

STATS 3DA3

Homework Assignment 6

Zichuan Xu(400374887), Junbo Tan(400310355)

2024-04-04

Introduction

The dataset we use for the assignment is from [Chronic Kidney Disease \(CKD\)](#), which can be used to predict the chronic kidney disease. The data has 24 features, we import this data as two parts. The first part is X, which is the explanatory variable. The another part is called y, which is response variable, i.e. the status of chronic kidney disease.

```
from ucimlrepo import fetch_ucirepo
import pandas as pd
# fetch dataset
chronic_kidney_disease = fetch_ucirepo(id=336)

# data (as pandas dataframes)
X = chronic_kidney_disease.data.features
y = chronic_kidney_disease.data.targets
X
print(X)

y
```

	age	bp	sg	al	su	rbc	pc	pcc	ba \
0	48.0	80.0	1.020	1.0	0.0	NaN	normal	notpresent	notpresent
1	7.0	50.0	1.020	4.0	0.0	NaN	normal	notpresent	notpresent
2	62.0	80.0	1.010	2.0	3.0	normal	normal	notpresent	notpresent

3	48.0	70.0	1.005	4.0	0.0	normal	abnormal	present	notpresent
4	51.0	80.0	1.010	2.0	0.0	normal	normal	notpresent	notpresent
..
395	55.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent
396	42.0	70.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent
397	12.0	80.0	1.020	0.0	0.0	normal	normal	notpresent	notpresent
398	17.0	60.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent
399	58.0	80.0	1.025	0.0	0.0	normal	normal	notpresent	notpresent

	bgr	...	hemo	pcv	wbcc	rbcc	htn	dm	cad	appet	pe	ane
0	121.0	...	15.4	44.0	7800.0	5.2	yes	yes	no	good	no	no
1	NaN	...	11.3	38.0	6000.0	NaN	no	no	no	good	no	no
2	423.0	...	9.6	31.0	7500.0	NaN	no	yes	no	poor	no	yes
3	117.0	...	11.2	32.0	6700.0	3.9	yes	no	no	poor	yes	yes
4	106.0	...	11.6	35.0	7300.0	4.6	no	no	no	good	no	no
..
395	140.0	...	15.7	47.0	6700.0	4.9	no	no	no	good	no	no
396	75.0	...	16.5	54.0	7800.0	6.2	no	no	no	good	no	no
397	100.0	...	15.8	49.0	6600.0	5.4	no	no	no	good	no	no
398	114.0	...	14.2	51.0	7200.0	5.9	no	no	no	good	no	no
399	131.0	...	15.8	53.0	6800.0	6.1	no	no	no	good	no	no

[400 rows x 24 columns]

	class
0	ckd
1	ckd
2	ckd
3	ckd
4	ckd
...	...