopt EHRs, CKD prediction models can be integrated with EHRs to provide real-time risk assessments and facilitate targeted interventions.Machine learning and artificial intelligence: Machine learning and artificial intelligence techniques can be used to develop more accurate and sophisticated CKD prediction models, incorporating multiple data sources and risk factors.

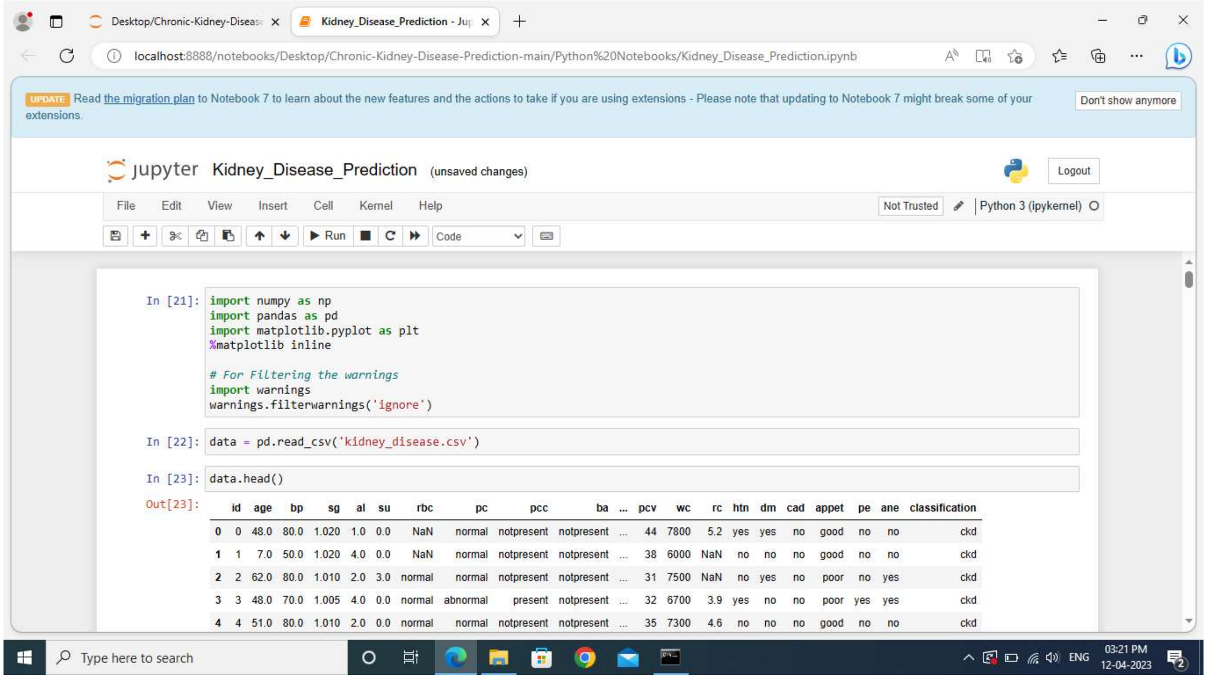
Personalized medicine: CKD prediction models can be used to provide personalized treatment plans based on individual patient risk profiles, incorporating patient preferences and goals.

Telemedicine and remote monitoring: CKD prediction models can be integrated with telemedicine and remote monitoring technologies to enable early detection and intervention, especially in underserved or remote areas.

Public health strategies: CKD prediction models can inform public health strategies for preventing and managing CKD, identifying high-risk populations and developing targeted interventions.

Overall, the future of CKD prediction holds great potential for improving healthcare outcomes and reducing the burden of the disease on public health. As technology and data continue to advance, CKD prediction models will become more sophisticated and accurate, enabling personalized and effective care for patients with CKD.

8.APPENDIX


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
Don't show anymore
jupyter Kidney_Disease_Prediction (unsaved changes)
Logout
File Edit View Insert Cell Kernel Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Type here to search
ENG
03:21 PM
12-04-2023

UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.

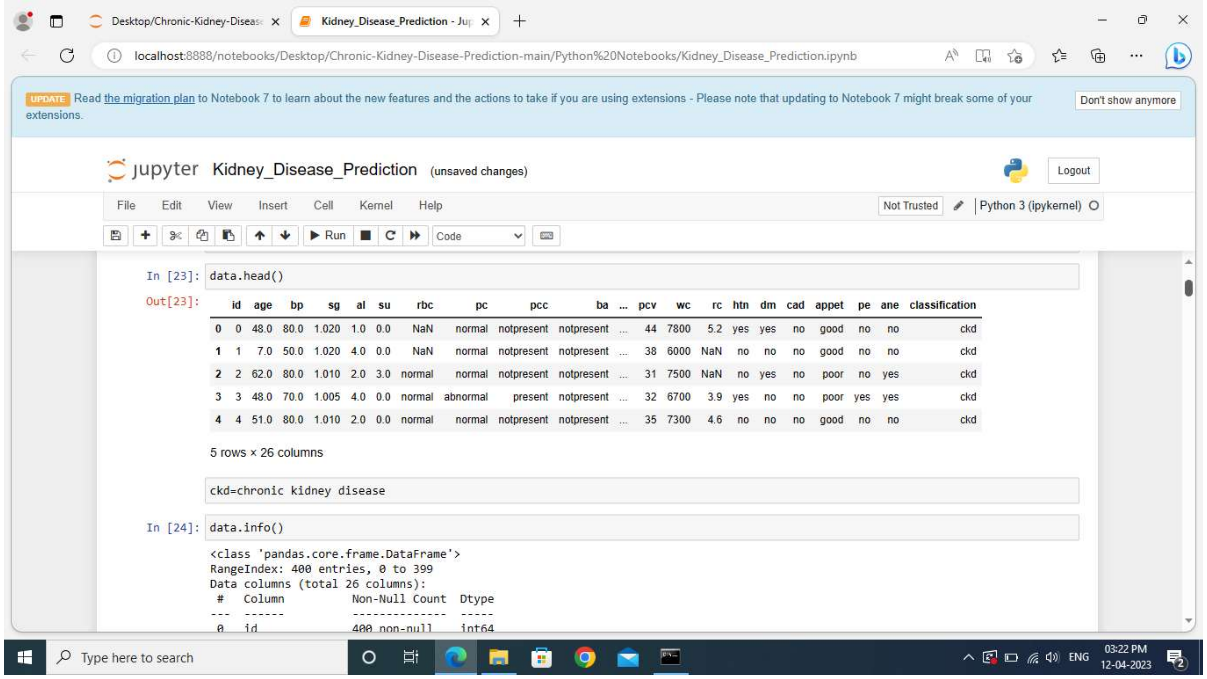
In [21]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline

# For Filtering the warnings
import warnings
warnings.filterwarnings('ignore')
```

In [22]:
```
data = pd.read_csv('kidney_disease.csv')
```

In [23]:
```
data.head()
```

Out[23]:

	id	age	bp	sg	al	su	rbc	pc	pcc	ba	...	pcv	wc	rc	htn	dm	cad	appet	pe	ane	classification
0	0	48.0	80.0	1.020	1.0	0.0	NaN	normal	notpresent	notpresent	...	44	7800	5.2	yes	yes	no	good	no	no	ckd
1	1	7.0	50.0	1.020	4.0	0.0	NaN	normal	notpresent	notpresent	...	38	6000	NaN	no	no	no	good	no	no	ckd
2	2	62.0	80.0	1.010	2.0	3.0	normal	normal	notpresent	notpresent	...	31	7500	NaN	no	yes	no	poor	no	yes	ckd
3	3	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent	...	32	6700	3.9	yes	no	no	poor	yes	yes	ckd
4	4	51.0	80.0	1.010	2.0	0.0	normal	normal	notpresent	notpresent	...	35	7300	4.6	no	no	no	good	no	no	ckd


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
Don't show anymore
jupyter Kidney_Disease_Prediction (unsaved changes)
Logout
File Edit View Insert Cell Kernel Help
Not Trusted
Python 3 (ipykernel)
Run
Code
Type here to search
ENG
03:22 PM
12-04-2023

UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.

In [23]:
```
data.head()
```

Out[23]:

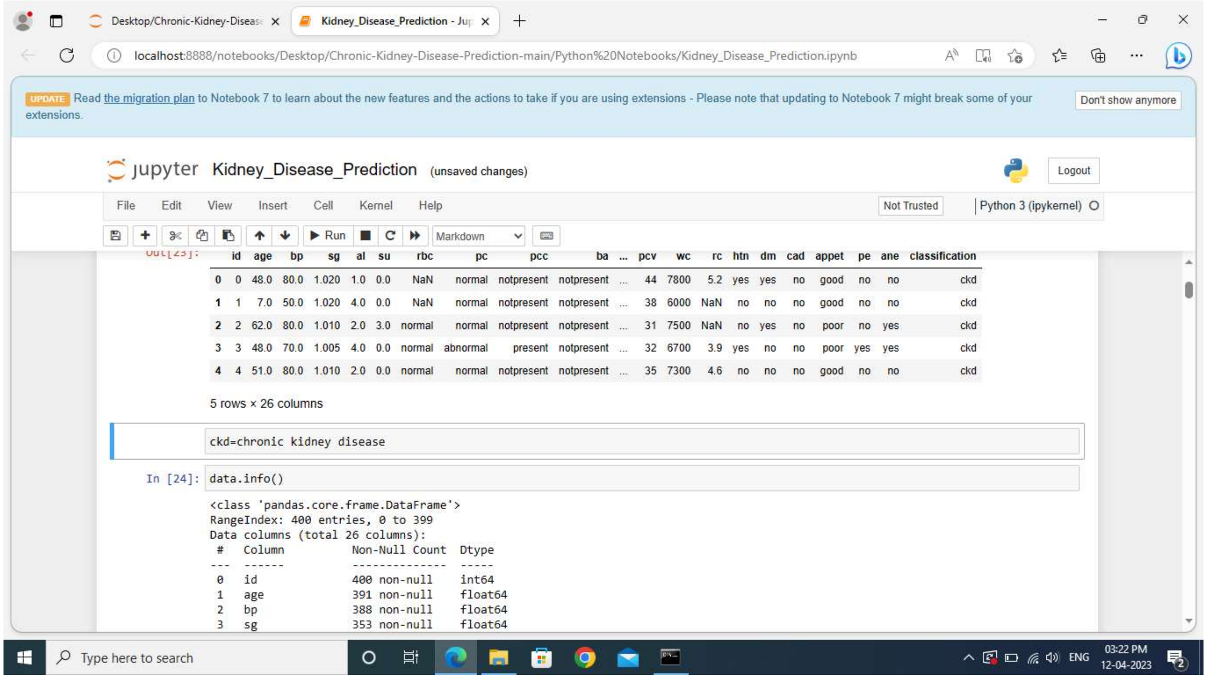
	id	age	bp	sg	al	su	rbc	pc	pcc	ba	...	pcv	wc	rc	htn	dm	cad	appet	pe	ane	classification
0	0	48.0	80.0	1.020	1.0	0.0	NaN	normal	notpresent	notpresent	...	44	7800	5.2	yes	yes	no	good	no	no	ckd
1	1	7.0	50.0	1.020	4.0	0.0	NaN	normal	notpresent	notpresent	...	38	6000	NaN	no	no	no	good	no	no	ckd
2	2	62.0	80.0	1.010	2.0	3.0	normal	normal	notpresent	notpresent	...	31	7500	NaN	no	yes	no	poor	no	yes	ckd
3	3	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent	...	32	6700	3.9	yes	no	no	poor	yes	yes	ckd
4	4	51.0	80.0	1.010	2.0	0.0	normal	normal	notpresent	notpresent	...	35	7300	4.6	no	no	no	good	no	no	ckd

5 rows × 26 columns

```
ckd=chronic kidney disease
```

In [24]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 400 entries, 0 to 399
Data columns (total 26 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   id              400 non-null    int64
```


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
jupyter
File
Edit
View
Insert
Cell
Kernel
Help
Not Trusted
Python 3 (ipykernel)
Logout
Run
Markdown
Type here to search
ENG
03:22 PM
12-04-2023

UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. Don't show anymore

Kidney_Disease_Prediction (unsaved changes)

Out[23]:

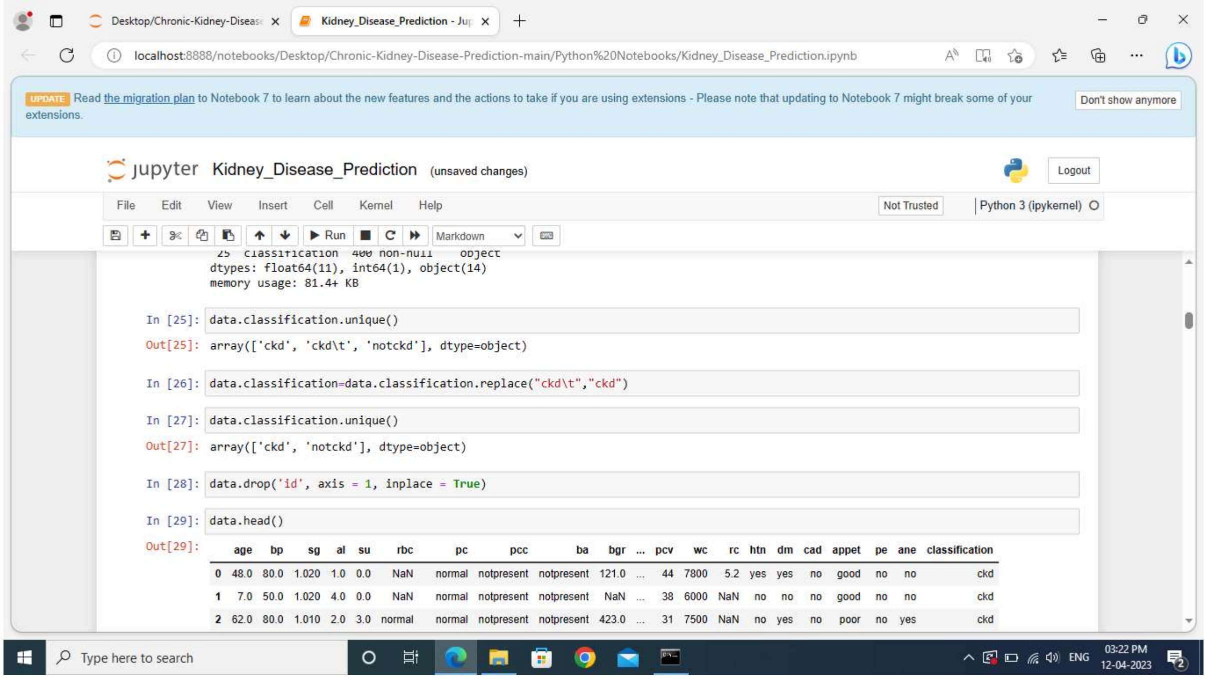
	id	age	bp	sg	al	su	rbc	pc	pcc	ba	...	pcv	wc	rc	htn	dm	cad	appet	pe	ane	classification
0	0	48.0	80.0	1.020	1.0	0.0	NaN	normal	notpresent	notpresent	...	44	7800	5.2	yes	yes	no	good	no	no	ckd
1	1	7.0	50.0	1.020	4.0	0.0	NaN	normal	notpresent	notpresent	...	38	6000	NaN	no	no	no	good	no	no	ckd
2	2	62.0	80.0	1.010	2.0	3.0	normal	normal	notpresent	notpresent	...	31	7500	NaN	no	yes	no	poor	no	yes	ckd
3	3	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent	...	32	6700	3.9	yes	no	no	poor	yes	yes	ckd
4	4	51.0	80.0	1.010	2.0	0.0	normal	normal	notpresent	notpresent	...	35	7300	4.6	no	no	no	good	no	no	ckd

5 rows × 26 columns

```
ckd=chronic kidney disease
```

In [24]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 400 entries, 0 to 399
Data columns (total 26 columns):
 #   Column          Non-Null Count  Dtype  
---  ------          --------------  -----  
 0   id              400 non-null    int64  
 1   age             391 non-null    float64
 2   bp              388 non-null    float64
 3   sg              353 non-null    float64
```


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
jupyter
File
Edit
View
Insert
Cell
Kernel
Help
Not Trusted
Python 3 (ipykernel)
Logout
Run
Markdown
Type here to search
ENG
03:22 PM
12-04-2023

UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions. Don't show anymore

Kidney_Disease_Prediction (unsaved changes)

```
 25  classification  400 non-null    object 
dtypes: float64(11), int64(1), object(14)
memory usage: 81.4+ KB
```

In [25]:
```
data.classification.unique()
```

Out[25]:
```
array(['ckd', 'ckd\t', 'notckd'], dtype=object)
```

In [26]:
```
data.classification=data.classification.replace("ckd\t","ckd")
```

In [27]:
```
data.classification.unique()
```

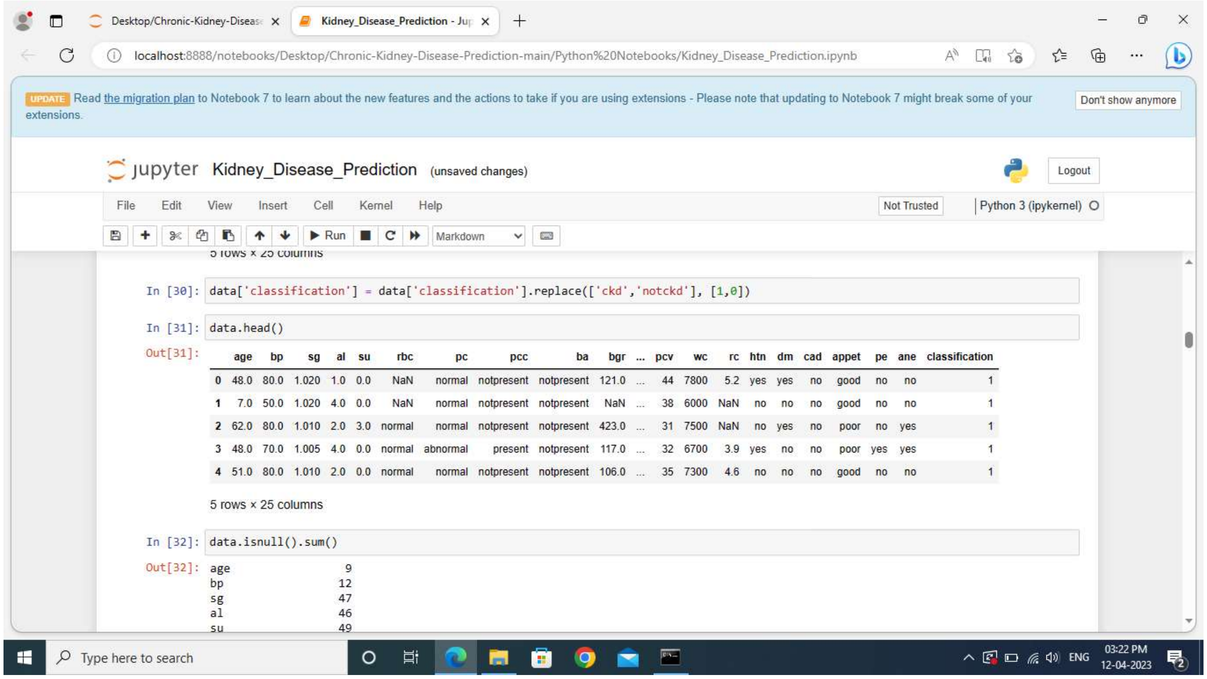
Out[27]:
```
array(['ckd', 'notckd'], dtype=object)
```

In [28]:
```
data.drop('id', axis = 1, inplace = True)
```

In [29]:
```
data.head()
```

Out[29]:

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	pcv	wc	rc	htn	dm	cad	appet	pe	ane	classification
0	48.0	80.0	1.020	1.0	0.0	NaN	normal	notpresent	notpresent	121.0	...	44	7800	5.2	yes	yes	no	good	no	no	ckd
1	7.0	50.0	1.020	4.0	0.0	NaN	normal	notpresent	notpresent	NaN	...	38	6000	NaN	no	no	no	good	no	no	ckd
2	62.0	80.0	1.010	2.0	3.0	normal	normal	notpresent	notpresent	423.0	...	31	7500	NaN	no	yes	no	poor	no	yes	ckd


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.
Don't show anymore
jupyter Kidney_Disease_Prediction (unsaved changes)
Logout
File Edit View Insert Cell Kernel Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Type here to search
ENG
03:22 PM
12-04-2023

5 rows × 25 columns

```
In [30]: data['classification'] = data['classification'].replace(['ckd','notckd'], [1,0])
```

```
In [31]: data.head()
```

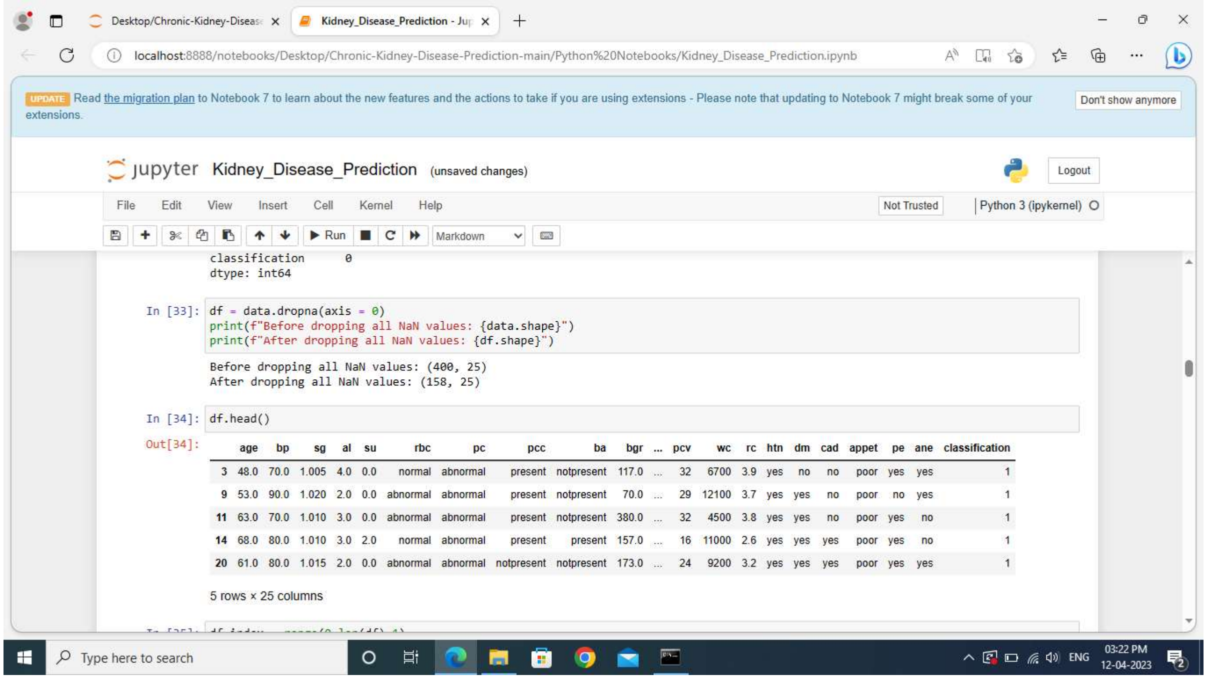
Out[31]:

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	pcv	wc	rc	htn	dm	cad	appet	pe	ane	classification
0	48.0	80.0	1.020	1.0	0.0	NaN	normal	notpresent	notpresent	121.0	...	44	7800	5.2	yes	yes	no	good	no	no	1
1	7.0	50.0	1.020	4.0	0.0	NaN	normal	notpresent	notpresent	NaN	...	38	6000	NaN	no	no	no	good	no	no	1
2	62.0	80.0	1.010	2.0	3.0	normal	normal	notpresent	notpresent	423.0	...	31	7500	NaN	no	yes	no	poor	no	yes	1
3	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent	117.0	...	32	6700	3.9	yes	no	no	poor	yes	yes	1
4	51.0	80.0	1.010	2.0	0.0	normal	normal	notpresent	notpresent	106.0	...	35	7300	4.6	no	no	no	good	no	no	1

5 rows × 25 columns

```
In [32]: data.isnull().sum()
```

```
Out[32]: age              9
         bp              12
         sg              47
         al              46
         su              49
```


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.
Don't show anymore
jupyter Kidney_Disease_Prediction (unsaved changes)
Logout
File Edit View Insert Cell Kernel Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Type here to search
ENG
03:22 PM
12-04-2023

```
classification      0
dtype: int64
```

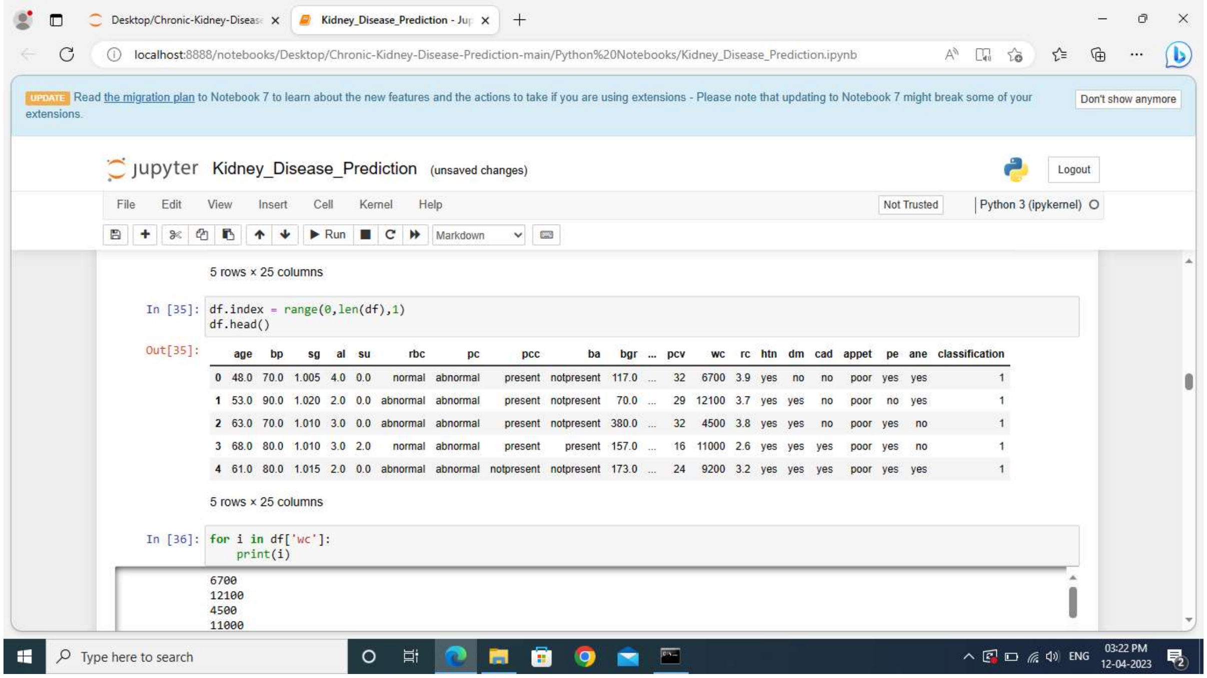
```
In [33]: df = data.dropna(axis = 0)
         print(f"Before dropping all NaN values: {data.shape}")
         print(f"After dropping all NaN values: {df.shape}")

         Before dropping all NaN values: (400, 25)
         After dropping all NaN values: (158, 25)
```

```
In [34]: df.head()
```

Out[34]:

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	pcv	wc	rc	htn	dm	cad	appet	pe	ane	classification
3	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent	117.0	...	32	6700	3.9	yes	no	no	poor	yes	yes	1
9	53.0	90.0	1.020	2.0	0.0	abnormal	abnormal	present	notpresent	70.0	...	29	12100	3.7	yes	yes	no	poor	no	yes	1
11	63.0	70.0	1.010	3.0	0.0	abnormal	abnormal	present	notpresent	380.0	...	32	4500	3.8	yes	yes	no	poor	yes	no	1
14	68.0	80.0	1.010	3.0	2.0	normal	abnormal	present	present	157.0	...	16	11000	2.6	yes	yes	yes	poor	yes	no	1
20	61.0	80.0	1.015	2.0	0.0	abnormal	abnormal	notpresent	notpresent	173.0	...	24	9200	3.2	yes	yes	yes	poor	yes	yes	1

5 rows × 25 columns


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
UPDATE
Don't show anymore
jupyter Kidney_Disease_Prediction (unsaved changes)
Logout
File Edit View Insert Cell Kernel Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Type here to search
ENG
03:22 PM
12-04-2023

Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.

5 rows × 25 columns

In [35]:
```
df.index = range(0,len(df),1)
df.head()
```

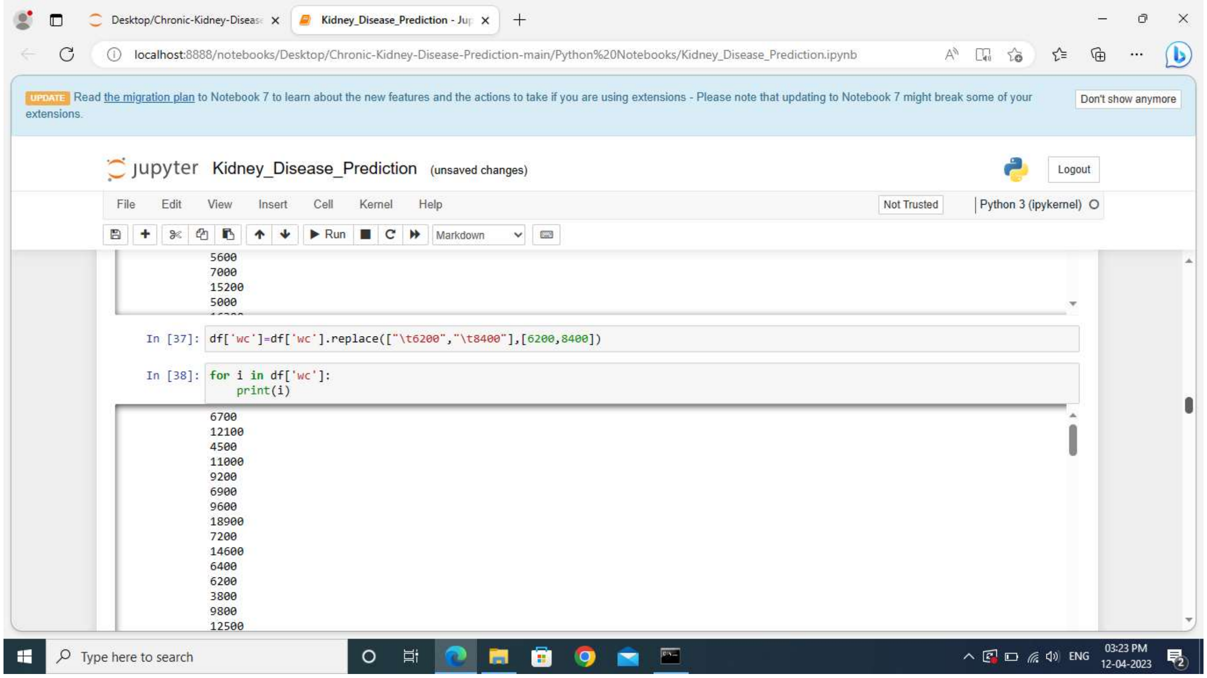
Out[35]:

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	pcv	wc	rc	htn	dm	cad	appet	pe	ane	classification
0	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent	117.0	...	32	6700	3.9	yes	no	no	poor	yes	yes	1
1	53.0	90.0	1.020	2.0	0.0	abnormal	abnormal	present	notpresent	70.0	...	29	12100	3.7	yes	yes	no	poor	no	yes	1
2	63.0	70.0	1.010	3.0	0.0	abnormal	abnormal	present	notpresent	380.0	...	32	4500	3.8	yes	yes	no	poor	yes	no	1
3	68.0	80.0	1.010	3.0	2.0	normal	abnormal	present	present	157.0	...	16	11000	2.6	yes	yes	yes	poor	yes	no	1
4	61.0	80.0	1.015	2.0	0.0	abnormal	abnormal	notpresent	notpresent	173.0	...	24	9200	3.2	yes	yes	yes	poor	yes	yes	1

5 rows × 25 columns

In [36]:
```
for i in df['wc']:
    print(i)
```

```
6700
12100
4500
11000
```


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
UPDATE
Don't show anymore
jupyter Kidney_Disease_Prediction (unsaved changes)
Logout
File Edit View Insert Cell Kernel Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Type here to search
ENG
03:23 PM
12-04-2023

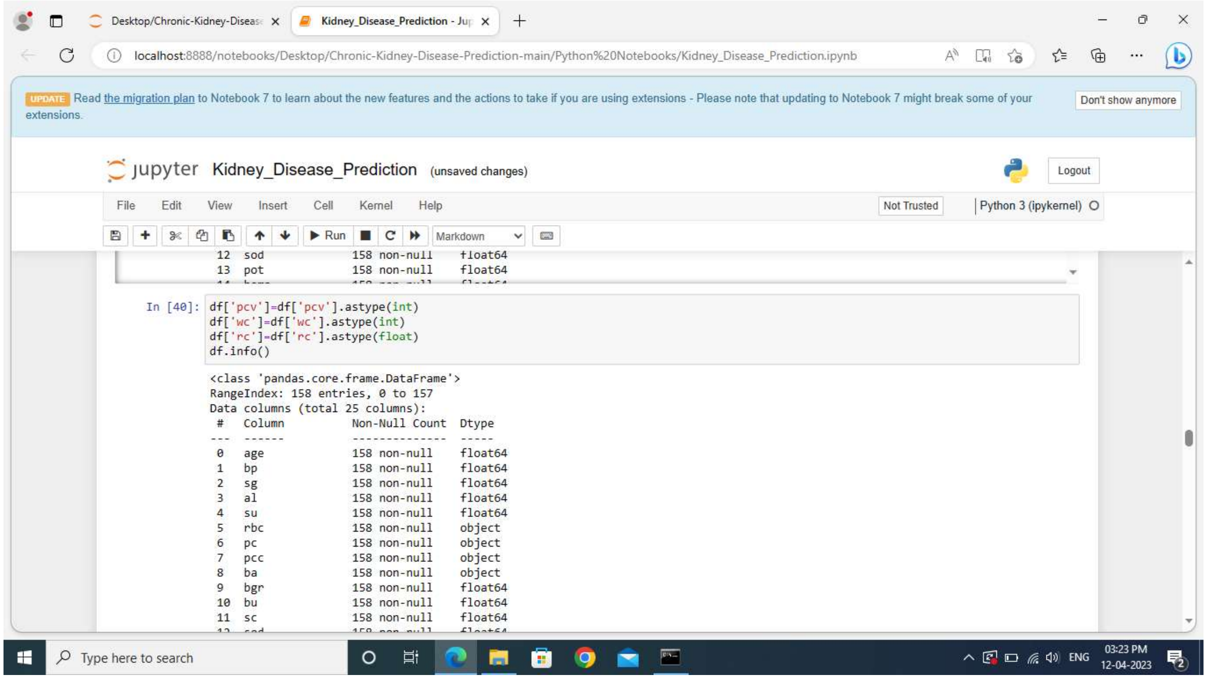
Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.

```
5600
7000
15200
5000
```

In [37]:
```
df['wc']=df['wc'].replace(["\t6200","\t8400"],[6200,8400])
```

In [38]:
```
for i in df['wc']:
    print(i)
```

```
6700
12100
4500
11000
9200
6900
9600
18900
7200
14600
6400
6200
3800
9800
12500
```


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
Don't show anymore
jupyter Kidney_Disease_Prediction (unsaved changes)
Logout
File Edit View Insert Cell Kernel Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Type here to search
ENG
03:23 PM
12-04-2023

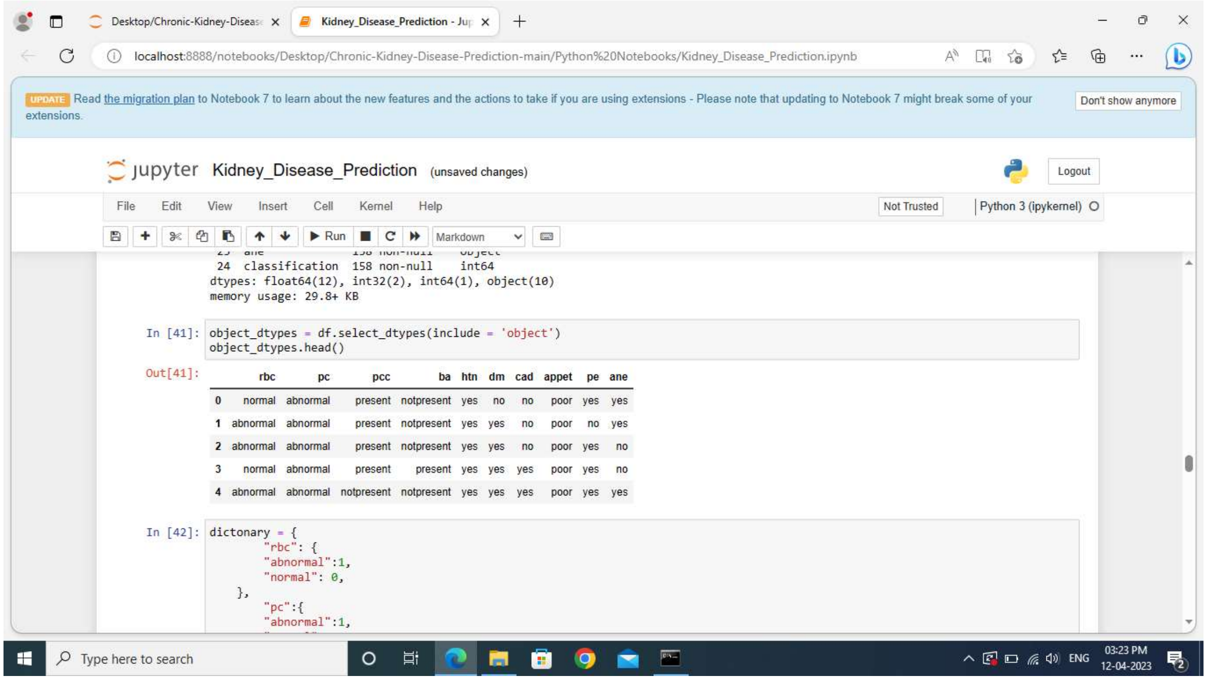
UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.

```
 12  sod           158 non-null    float64
 13  pot           158 non-null    float64
```

In [40]:

```
df['pcv']=df['pcv'].astype(int)
df['wc']=df['wc'].astype(int)
df['rc']=df['rc'].astype(float)
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 158 entries, 0 to 157
Data columns (total 25 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   age            158 non-null    float64
 1   bp             158 non-null    float64
 2   sg             158 non-null    float64
 3   al             158 non-null    float64
 4   su             158 non-null    float64
 5   rbc            158 non-null    object
 6   pc             158 non-null    object
 7   pcc            158 non-null    object
 8   ba             158 non-null    object
 9   bgr            158 non-null    float64
 10  bu             158 non-null    float64
 11  sc             158 non-null    float64
```


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
Don't show anymore
jupyter Kidney_Disease_Prediction (unsaved changes)
Logout
File Edit View Insert Cell Kernel Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Type here to search
ENG
03:23 PM
12-04-2023

UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.

```
 24  classification  158 non-null    int64
dtypes: float64(12), int32(2), int64(1), object(10)
memory usage: 29.8+ KB
```

In [41]:

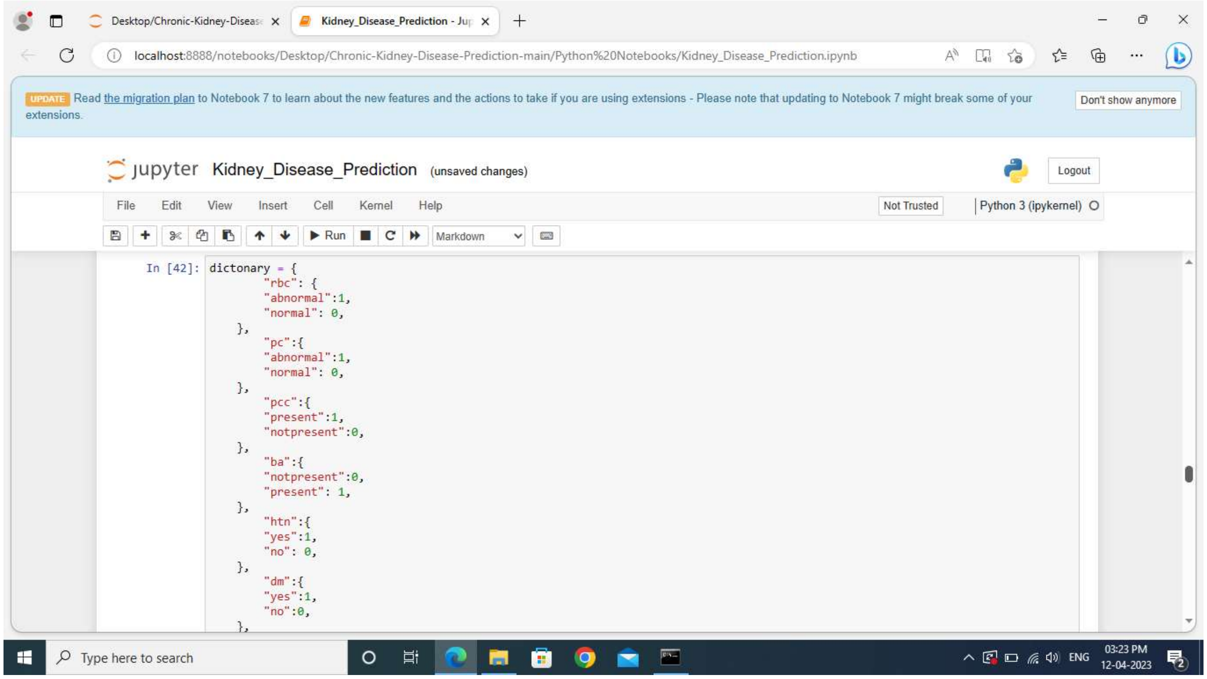
```
object_dtypes = df.select_dtypes(include = 'object')
object_dtypes.head()
```

Out[41]:

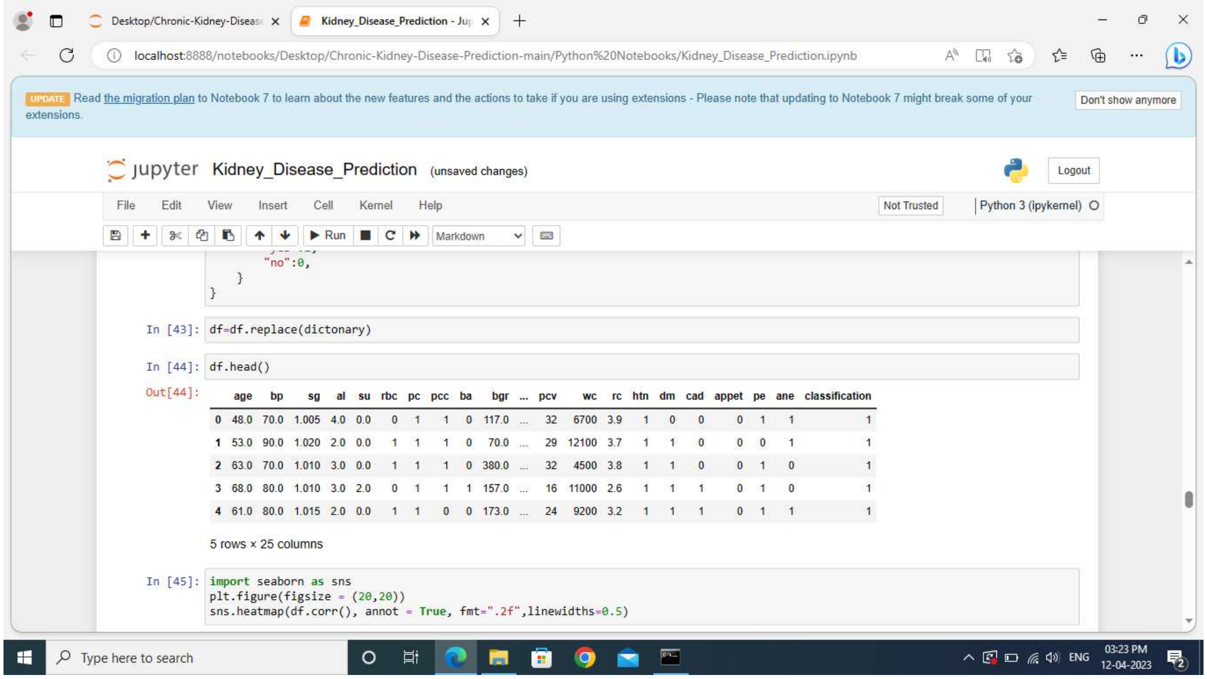
	rbc	pc	pcc	ba	htn	dm	cad	appet	pe	ane
0	normal	abnormal	present	notpresent	yes	no	no	poor	yes	yes
1	abnormal	abnormal	present	notpresent	yes	yes	no	poor	no	yes
2	abnormal	abnormal	present	notpresent	yes	yes	no	poor	yes	no
3	normal	abnormal	present	present	yes	yes	yes	poor	yes	no
4	abnormal	abnormal	notpresent	notpresent	yes	yes	yes	poor	yes	yes

In [42]:

```
dictonary = {
        "rbc": {
        "abnormal":1,
        "normal": 0,
    },
        "pc":{
        "abnormal":1,
```


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.
Don't show anymore
jupyter Kidney_Disease_Prediction (unsaved changes)
Logout
File Edit View Insert Cell Kernel Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Type here to search
ENG
03:23 PM
12-04-2023

```
In [42]: dictonary = {
                "rbc": {
                "abnormal":1,
                "normal": 0,
            },
                "pc":{
                "abnormal":1,
                "normal": 0,
            },
                "pcc":{
                "present":1,
                "notpresent":0,
            },
                "ba":{
                "notpresent":0,
                "present": 1,
            },
                "htn":{
                "yes":1,
                "no": 0,
            },
                "dm":{
                "yes":1,
                "no":0,
            },
```

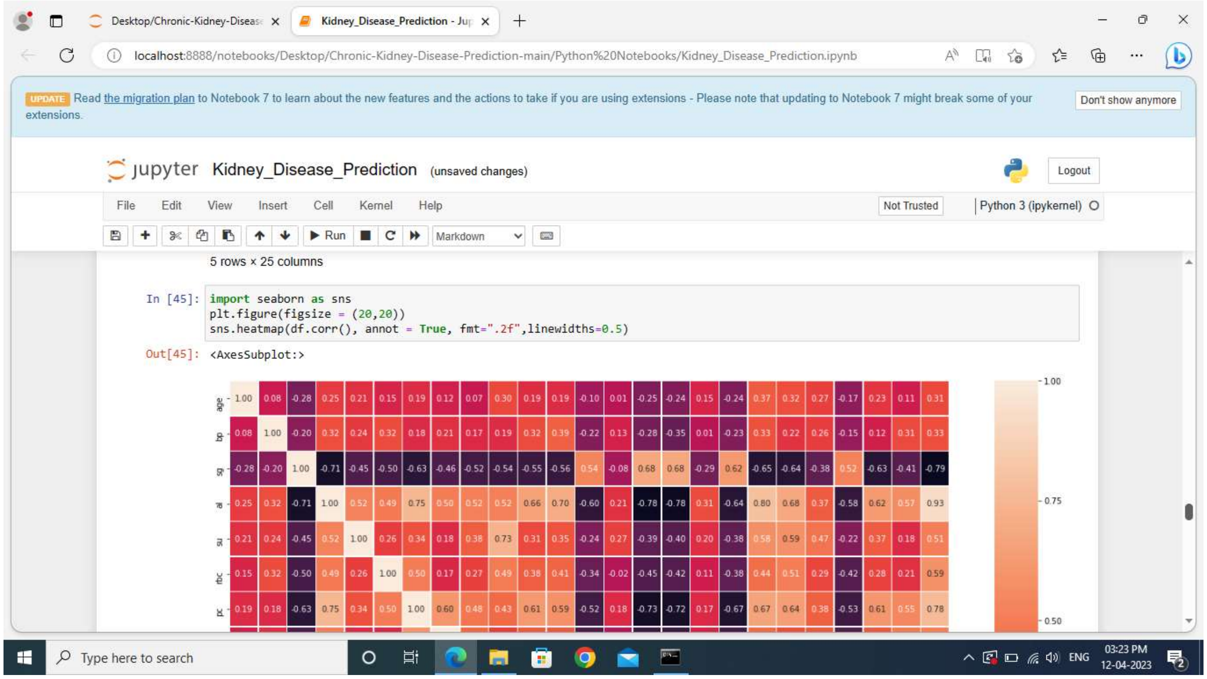

Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.
Don't show anymore
jupyter Kidney_Disease_Prediction (unsaved changes)
Logout
File Edit View Insert Cell Kernel Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Type here to search
ENG
03:23 PM
12-04-2023

```
                "no":0,
            }
        }

In [43]: df=df.replace(dictonary)

In [44]: df.head()

Out[44]:
```

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	pcv	wc	rc	htn	dm	cad	appet	pe	ane	classification
0	48.0	70.0	1.005	4.0	0.0	0	1	1	0	117.0	...	32	6700	3.9	1	0	0	0	1	1	1
1	53.0	90.0	1.020	2.0	0.0	1	1	1	0	70.0	...	29	12100	3.7	1	1	0	0	0	1	1
2	63.0	70.0	1.010	3.0	0.0	1	1	1	0	380.0	...	32	4500	3.8	1	1	0	0	1	0	1
3	68.0	80.0	1.010	3.0	2.0	0	1	1	1	157.0	...	16	11000	2.6	1	1	1	0	1	0	1
4	61.0	80.0	1.015	2.0	0.0	1	1	0	0	173.0	...	24	9200	3.2	1	1	1	0	1	1	1

5 rows × 25 columns

```
In [45]: import seaborn as sns
         plt.figure(figsize = (20,20))
         sns.heatmap(df.corr(), annot = True, fmt=".2f",linewidths=0.5)
```


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
Don't show anymore
jupyter Kidney_Disease_Prediction (unsaved changes)
Logout
File
Edit
View
Insert
Cell
Kernel
Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Type here to search
ENG
03:23 PM
12-04-2023

UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.

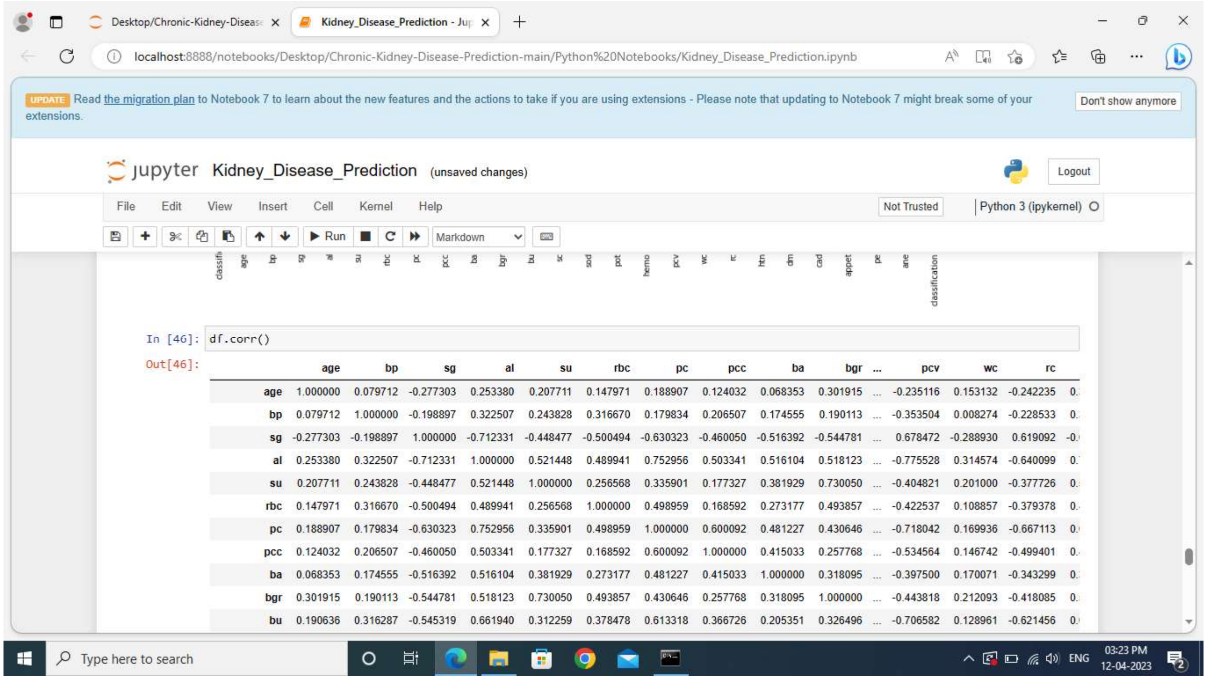
5 rows × 25 columns

In [45]:

```
import seaborn as sns
plt.figure(figsize = (20,20))
sns.heatmap(df.corr(), annot = True, fmt=".2f",linewidths=0.5)
```

Out[45]: `<AxesSubplot:>`

age	1.00	0.08	-0.28	0.25	0.21	0.15	0.19	0.12	0.07	0.30	0.19	0.19	-0.10	0.01	-0.25	-0.24	0.15	-0.24	0.37	0.32	0.27	-0.17	0.23	0.11	0.31
bp	0.08	1.00	-0.20	0.32	0.24	0.32	0.18	0.21	0.17	0.19	0.32	0.39	-0.22	0.13	-0.28	-0.35	0.01	-0.23	0.33	0.22	0.26	-0.15	0.12	0.31	0.33
sg	-0.28	-0.20	1.00	-0.71	-0.45	-0.50	-0.63	-0.46	-0.52	-0.54	-0.55	-0.56	0.54	-0.08	0.68	0.68	-0.29	0.62	-0.65	-0.64	-0.38	0.52	-0.63	-0.41	-0.79
al	0.25	0.32	-0.71	1.00	0.52	0.49	0.75	0.50	0.52	0.52	0.66	0.70	-0.60	0.21	-0.78	-0.78	0.31	-0.64	0.80	0.68	0.37	-0.58	0.62	0.57	0.93
su	0.21	0.24	-0.45	0.52	1.00	0.26	0.34	0.18	0.38	0.73	0.31	0.35	-0.24	0.27	-0.39	-0.40	0.20	-0.38	0.58	0.59	0.47	-0.22	0.37	0.18	0.51
rbc	0.15	0.32	-0.50	0.49	0.26	1.00	0.50	0.17	0.27	0.49	0.38	0.41	-0.34	-0.02	-0.45	-0.42	0.11	-0.38	0.44	0.51	0.29	-0.42	0.28	0.21	0.59
pc	0.19	0.18	-0.63	0.75	0.34	0.50	1.00	0.60	0.48	0.43	0.61	0.59	-0.52	0.18	-0.73	-0.72	0.17	-0.67	0.67	0.64	0.38	-0.53	0.61	0.55	0.78


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
Don't show anymore
jupyter Kidney_Disease_Prediction (unsaved changes)
Logout
File
Edit
View
Insert
Cell
Kernel
Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
classifi
age
bp
sg
al
su
rbc
pc
pcc
ba
bgr
bu
sc
sod
pot
hemo
pcv
wc
rc
htn
dm
cad
appet
pe
ane
classification
Type here to search
ENG
03:23 PM
12-04-2023

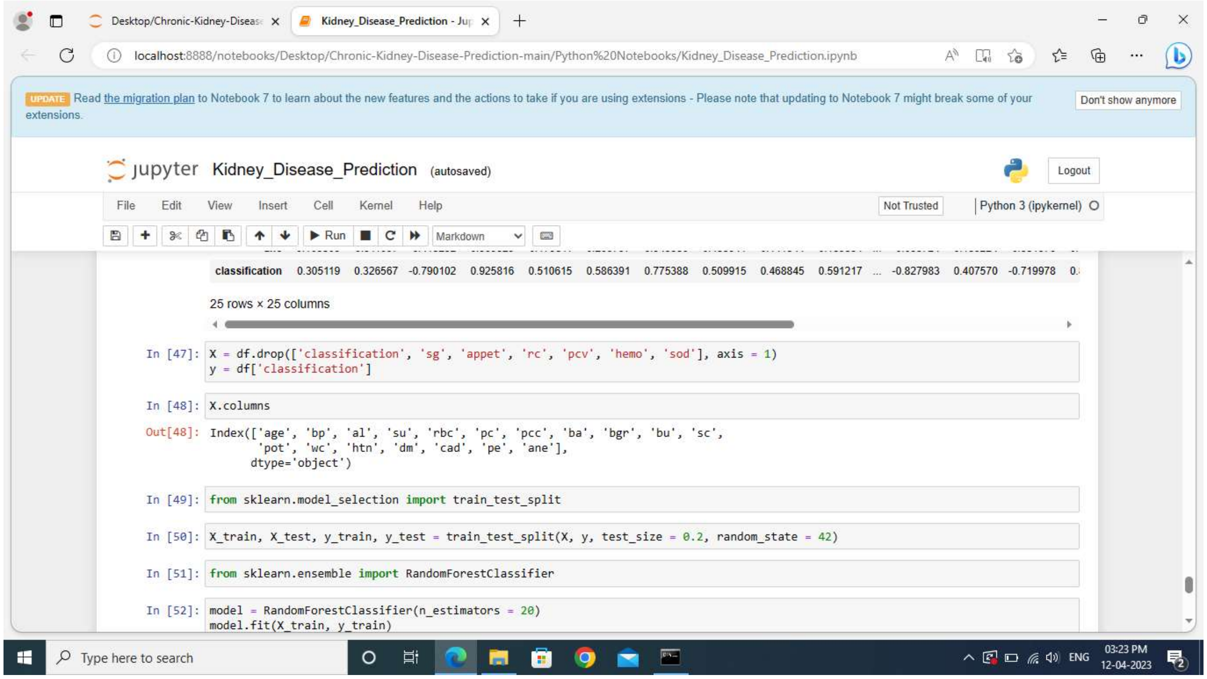
UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.

In [46]:

```
df.corr()
```

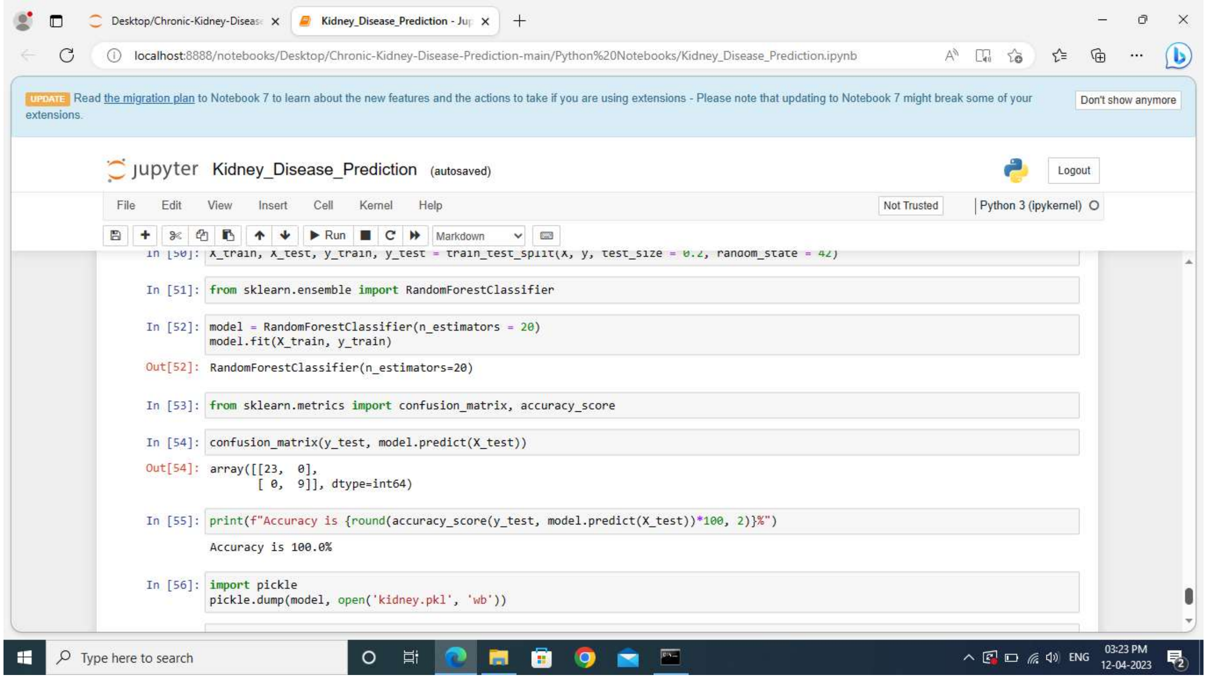
Out[46]:

	age	bp	sg	al	su	rbc	pc	pcc	ba	bgr	...	pcv	wc	rc	
age	1.000000	0.079712	-0.277303	0.253380	0.207711	0.147971	0.188907	0.124032	0.068353	0.301915	...	-0.235116	0.153132	-0.242235	0.
bp	0.079712	1.000000	-0.198897	0.322507	0.243828	0.316670	0.179834	0.206507	0.174555	0.190113	...	-0.353504	0.008274	-0.228533	0.
sg	-0.277303	-0.198897	1.000000	-0.712331	-0.448477	-0.500494	-0.630323	-0.460050	-0.516392	-0.544781	...	0.678472	-0.288930	0.619092	-0.
al	0.253380	0.322507	-0.712331	1.000000	0.521448	0.489941	0.752956	0.503341	0.516104	0.518123	...	-0.775528	0.314574	-0.640099	0.
su	0.207711	0.243828	-0.448477	0.521448	1.000000	0.256568	0.335901	0.177327	0.381929	0.730050	...	-0.404821	0.201000	-0.377726	0.
rbc	0.147971	0.316670	-0.500494	0.489941	0.256568	1.000000	0.498959	0.168592	0.273177	0.493857	...	-0.422537	0.108857	-0.379378	0.
pc	0.188907	0.179834	-0.630323	0.752956	0.335901	0.498959	1.000000	0.600092	0.481227	0.430646	...	-0.718042	0.169936	-0.667113	0.
pcc	0.124032	0.206507	-0.460050	0.503341	0.177327	0.168592	0.600092	1.000000	0.415033	0.257768	...	-0.534564	0.146742	-0.499401	0.
ba	0.068353	0.174555	-0.516392	0.516104	0.381929	0.273177	0.481227	0.415033	1.000000	0.318095	...	-0.397500	0.170071	-0.343299	0.
bgr	0.301915	0.190113	-0.544781	0.518123	0.730050	0.493857	0.430646	0.257768	0.318095	1.000000	...	-0.443818	0.212093	-0.418085	0.
bu	0.190636	0.316287	-0.545319	0.661940	0.312259	0.378478	0.613318	0.366726	0.205351	0.326496	...	-0.706582	0.128961	-0.621456	0.

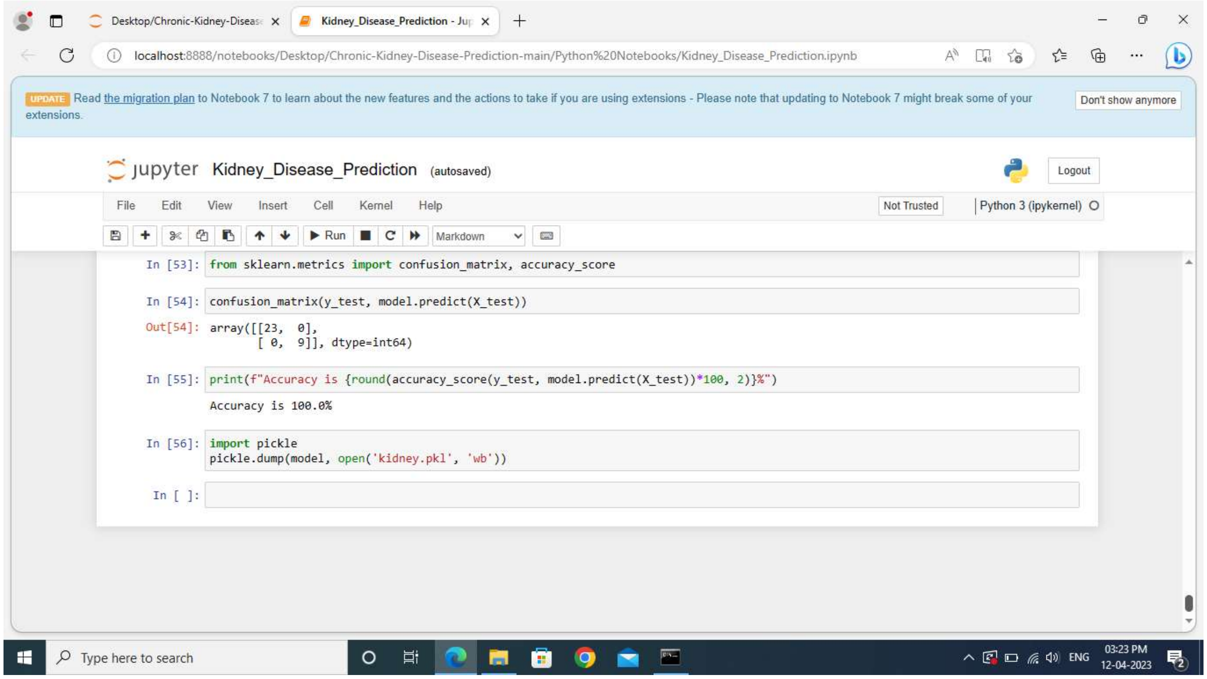

Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.
Don't show anymore
jupyter Kidney_Disease_Prediction (autosaved)
Logout
File Edit View Insert Cell Kernel Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Type here to search
ENG
03:23 PM
12-04-2023

```
classification  0.305119  0.326567  -0.790102  0.925816  0.510615  0.586391  0.775388  0.509915  0.468845  0.591217  ...  -0.827983  0.407570  -0.719978  0.

25 rows × 25 columns
```

```
In [47]: X = df.drop(['classification', 'sg', 'appet', 'rc', 'pcv', 'hemo', 'sod'], axis = 1)
         y = df['classification']

In [48]: X.columns

Out[48]: Index(['age', 'bp', 'al', 'su', 'rbc', 'pc', 'pcc', 'ba', 'bgr', 'bu', 'sc',
                'pot', 'wc', 'htn', 'dm', 'cad', 'pe', 'ane'],
               dtype='object')

In [49]: from sklearn.model_selection import train_test_split

In [50]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 42)

In [51]: from sklearn.ensemble import RandomForestClassifier

In [52]: model = RandomForestClassifier(n_estimators = 20)
         model.fit(X_train, y_train)
```


Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.
Don't show anymore
jupyter Kidney_Disease_Prediction (autosaved)
Logout
File Edit View Insert Cell Kernel Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Type here to search
ENG
03:23 PM
12-04-2023

```
In [50]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 42)

In [51]: from sklearn.ensemble import RandomForestClassifier

In [52]: model = RandomForestClassifier(n_estimators = 20)
         model.fit(X_train, y_train)

Out[52]: RandomForestClassifier(n_estimators=20)

In [53]: from sklearn.metrics import confusion_matrix, accuracy_score

In [54]: confusion_matrix(y_test, model.predict(X_test))

Out[54]: array([[23,  0],
                [ 0,  9]], dtype=int64)

In [55]: print(f"Accuracy is {round(accuracy_score(y_test, model.predict(X_test))*100, 2)}%")

         Accuracy is 100.0%

In [56]: import pickle
         pickle.dump(model, open('kidney.pkl', 'wb'))
```



Desktop/Chronic-Kidney-Disease
Kidney_Disease_Prediction - Jup
localhost:8888/notebooks/Desktop/Chronic-Kidney-Disease-Prediction-main/Python%20Notebooks/Kidney_Disease_Prediction.ipynb
UPDATE Read the migration plan to Notebook 7 to learn about the new features and the actions to take if you are using extensions - Please note that updating to Notebook 7 might break some of your extensions.
Don't show anymore
jupyter Kidney_Disease_Prediction (autosaved)
Logout
File Edit View Insert Cell Kernel Help
Not Trusted
Python 3 (ipykernel)
Run
Markdown
Type here to search
ENG
03:23 PM
12-04-2023

```
In [53]: from sklearn.metrics import confusion_matrix, accuracy_score

In [54]: confusion_matrix(y_test, model.predict(X_test))

Out[54]: array([[23,  0],
                [ 0,  9]], dtype=int64)

In [55]: print(f"Accuracy is {round(accuracy_score(y_test, model.predict(X_test))*100, 2)}%")

         Accuracy is 100.0%

In [56]: import pickle
         pickle.dump(model, open('kidney.pkl', 'wb'))

In [ ]:
```