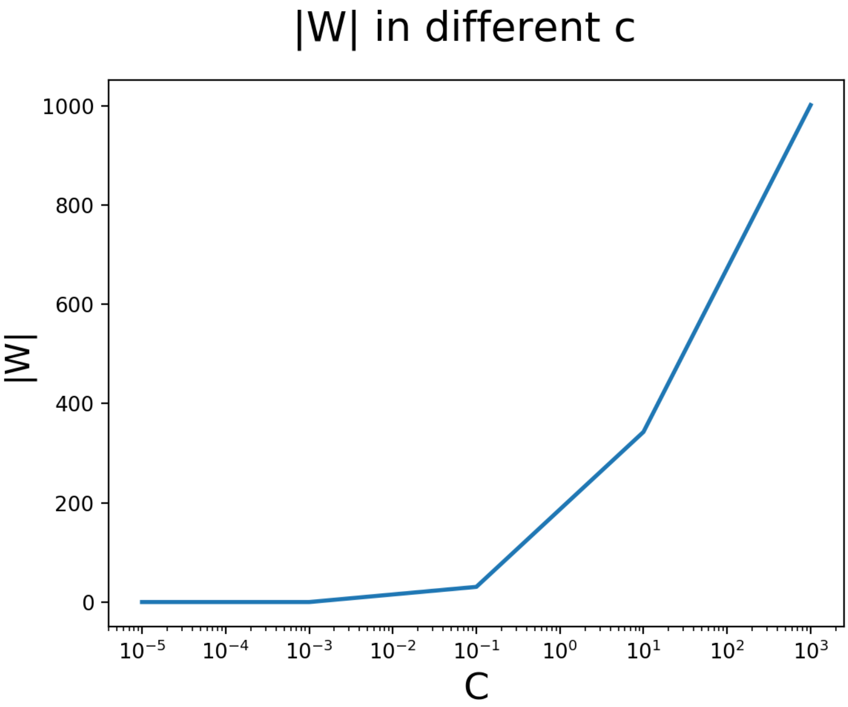
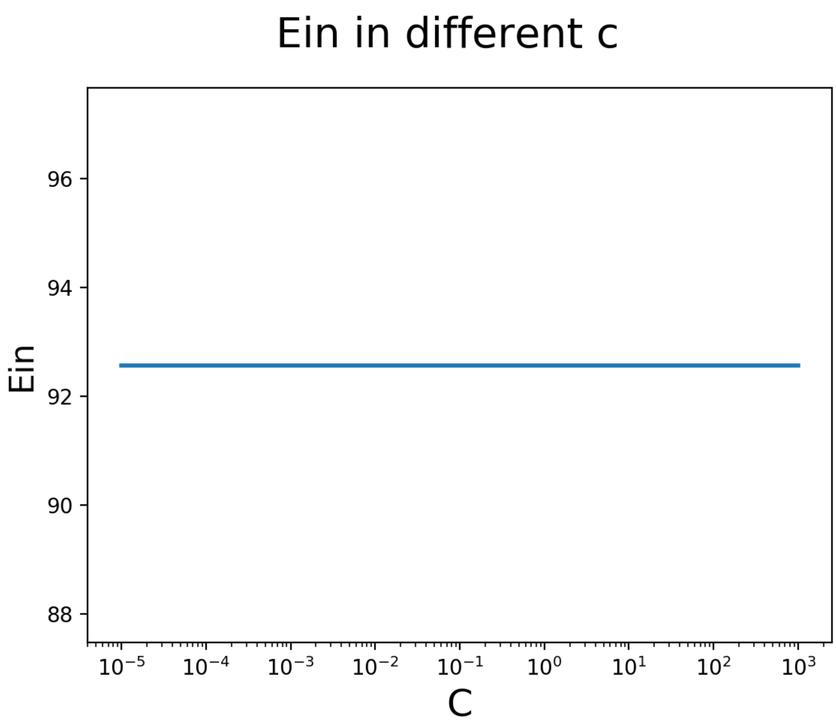
11.

The |w| becomes larger while C increases. This makes sense because larger C implies lower tolerance to error. In order to make sure error is lower, |w| have to be larger.



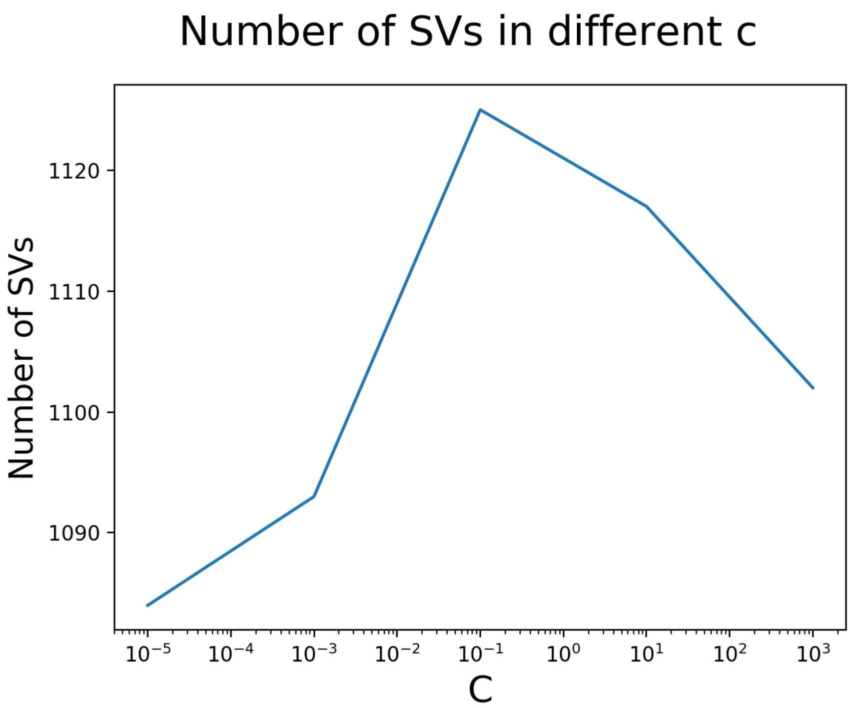
12.

The Eins in different C are all the same. The reason that causes this result might be the limitation of polynomial kernel. Since the kernel is not able to separate all data points, larger C won't help at all.



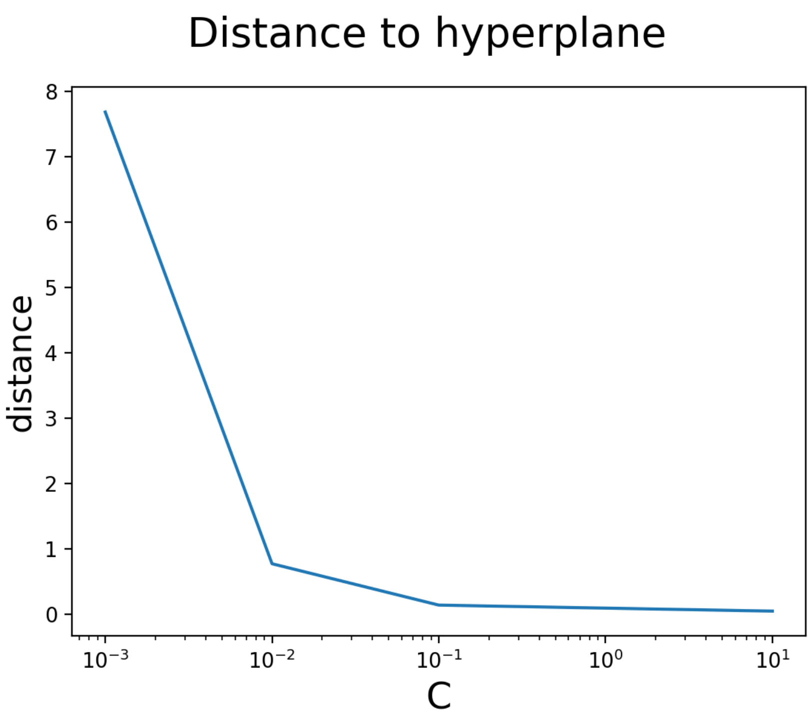
13.

When C = 0.1, it has the most support vectors. While C increases after it reaches 0.1, the number of support vectors starts to decreases. Although the difference in the figure might seem big, the number of SVs are actually very close. They are all around 1000 to 1100.



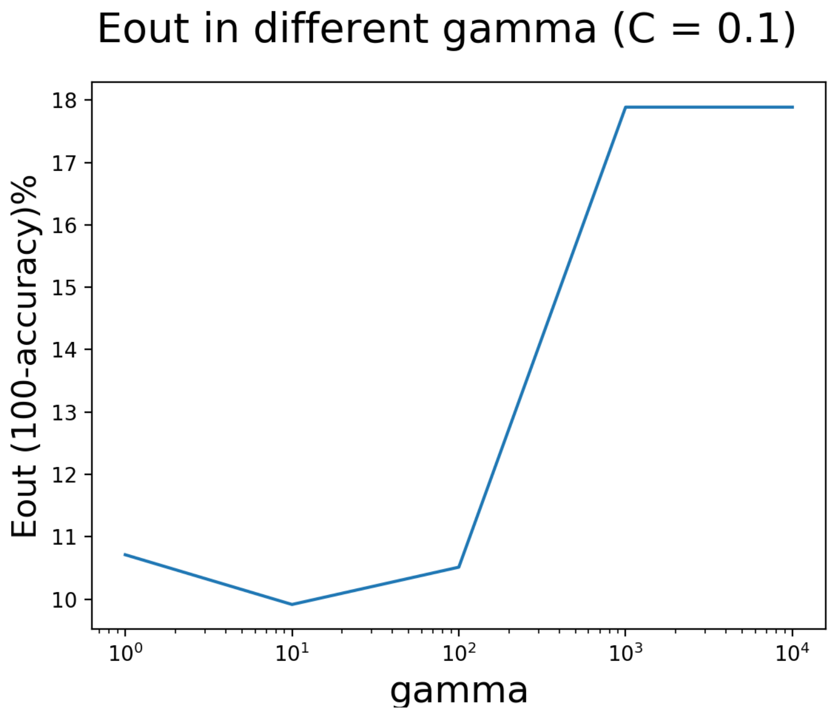
14.

The distance is big when C is small. When C is big, the hyperplane becomes complicated to separate the data points. Therefore, the distance turns smaller.



15.

We can see that when gamma = 10, it has the lowest Eout. When gamma becomes larger, the kernel becomes a “0 1” kernel. It is powerful, it can lead to overfitting. In this case, Eout grows sharply while gamma increases. This might be overfitting.



16.

This result matches with Problem 15. When C = 10, the Eval is often better than others. In problem 15, when C = 10, the Eout is the lowest.