

Principal Component Analysis

Dataset:

```
# print the updated data  
df2
```

	age	bp	al	su	bgr	bu	sc	sod	pot	hrmo	...	pc_normal	pcc_present	ba_present	htn_yes
0	2.000000	76.459948	3.0	0.0	148.112676	57.482105	3.077356	137.528754	4.627244	12.518156	...	False	False	False	False
1	3.000000	76.459948	2.0	0.0	148.112676	22.000000	0.700000	137.528754	4.627244	10.700000	...	True	False	False	False
2	4.000000	76.459948	1.0	0.0	99.000000	23.000000	0.600000	138.000000	4.400000	12.000000	...	True	False	False	False
3	5.000000	76.459948	1.0	0.0	148.112676	16.000000	0.700000	138.000000	3.200000	8.100000	...	True	False	False	False
4	5.000000	50.000000	0.0	0.0	148.112676	25.000000	0.600000	137.528754	4.627244	11.800000	...	True	False	False	False
...
394	51.492308	70.000000	0.0	0.0	219.000000	36.000000	1.300000	139.000000	3.700000	12.500000	...	True	False	False	False
395	51.492308	70.000000	0.0	2.0	220.000000	68.000000	2.800000	137.528754	4.627244	8.700000	...	True	False	False	True
396	51.492308	70.000000	3.0	0.0	110.000000	115.000000	6.000000	134.000000	2.700000	9.100000	...	True	False	False	True
397	51.492308	90.000000	0.0	0.0	207.000000	80.000000	6.800000	142.000000	5.500000	8.500000	...	True	False	False	True
398	51.492308	80.000000	0.0	0.0	100.000000	49.000000	1.000000	140.000000	5.000000	16.300000	...	True	False	False	False

399 rows x 28 columns

Classification Analysis:

```
# result of n = 2  
result
```

	Logistic	SVMI	SVMnl	KNN	Navie	Decision	Random
pca: 2	0.99	0.99	1.0	1.0	1.0	0.99	0.99

Here, SVMnl, KNN and Naive provide best accuracy than other algorithms

Regression Analysis:

```
# result of n=2  
result
```

	Linear	SVMI	SVMnl	Decision	Random
pca:2	0.583459	0.519554	0.932962	0.956597	0.979167

Here, Random forest provides best r2 score than other algorithms