Name: Chim ka long
SID: 1155094482
Assignment 2

1). Train ():

```
C:\Users\jackc\Desktop\lab report\CSCI3320\asgn2>python ex1.py
Train:
Number of wrong predictions is: 0
```

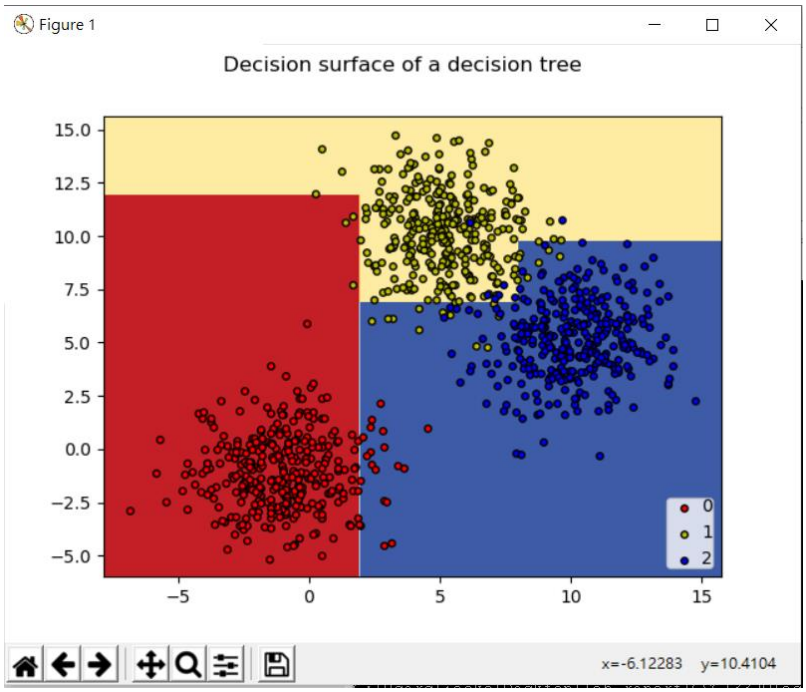
Train_matrix():

```
C:\Users\jackc\Desktop\lab report\CSCI3320\asgn2>python ex1.py
Train_Matrix:
Number of wrong predictions is: 0
```

I test 5 times for train() and train_matrix() respectively. All of them have 0 wrong predictions.

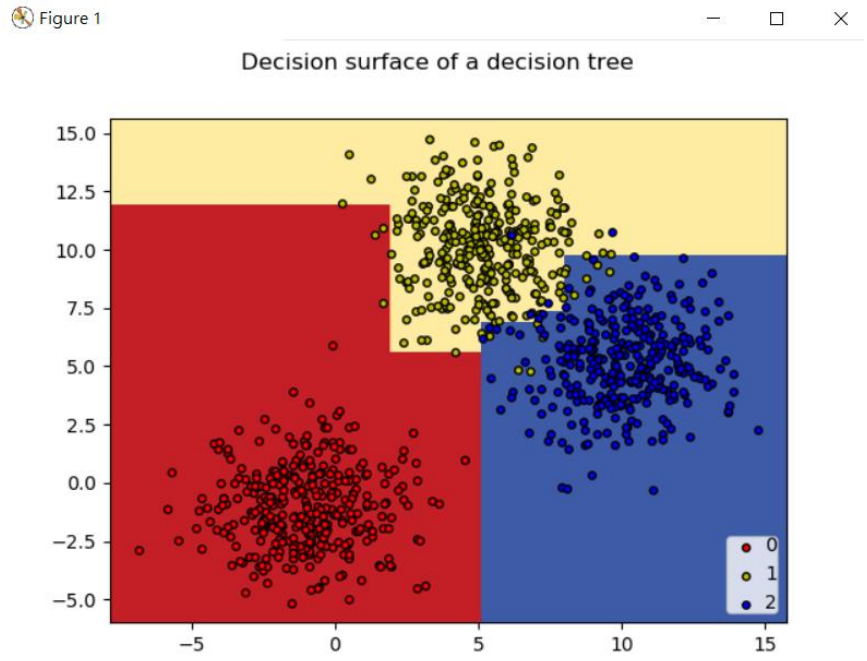
2). For max_depth = 3 :

```
C:\Users\jackc\Desktop\lab report\CSCI3320\asgn2>python ex2.py
Number of wrong predictions is: 18
```



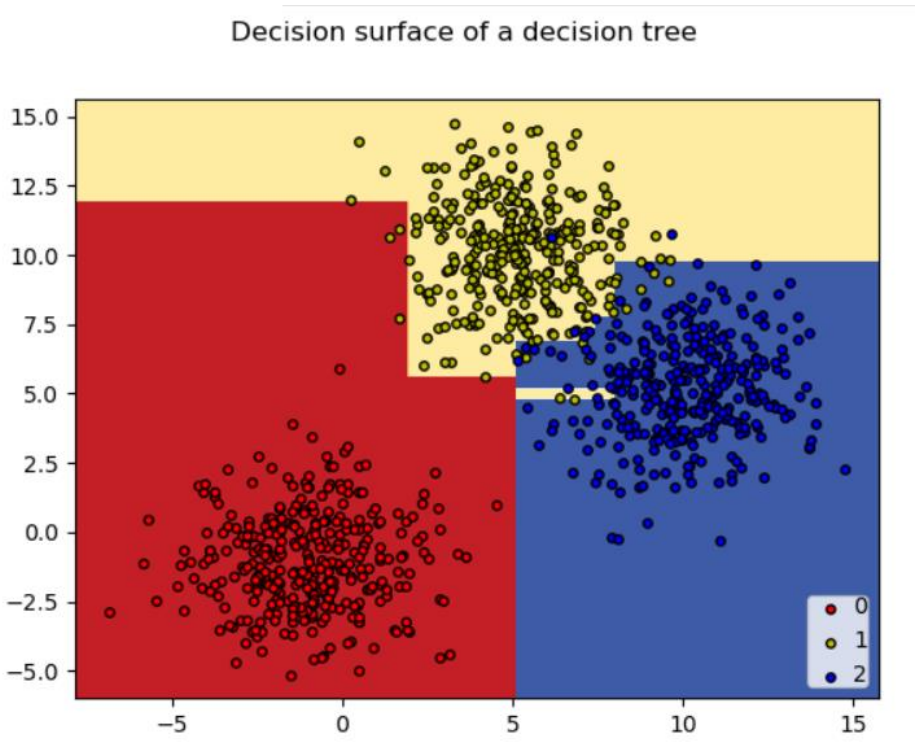
For max_depth = 5:

```
C:\Users\jackc\Desktop\lab report\CSCI3320\asgn2>python ex2.py
Number of wrong predictions is: 9
```



For max_depth = 7:

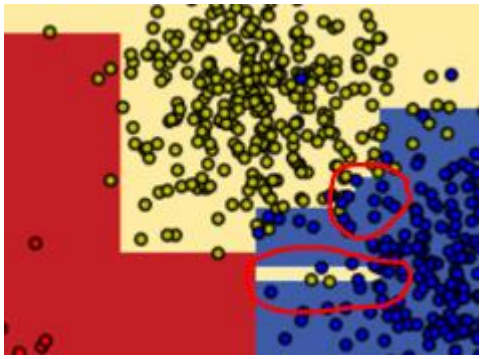
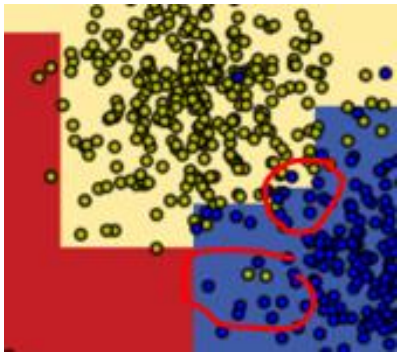
```
C:\Users\jackc\Desktop\lab report\CSCI3320\asgn2>python ex2.py
Number of wrong predictions is: 10
```



Max_depth	3	5	7
Number of wrong prediction	18	9	10

From figure, we can see decision tree classify more regions detail. However, the error is minimum when maximum depth is 5, due to bias-variance dilemma. Very high complexity can decrease error in train set but may increase error in test set.

When we compare the difference between figures of max_depth = 5, 7.

There is small yellow area surrounded by large blue area, because model of high complexity is sensitive to those outlier points in train set. However, those small area may wrongly classify points in test set.

Enviroment:
Python 3.5.3

Package	Version
imageio	2.8.0
kiwisolver	1.1.0
matplotlib	3.0.3
numpy	1.18.2
Pillow	7.0.0
pip	20.0.2
scikit-learn	0.22.2.post1
scipy	1.4.1