

Doc2.pdf x localhost x CKD.pkl (edit) x Assignment-4 x Assignment-4 x Downloads/ x project - Jupyter x IBM-40609-1 x + -

localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (pykernel)

In [1]:

```
#importing libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import os
```

In [2]:

```
os.getcwd()
```

Out[2]: 'C:\\Users\\ELCOT\\Downloads'

In [5]:

```
path='C:\\Users\\ELCOT\\Downloads\\'
data=pd.read_csv(path+'chronickidneydisease.csv')
data.head(10)
```

Out[5]:

	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	5	60.0	90.0	1.015	3.0	0.0	NaN	NaN	notpresent	notpresent	...	39	7800	4.4	yes	yes	no	good	yes	no	ckd
6	6	68.0	70.0	1.010	0.0	0.0	NaN	normal	notpresent	notpresent	...	36	NaN	NaN	no	no	no	good	no	no	ckd
7	7	24.0	NaN	1.015	2.0	4.0	normal	abnormal	notpresent	notpresent	...	44	6900	5	no	yes	no	good	yes	no	ckd
8	8	52.0	100.0	1.015	3.0	0.0	normal	abnormal	present	notpresent	...	33	9600	4.0	yes	yes	no	good	no	yes	ckd

Doc2.pdf x localhost x CKD.pkl (edit) x Assignment-4 x Assignment-4 x Downloads/ x project - Jupyter x IBM-40609-1 x + -

localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (pykernel)

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	5	60.0	90.0	1.015	3.0	0.0	NaN	NaN	notpresent	notpresent	...	39	7800	4.4	yes	yes	no	good	yes	no	ckd
6	6	68.0	70.0	1.010	0.0	0.0	NaN	normal	notpresent	notpresent	...	36	NaN	NaN	no	no	no	good	no	no	ckd
7	7	24.0	NaN	1.015	2.0	4.0	normal	abnormal	notpresent	notpresent	...	44	6900	5	no	yes	no	good	yes	no	ckd
8	8	52.0	100.0	1.015	3.0	0.0	normal	abnormal	present	notpresent	...	33	9600	4.0	yes	yes	no	good	no	yes	ckd
9	9	53.0	90.0	1.020	2.0	0.0	abnormal	abnormal	present	notpresent	...	29	12100	3.7	yes	yes	no	poor	no	yes	ckd

10 rows x 26 columns

In [6]:

```
data.tail(10)
```

Out[6]:

	id	age	bp	sg	al	su	rbc	pc	pcc	ba	...	pcv	wc	rc	htn	dm	cad	appet	pe	ane	classification
390	390	52.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	52	6300	5.3	no	no	no	good	no	no	notckd
391	391	36.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	44	5800	6.3	no	no	no	good	no	no	notckd
392	392	57.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	...	46	6600	5.5	no	no	no	good	no	no	notckd
393	393	43.0	60.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	54	7400	5.4	no	no	no	good	no	no	notckd
394	394	50.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	...	45	9500	4.6	no	no	no	good	no	no	notckd
395	395	55.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	...	47	6700	4.9	no	no	no	good	no	no	notckd
396	396	42.0	70.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	54	7800	6.2	no	no	no	good	no	no	notckd
397	397	12.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	...	49	6600	5.4	no	no	no	good	no	no	notckd
398	398	17.0	60.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	51	7200	5.9	no	no	no	good	no	no	notckd
399	399	58.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	...	53	6800	6.1	no	no	no	good	no	no	notckd

localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (pykernel)

```
396 396 42.0 70.0 1.025 0.0 0.0 normal normal notpresent notpresent ... 54 7800 6.2 no no no good no no notkdd
397 397 12.0 80.0 1.020 0.0 0.0 normal normal notpresent notpresent ... 49 6600 5.4 no no no good no no notkdd
398 398 17.0 60.0 1.025 0.0 0.0 normal normal notpresent notpresent ... 51 7200 5.9 no no no good no no notkdd
399 399 58.0 80.0 1.025 0.0 0.0 normal normal notpresent notpresent ... 53 6800 6.1 no no no good no no notkdd
```

10 rows x 26 columns

In [7]: data.shape

Out[7]: (400, 26)

In [8]: data.columns=['id','age','blood\_pressure','specific\_gravity','albumin','sugar','red\_blood\_cells',  
'pus\_cell','pus\_cell\_clumps','bacteria','blood\_glucose\_random',  
'blood\_urea','serum\_creatinine','sodium','potassium','hemoglobin','packed\_cell\_volume',  
'white\_blood\_cell\_count','red\_blood\_cell\_count','hypertension','diabetesmellitus','coronary\_artery\_disease',  
'appetite','pedal\_edema','anemia','class']

In [9]: data.columns

Out[9]: Index(['id', 'age', 'blood\_pressure', 'specific\_gravity', 'albumin', 'sugar',  
'red\_blood\_cells', 'pus\_cell', 'pus\_cell\_clumps', 'bacteria',  
'blood\_glucose\_random', 'blood\_urea', 'serum\_creatinine', 'sodium',  
'potassium', 'hemoglobin', 'packed\_cell\_volume',  
'white\_blood\_cell\_count', 'red\_blood\_cell\_count', 'hypertension',  
'diabetesmellitus', 'coronary\_artery\_disease', 'appetite',  
'pedal\_edema', 'anemia', 'class'],  
dtype='object')

In [10]: data.info()

localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (pykernel)

In [10]: 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   blood_pressure        388 non-null   float64
3   specific_gravity      353 non-null   float64
4   albumin               354 non-null   float64
5   sugar                 351 non-null   float64
6   red_blood_cells       248 non-null   object
7   pus_cell              335 non-null   object
8   pus_cell_clumps       396 non-null   object
9   bacteria              396 non-null   object
10  blood_glucose_random   356 non-null   float64
11  blood_urea            381 non-null   float64
12  serum_creatinine      383 non-null   float64
13  sodium               313 non-null   float64
14  potassium             312 non-null   float64
15  hemoglobin            348 non-null   float64
16  packed_cell_volume     330 non-null   object
17  white_blood_cell_count 295 non-null   object
18  red_blood_cell_count   270 non-null   object
19  hypertension          398 non-null   object
20  diabetesmellitus       398 non-null   object
21  coronary_artery_disease 398 non-null   object
22  appetite              399 non-null   object
23  pedal_edema           399 non-null   object
24  anemia                399 non-null   object
25  class                 400 non-null   object
dtypes: float64(11), int64(1), object(14)
```

localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help

Code

```
22 appetite 399 non-null object
23 pedal_edema 399 non-null object
24 anemia 399 non-null object
25 class 400 non-null object
dtypes: float64(11), int64(1), object(14)
memory usage: 81.4+ KB
```

In [11]: data.drop(['id'],axis=1,inplace=True)

Out[11]:

	age	blood_pressure	specific_gravity	albumin	sugar	red_blood_cells	pus_cell	pus_cell_clumps	bacteria	blood_glucose_random	packed_cell_volume	white_blood_cell_count
0	48.0	80.0	1.020	1.0	0.0	NaN	normal	notpresent	notpresent	121.0	...	44
1	7.0	50.0	1.020	4.0	0.0	NaN	normal	notpresent	notpresent	NaN	...	38
2	62.0	80.0	1.010	2.0	3.0	normal	normal	notpresent	notpresent	423.0	...	31
3	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent	117.0	...	32
4	51.0	80.0	1.010	2.0	0.0	normal	normal	notpresent	notpresent	108.0	...	35
...	...	...	...	...	...	...	...	...	...	...	...	...
395	55.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	140.0	...	47
396	42.0	70.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	75.0	...	54
397	12.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent	100.0	...	49
398	17.0	60.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	114.0	...	51
399	58.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent	131.0	...	53

400 rows x 25 columns

localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help

Code

```
In [12]: #target column
data['class'].unique()

Out[12]: array(['ckd', 'ckd\t', 'notckd'], dtype=object)

In [ ]:

In [13]: #rectify target column
data['class']=data['class'].replace('ckd\t','ckd')
data['class'].unique()

Out[13]: array(['ckd', 'notckd'], dtype=object)

In [14]: #fetching categorical column
cat=data.select_dtypes(include=['object']).columns.tolist()
cat

Out[14]: ['red_blood_cells',
'pus_cell',
'pus_cell_clumps',
'bacteria',
'packed_cell_volume',
'white_blood_cell_count',
'red_blood_cell_count',
'hypertension',
'diabetesmellitus',
'coronary_artery_disease',
'appetite',
'pedal_edema',
'anemia',
'class']
```

Doc2.pdf x localhost x CKD.pkl (edit) x Assignment-4 x Assignment-4 x Downloads/ x project - Jupyter x IBM-40609-1 x + -

localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (pykernel)

In [15]: #removing column which are not categorical  
cat.remove('red\_blood\_cell\_count')  
cat.remove('packed\_cell\_volume')  
cat.remove('white\_blood\_cell\_count')  
cat

Out[15]: ['red\_blood\_cells',  
'pus\_cell',  
'pus\_cell\_clumps',  
'bacteria',  
'hypertension',  
'diabetesmellitus',  
'coronary\_artery\_disease',  
'appetite',  
'pedal\_edema',  
'anemia',  
'class']

In [16]: num=data.select\_dtypes(include=['float64']).columns.tolist()#fetch numerical column

In [17]: num.remove('specific\_gravity')#remove which are not numerical  
num.remove('albumin')  
num.remove('sugar')  
num

Out[17]: ['age',  
'blood\_pressure',  
'blood\_glucose\_random',  
'blood\_urea',  
'serum\_creatinine',  
'sodium',  
'potassium',  
'hemoglobin']

Type here to search

Doc2.pdf x localhost x CKD.pkl (edit) x Assignment-4 x Assignment-4 x Downloads/ x project - Jupyter x IBM-40609-1 x + -

localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (pykernel)

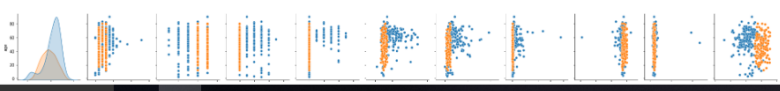
'serum\_creatinine',  
'sodium',  
'potassium',  
'hemoglobin']

In [18]: #adding column which is numerical  
num.append('red\_blood\_cell\_count')  
num.append('packed\_cell\_volume')  
num.append('white\_blood\_cell\_count')  
num

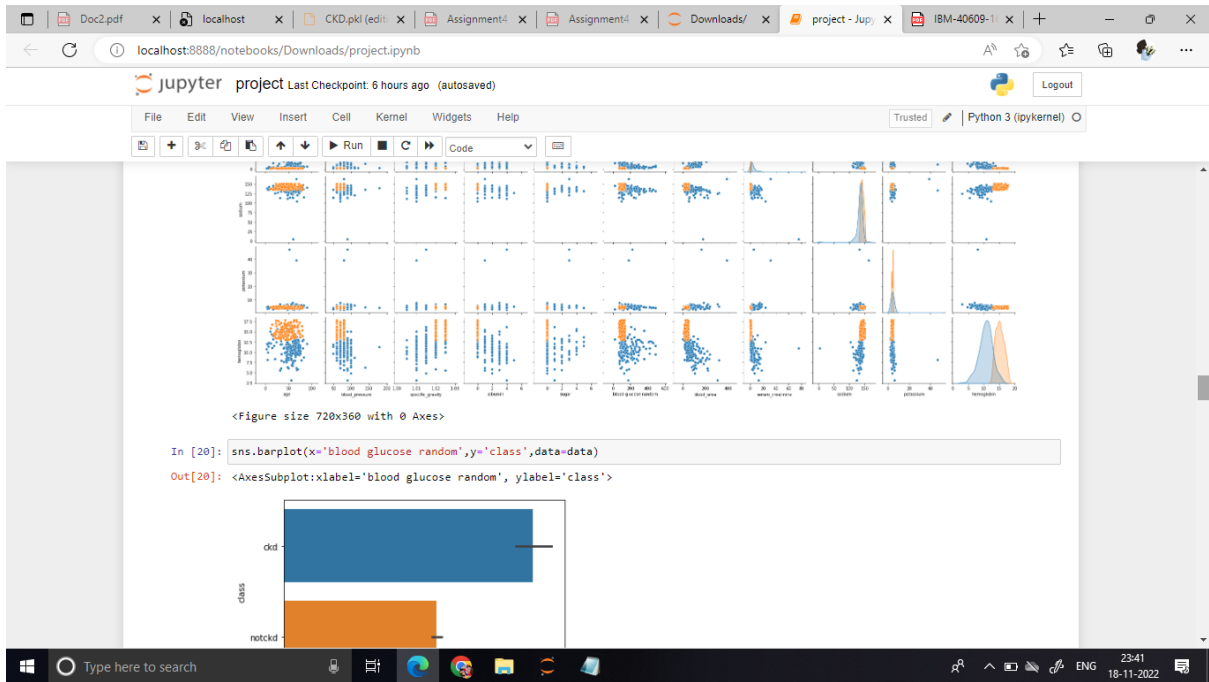
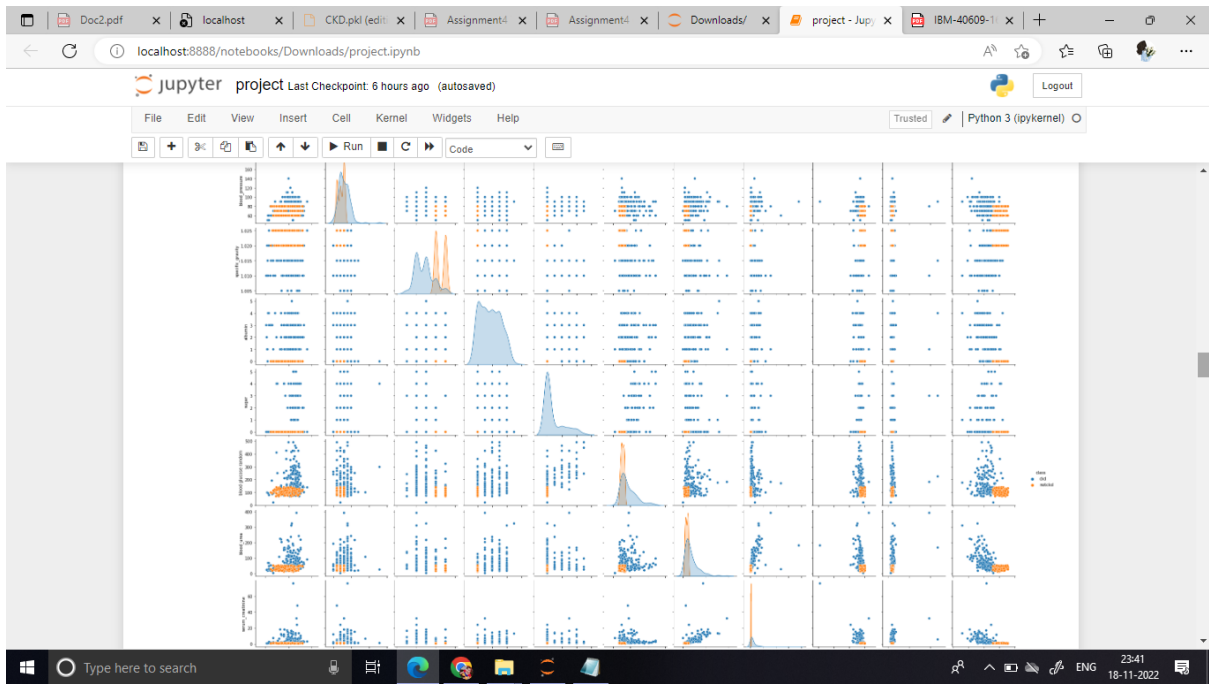
Out[18]: ['age',  
'blood\_pressure',  
'blood\_glucose\_random',  
'blood\_urea',  
'serum\_creatinine',  
'sodium',  
'potassium',  
'hemoglobin',  
'red\_blood\_cell\_count',  
'packed\_cell\_volume',  
'white\_blood\_cell\_count']

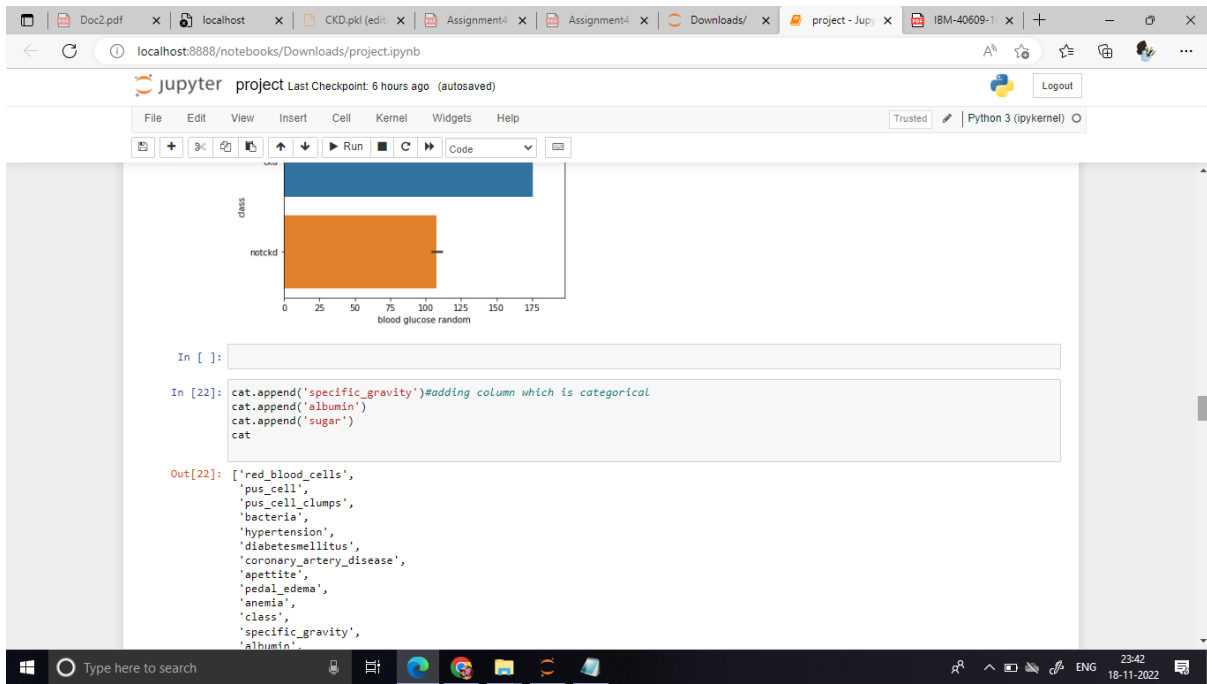
In [19]: sns.pairplot(data,hue='class')  
fig=plt.figure(figsize=(10,5))  
fig

Out[19]: <Figure size 720x360 with 0 Axes>



Type here to search





localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved)

File Edit View Insert Cell Kernel Widgets Help

Run Python 3 (pykernel)

```
'albumin',
'sugar',
'specific_gravity',
'albumin',
'sugar']
```

```
In [23]: a=data['coronary_artery_disease'].unique()
#b=data['sugar'].unique()
#c=data['albumin'].unique()
#d=data['specific_gravity'].unique()
#e=data['anemia'].unique()
#f=data['pedal_edema'].unique()#
#g=data['appetite'].unique()
h=data['diabetesmellitus'].unique()
#i=data['bacteria'].unique()
#j=data['hypertension'].unique()#
#k=data['red_blood_cell_count'].unique()
#l=data['pus_cell'].unique()
#m=data['pus_cell_clumps'].unique()
```

```
In [24]: a,h
```

```
Out[24]: (array(['no', 'yes', '\tno', nan], dtype=object),
array(['yes', 'no', ' yes', '\tno', '\tyes', nan], dtype=object))
```

```
In [25]: #rectifying the categorical column classes
data['coronary_artery_disease']=data['coronary_artery_disease'].replace('\tno','no')
data['coronary_artery_disease'].unique()
```

localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (pykernel)

```
Out[25]: array(['no', 'yes', nan], dtype=object)
```

```
In [26]: data['diabetesmellitus']=data.diabetesmellitus.replace(to_replace={'yes':'yes','\tno':'no'})
data['diabetesmellitus'].unique()
```

```
Out[26]: array(['yes', 'no', ' yes', nan], dtype=object)
```

```
In [27]: #handling missing value
data.isna().sum()
```

```
Out[27]: age                9
blood_pressure         12
specific_gravity       47
albumin               46
sugar                 49
red_blood_cells       152
pus_cell              65
pus_cell_clumps        4
bacteria              4
blood_glucose_random   44
blood_urea            19
serum_creatinine      17
sodium                87
potassium             88
hemoglobin            52
packed_cell_volume     70
white_blood_cell_count 105
red_blood_cell_count   130
hypertension           2
diabetesmellitus        2
coronary_artery_disease 1
appetite               1
pedal_edema            1
```

localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (pykernel)

```
pedal_edema      0
anemia           0
class            0
dtype: int64
```

```
In [ ]:
```

```
In [32]: from sklearn.preprocessing import LabelEncoder
for i in cat:
    print('label of encoder= ',i)
    lei=LabelEncoder()
    print(data[i])
    data[i]=lei.fit_transform(data[i])
    print(data[i])
    print('***100')
```

```
label of encoder= red_blood_cells
0      normal
1      normal
2      normal
3      normal
4      normal
...
395    normal
396    normal
397    normal
398    normal
399    normal
Name: red_blood_cells, Length: 400, dtype: object
0      1
1      1
2      1
```

Doc2.pdf x localhost x CKD.pkl (edit) x Assignment-4 x Assignment-4 x Downloads/ x project - Jup: x IBM-40609-1 x + -

localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (pykernel)

```
390 normal
397 normal
398 normal
399 normal
Name: red_blood_cells, Length: 400, dtype: object
0 1
1 1
2 1
3 1
4 1
...
```

In [33]: data.corr().T

Out[33]:

	age	blood_pressure	specific_gravity	albumin	sugar	red_blood_cells	pus_cell	pus_cell_clumps	bacteria	blood glucose random	...
age	1.000000	0.146830	-0.161959	0.087907	0.185221	-0.019305	-0.102595	0.154774	0.040733	0.213779	...
blood_pressure	0.146830	1.000000	-0.164422	0.122541	0.190218	-0.151369	-0.156856	0.059560	0.112173	0.149100	...
specific_gravity	-0.161959	-0.164422	1.000000	-0.479962	-0.292053	0.253894	0.365353	-0.306426	-0.231704	-0.317893	...
albumin	0.087907	0.122541	-0.479962	1.000000	0.287751	-0.394844	-0.561713	0.417868	0.377935	0.310481	...
sugar	0.185221	0.190218	-0.292053	0.287751	1.000000	-0.092940	-0.190062	0.168091	0.119399	0.629809	...
red_blood_cells	-0.019305	-0.151369	0.253894	-0.394844	-0.092940	1.000000	0.377394	-0.102948	-0.184402	-0.153076	...
pus_cell	-0.102595	-0.156856	0.365353	-0.561713	-0.190062	0.377394	1.000000	-0.520118	-0.330401	-0.262259	...
pus_cell_clumps	0.154774	0.059560	-0.306426	0.417868	0.168091	-0.102948	-0.520118	1.000000	0.275082	0.197593	...
bacteria	0.040733	0.112173	-0.231704	0.377935	0.119399	-0.184402	-0.330401	0.275082	1.000000	0.085940	...
blood glucose random	0.213779	0.149100	-0.317893	0.310481	0.629809	-0.153076	-0.262259	0.197593	0.085940	1.000000	...
blood_urea	0.190669	0.183970	-0.249370	0.346935	0.126043	-0.236322	-0.344048	0.184415	0.158444	0.127489	...

Type here to search

23:43 18-11-2022

Doc2.pdf x localhost x CKD.pkl (edit) x Assignment-4 x Assignment-4 x Downloads/ x project - Jup: x IBM-40609-1 x + -

localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (pykernel)

```
red_blood_cells -0.019305 -0.151369 0.253894 -0.394844 -0.092940 1.000000 0.377394 -0.102948 -0.184402 -0.153076 ...
pus_cell -0.102595 -0.156856 0.365353 -0.561713 -0.190062 0.377394 1.000000 -0.520118 -0.330401 -0.262259 ...
pus_cell_clumps 0.154774 0.059560 -0.306426 0.417868 0.168091 -0.102948 -0.520118 1.000000 0.275082 0.197593 ...
bacteria 0.040733 0.112173 -0.231704 0.377935 0.119399 -0.184402 -0.330401 0.275082 1.000000 0.085940 ...
blood glucose random 0.213779 0.149100 -0.317893 0.310481 0.629809 -0.153076 -0.262259 0.197593 0.085940 1.000000 ...
blood_urea 0.190669 0.183970 -0.249370 0.346935 0.126043 -0.236322 -0.344048 0.184415 0.158444 0.127489 ...
serum_creatinine 0.128292 0.144359 -0.176146 0.160252 0.094565 -0.138394 -0.157896 0.049940 0.050830 0.082242 ...
sodium -0.086552 -0.103220 0.217473 -0.228078 -0.053452 0.140572 0.173323 -0.142135 -0.081733 -0.154393 ...
potassium 0.049315 0.066648 -0.063324 0.111614 0.180067 0.018192 -0.158750 -0.006316 -0.002688 0.056695 ...
hemoglobin -0.182737 -0.279535 0.492143 -0.474211 -0.156876 0.280990 0.411500 -0.275763 -0.204954 -0.269149 ...
packed_cell_volume -0.213961 -0.286473 0.491157 -0.470819 -0.180658 0.278263 0.417339 -0.297980 -0.188624 -0.259831 ...
white_blood_cell_count 0.118077 0.039957 -0.238735 0.214295 0.161132 -0.021104 -0.107753 0.158224 0.095702 0.137186 ...
red_blood_cell_count -0.202451 -0.209598 0.371683 -0.368210 -0.149470 0.166161 0.365479 -0.238772 -0.186593 -0.194327 ...
hypertension 0.396749 0.270447 -0.323643 0.406057 0.254268 -0.140538 -0.291719 0.195623 0.089046 0.369849 ...
diabetesmellitus 0.357048 0.231573 -0.352804 0.312978 0.431277 -0.148374 -0.204596 0.167585 0.081995 0.504587 ...
coronary_artery_disease 0.228872 0.086618 -0.135814 0.200957 0.229301 -0.111493 -0.172295 0.188029 0.162395 0.212433 ...
appetite 0.154683 0.177601 -0.230975 0.303145 0.069216 -0.160668 -0.274985 0.189688 0.149126 0.176132 ...
pedal_edema 0.095099 0.054083 -0.253803 0.411080 0.116442 -0.199285 -0.350227 0.104356 0.134732 0.101502 ...
anemia 0.056215 0.194962 -0.184155 0.229558 0.042464 -0.107625 -0.260566 0.175861 0.052208 0.126854 ...
class -0.230926 -0.290600 0.659504 -0.531562 -0.294555 0.282642 0.375154 -0.265313 -0.186871 -0.401374 ...
```

25 rows x 25 columns

Type here to search

23:44 18-11-2022



localhost:8888/notebooks/Downloads/project.ipynb

jupyter project Last Checkpoint: 6 hours ago (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (pykernel)

Run

apetite	0.154683	0.177901	-0.230975	0.303145	0.069216	-0.160068	-0.274985	0.189688	0.149126	0.176132	...
pedal_edema	0.095099	0.054003	-0.253803	0.411080	0.116442	-0.199285	-0.350227	0.104356	0.134732	0.101502	...
anemia	0.056215	0.194962	-0.184155	0.229558	0.042464	-0.107625	-0.260566	0.175861	0.052208	0.126854	...
class	-0.230926	-0.290900	0.659504	-0.531562	-0.294555	0.282642	0.375154	-0.265313	-0.186871	-0.401374	...

25 rows x 25 columns

```
In [34]: selcols=['red_blood_cells','pus_cell',
'diabetesmellitus', 'coronary_artery_disease','blood_urea','pedal_edema','anemia',
'blood_glucose_random']

In [35]: x=pd.DataFrame(data,columns=selcols)
y=pd.DataFrame(data,columns=['class'])
print(x.shape)
print(y.shape)

(400, 8)
(400, 1)

In [36]: from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,random_state=2)
print(x_train.shape)
print(y_train.shape)
print(x_test.shape)
print(y_test.shape)

(320, 8)
(320, 1)
(80, 8)
(80, 1)
```

Type here to search

23:44 18-11-2022