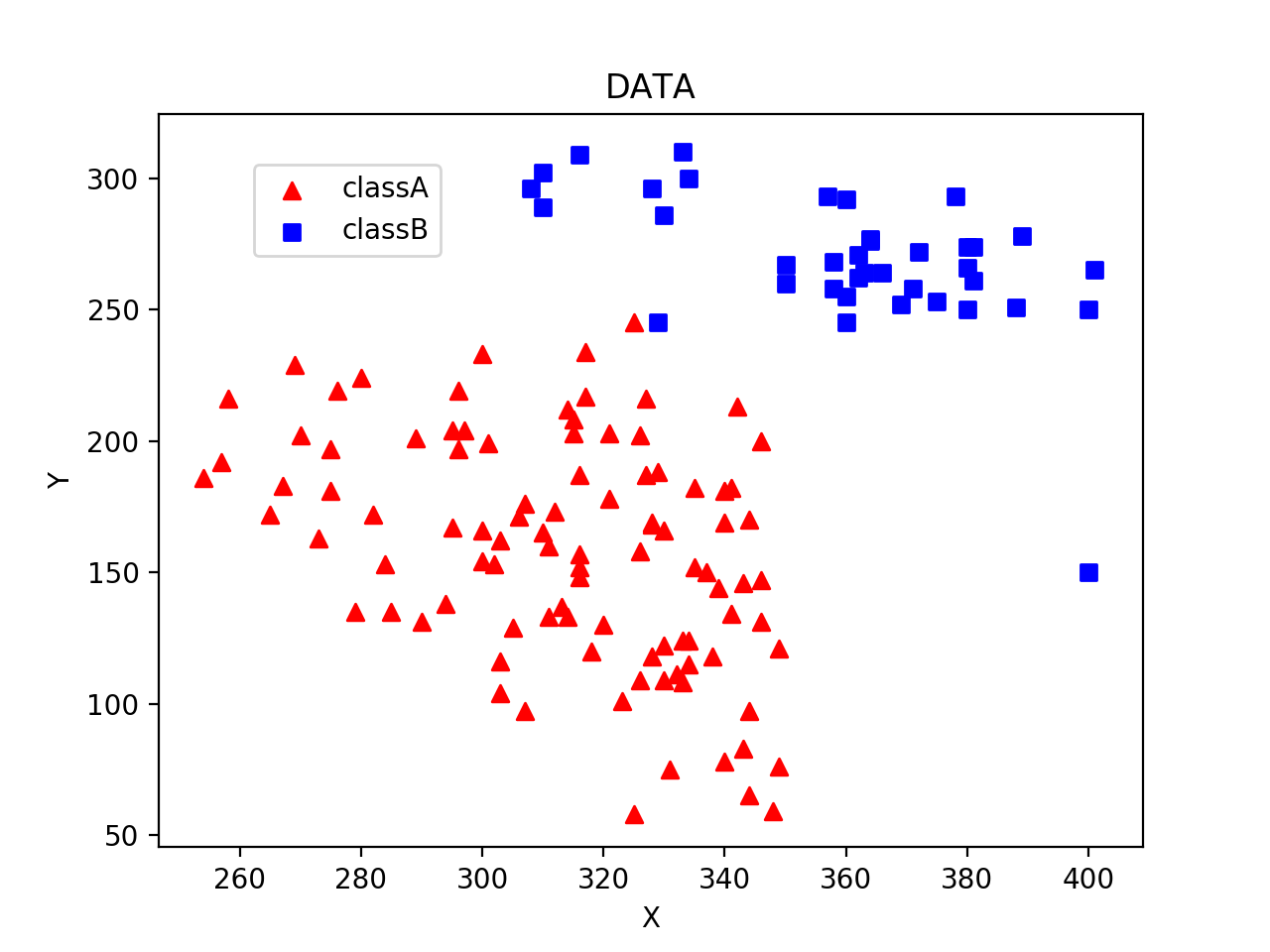
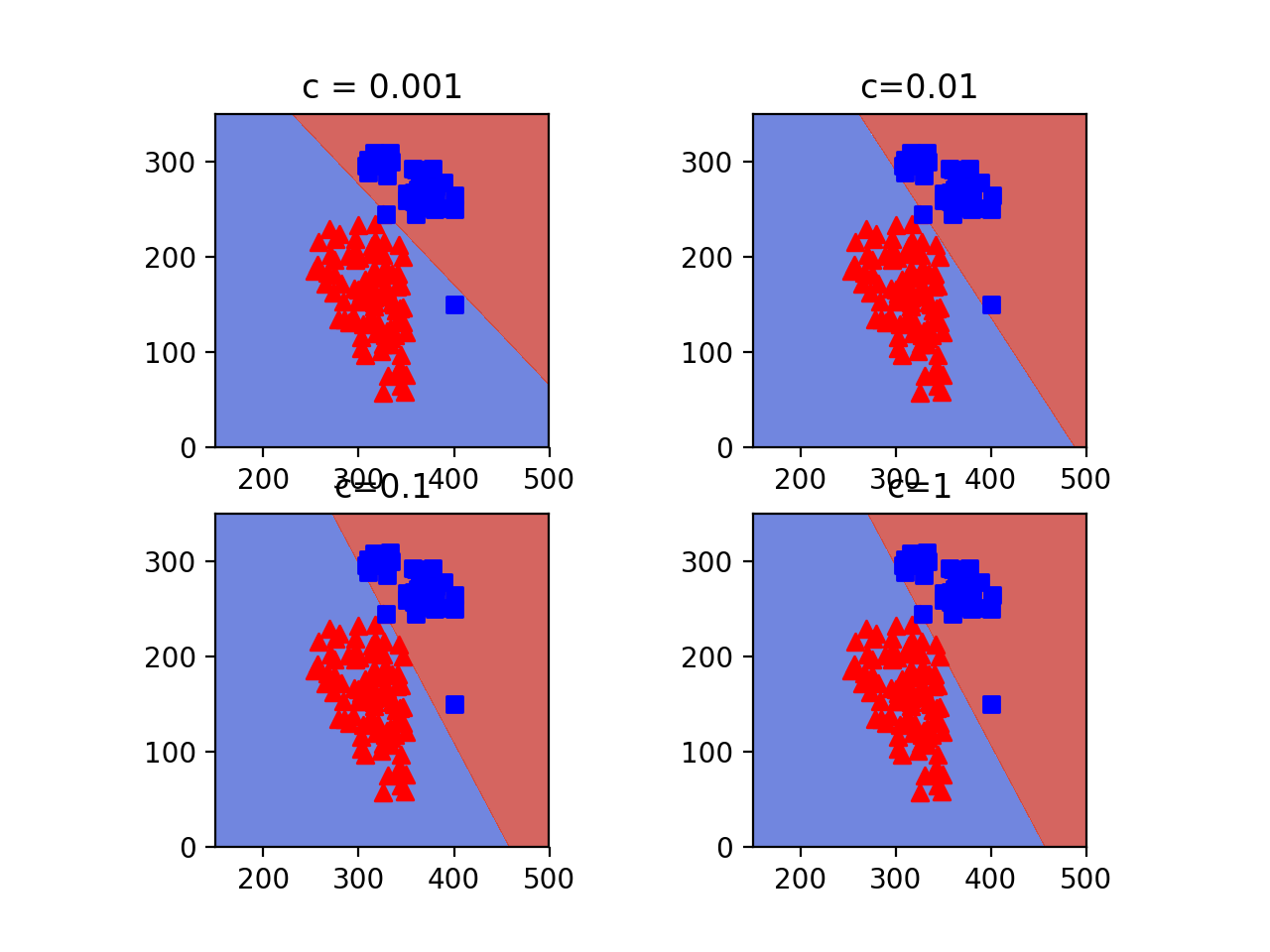
question 1

a.



c.

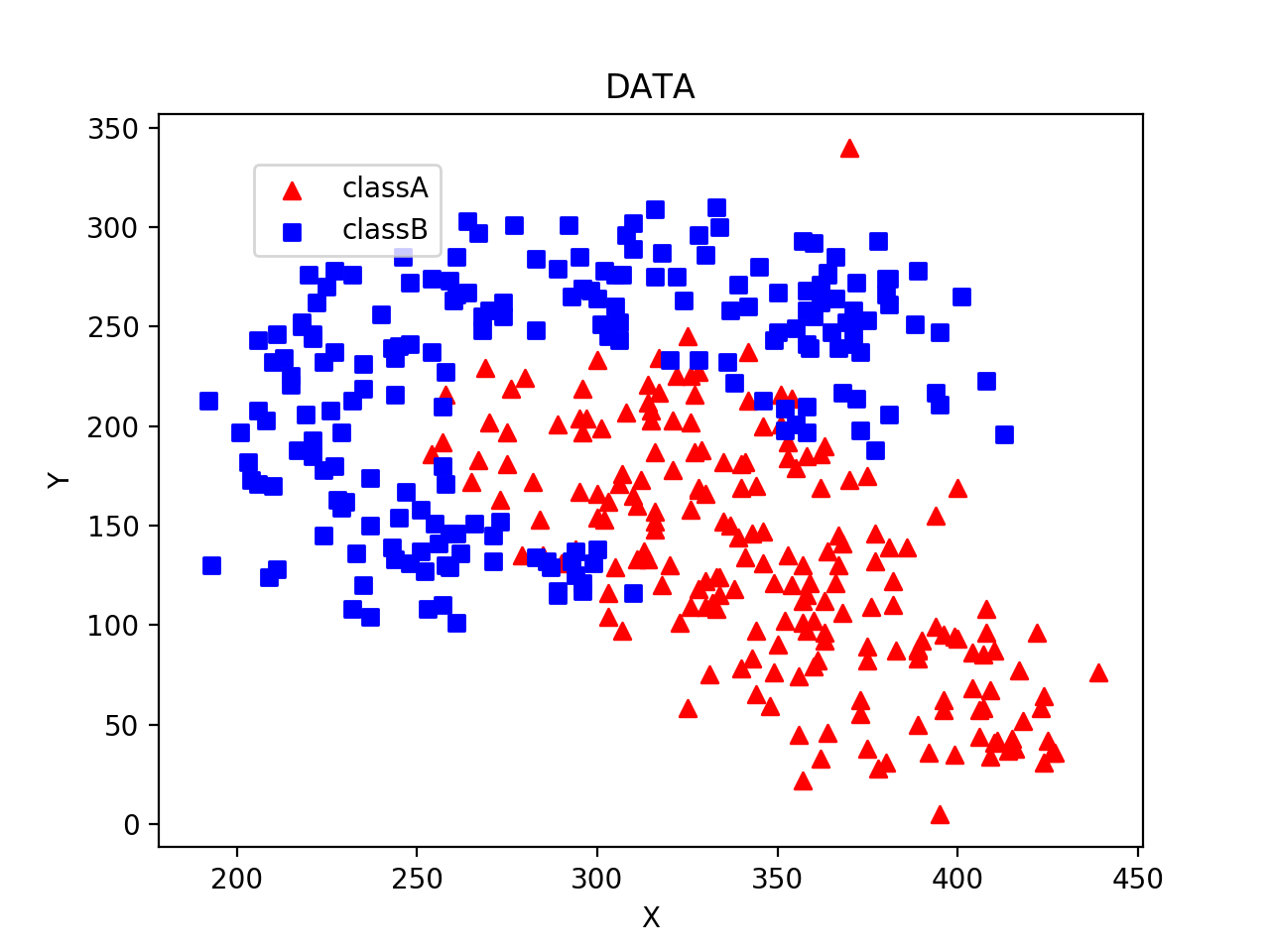
With C value increasing, we can see the division on plot is more accurate. It can split two class data well when C increase. It will more fit for the training dataset. However, it may cause over fitting and decrease the accuracy.

d.

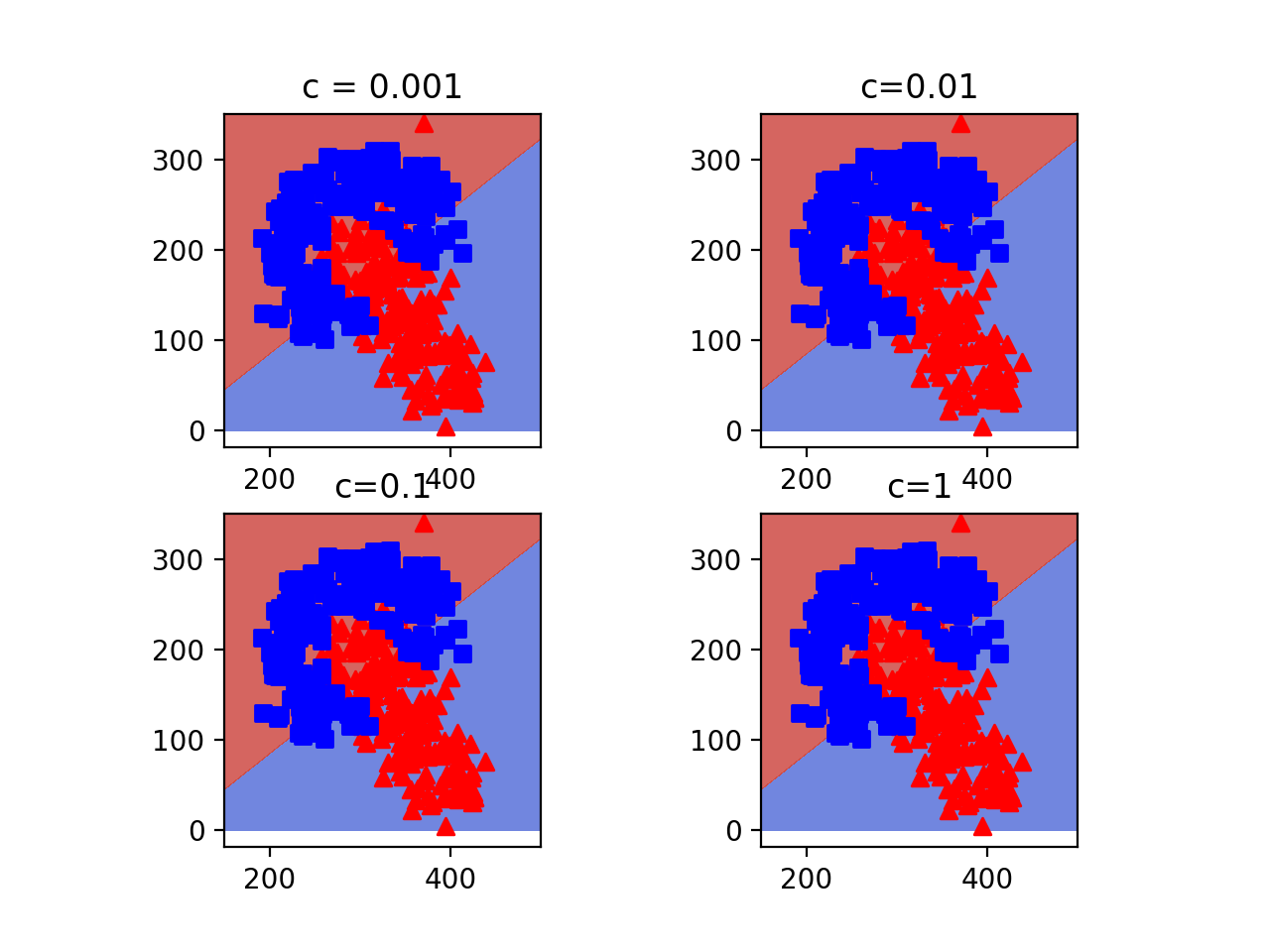
when c is 1, we can get the best result. When c increasing, the accuracy of training data increase. However, it may cause over fitting of test data and decrease the test accuracy.

Question 2

a.



b.



the accuracy shows below

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| C value | 0.1 | 1 | 10 | 100 |
| Accuracy | 0.79843902439 | 0.807420731707 | 0.79745121951 | 0.796768292682 |

we can see from the result, we can get the highest accuracy when c is 1.

c.

for this part, we select c=1 on SVM in our test.

d.

mean accuracy : 0.8677073170731708

variance 0.0032310

e. 