

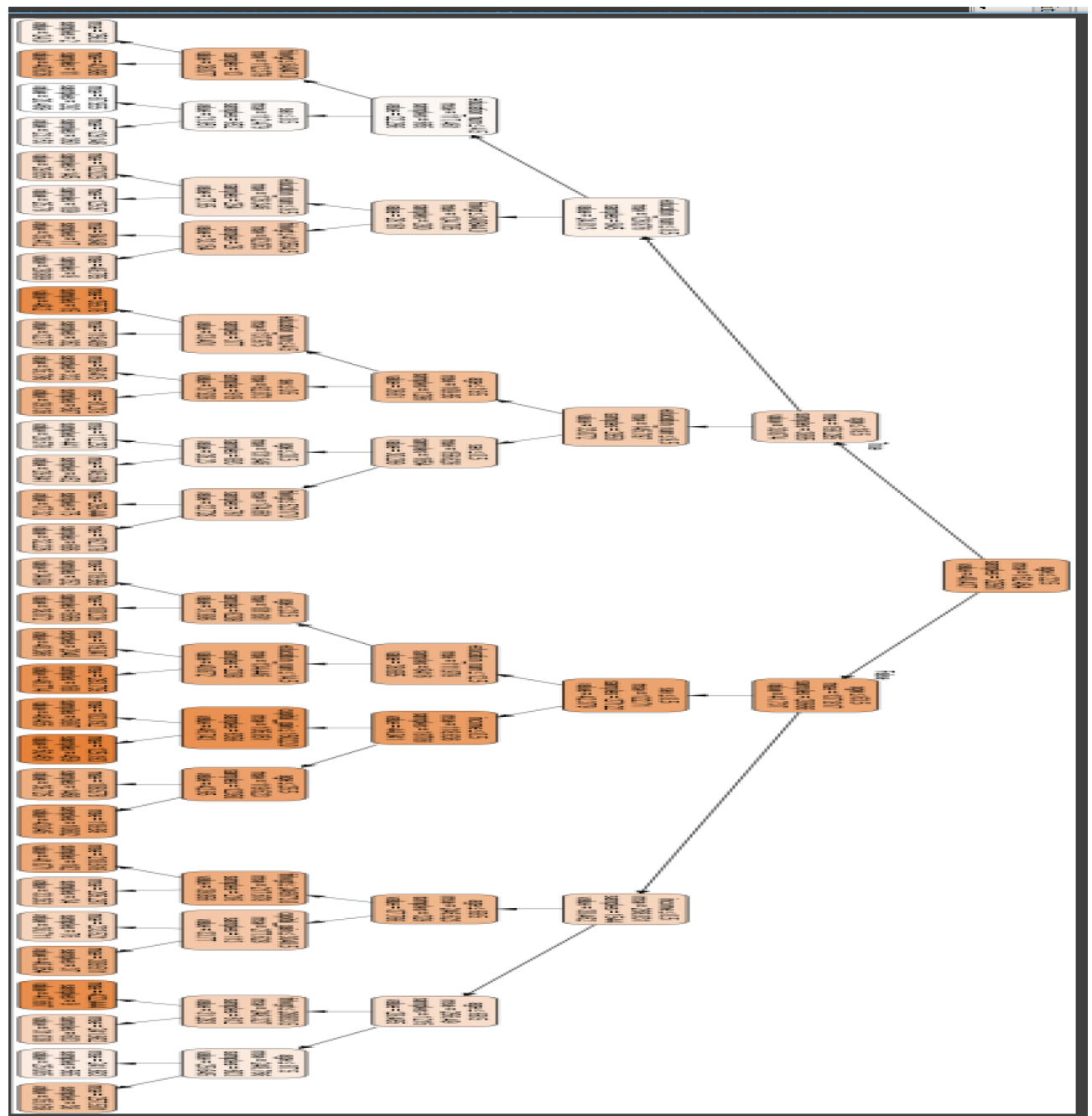
Statistical Learning Computer HW.B

106061703 藍國瑞

1.1.1

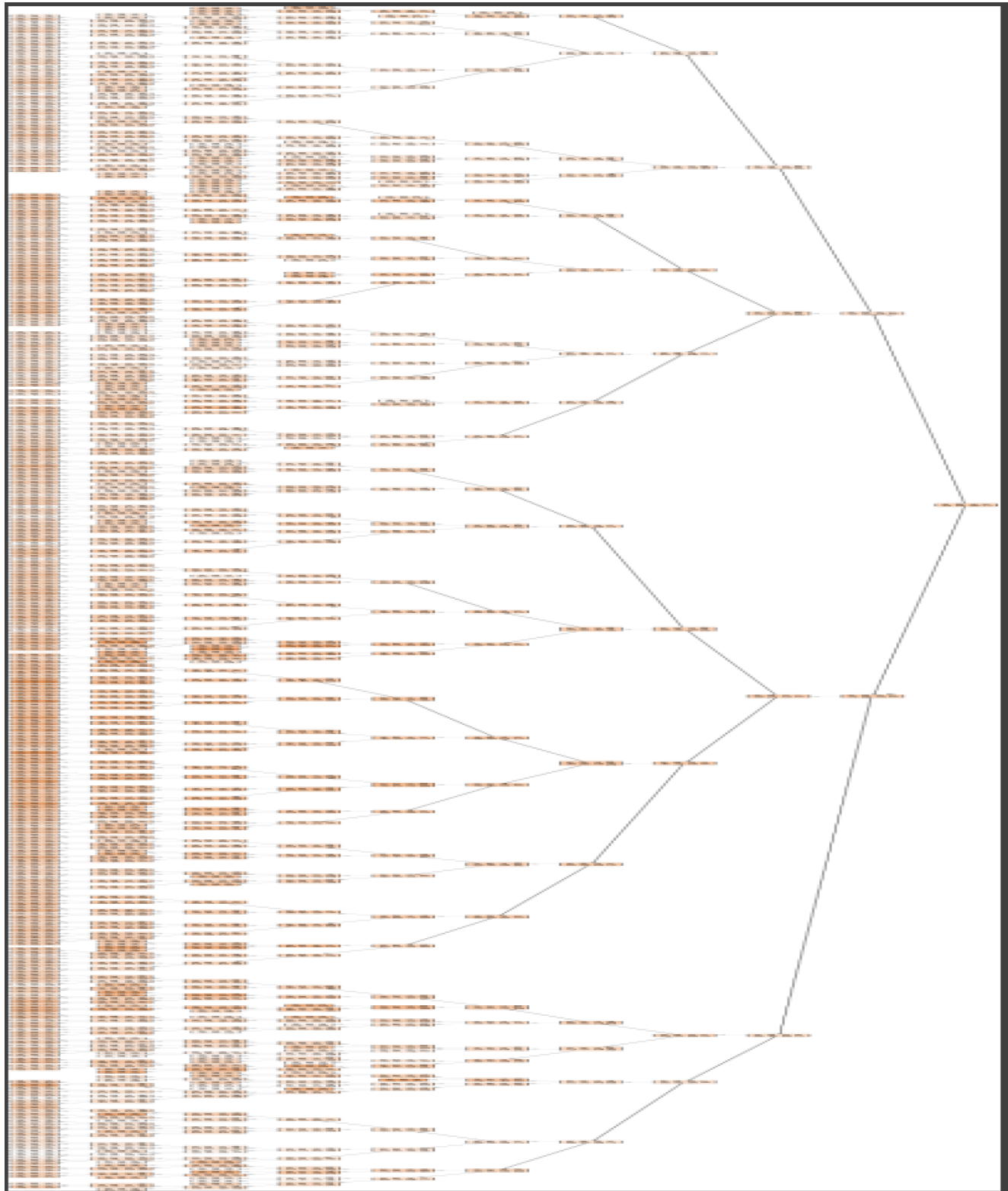
- (max depth, min samples split,min samples leaf)=(5,2,2)

Training mean square error	119.77199843514842
Test mean square error	123.63213395147034
R^2	0.21437448215705046



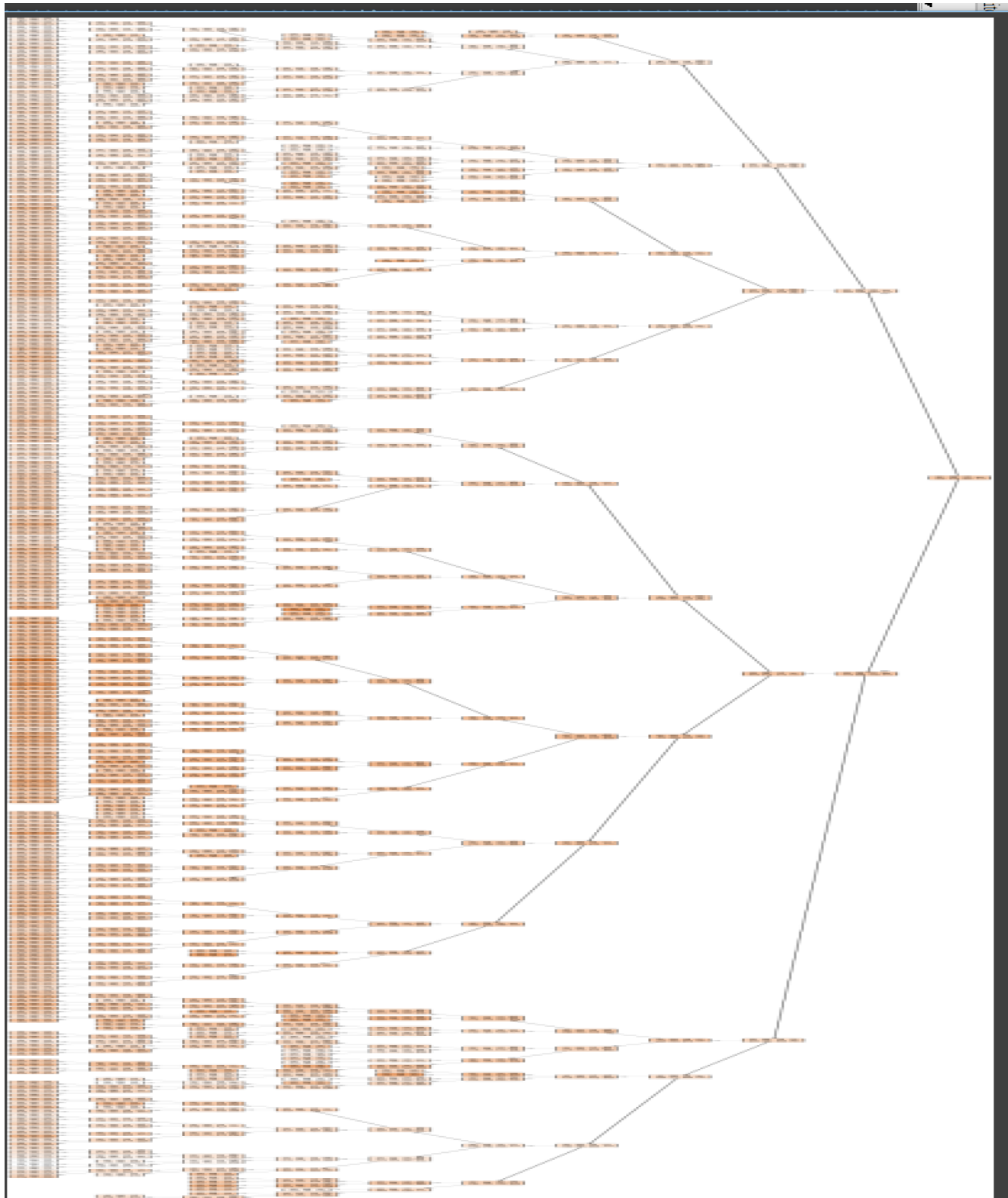
- (max depth, min samples split,min samples leaf)=(10,2,2)

Training mean square error	110.03689211528922
Test mean square error	128.27718201825272
R^2	0.2782303753851878



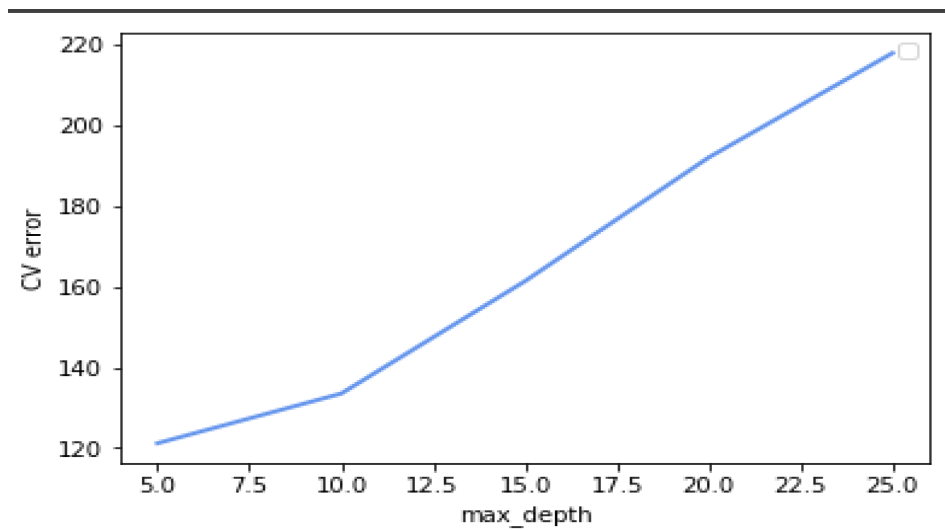
- (max depth, min samples split,min samples leaf)=(10,4,4)

Training mean square error	111.45494999253383
Test mean square error	127.15071253648969
R^2	0.268928848578446



1.1.2

Max depth	3-fold CV error
5	121.212
10	133.525
15	161.371
20	192.031
25	217.891



1.2

- Bagging

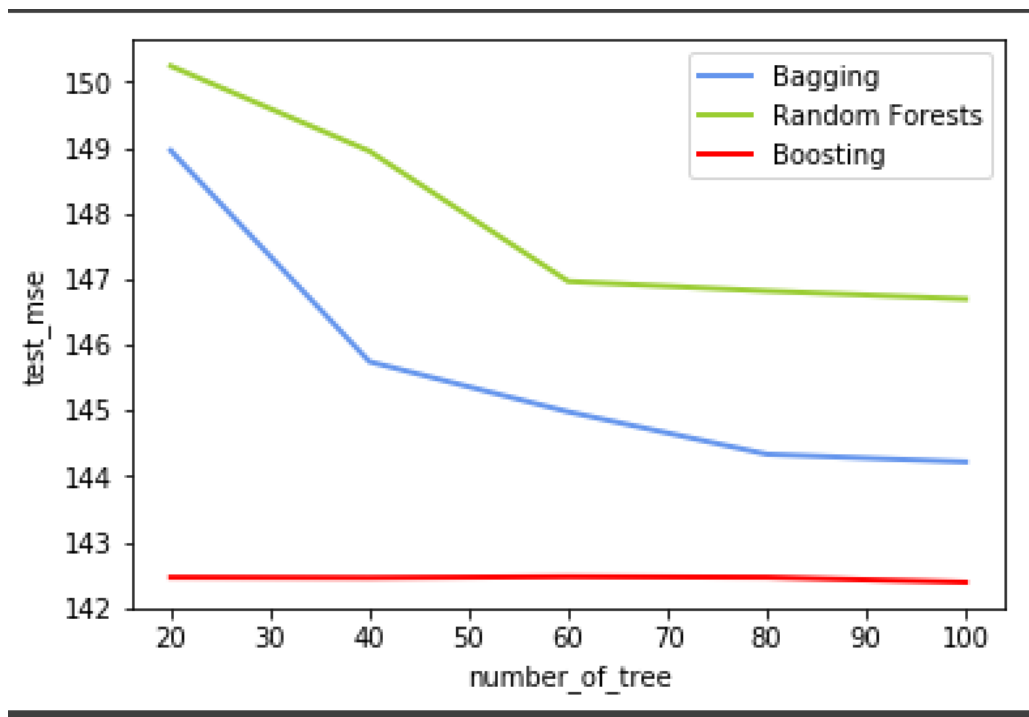
Number of tree	Test MSE
20	148.95455067721468
40	145.74115869680244
60	144.978868964886
80	144.3334646368326
100	144.21616018652705

- RandomForest

Number of tree	Test MSE
20	150.24074557071472
40	148.94564165569193
60	146.96332982398616
80	146.81787445498253
100	146.6959493083898

- Boosting

Number of tree	Test MSE
20	142.46105958194016
40	142.45543808723215
60	142.4703821814956
80	142.46189710327204
100	142.38483260849543



2.1

- $C=10^1$

- Number of support vectors: 200

- Confusion matrix

3454	0
0	3545

- $C=10^{-1}$

- Number of support vectors: 227

- Confusion matrix

3446	8
9	3536

- $C=10^{-2}$

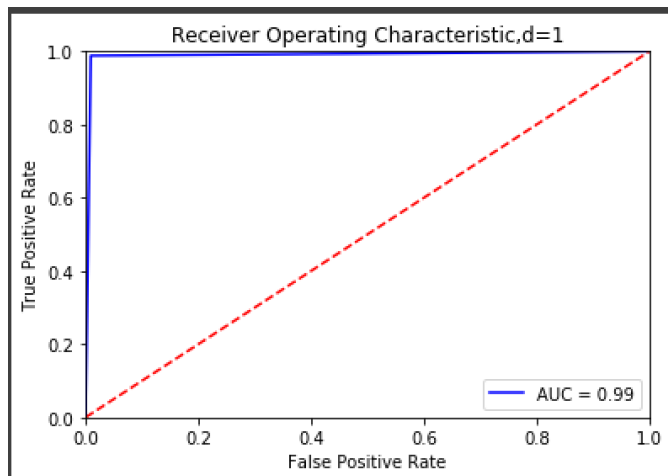
- Number of support vectors: 393

- Confusion matrix

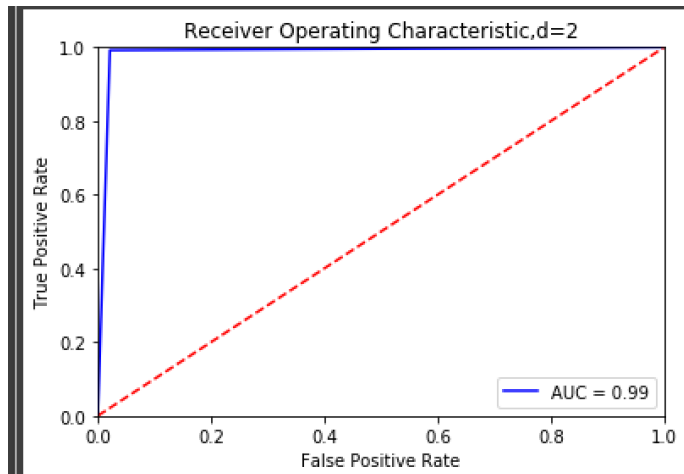
3435	19
25	3520

2.2

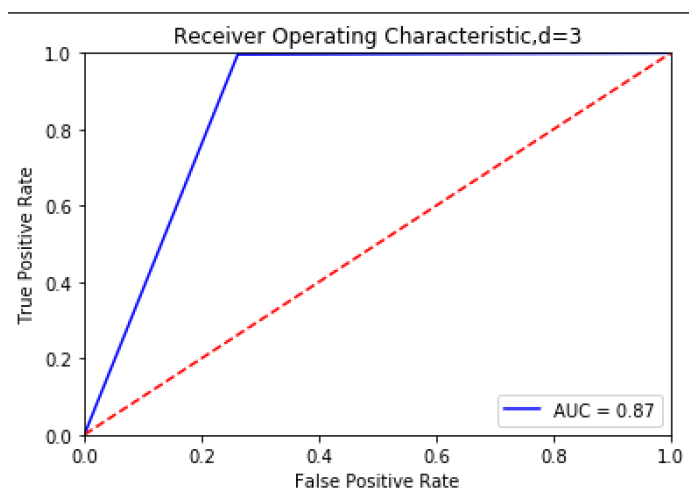
- $d=1$



- $d=2$

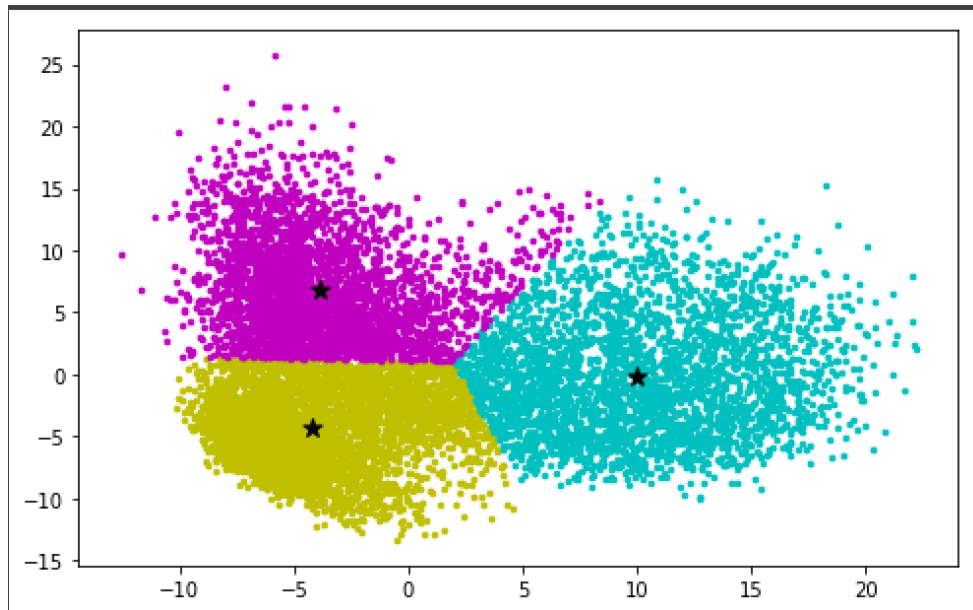


- $d=3$



2.3

- K-means clustering



- True label

