



دانشگاه یزد

گروه مهندسی کامپیوتر

رشته تحصیلی: هوش مصنوعی و رباتیکز

نام درس: یادگیری ماشین

تکلیف شماره ۳

استاد مربوطه: دکتر مهدی یزدیان

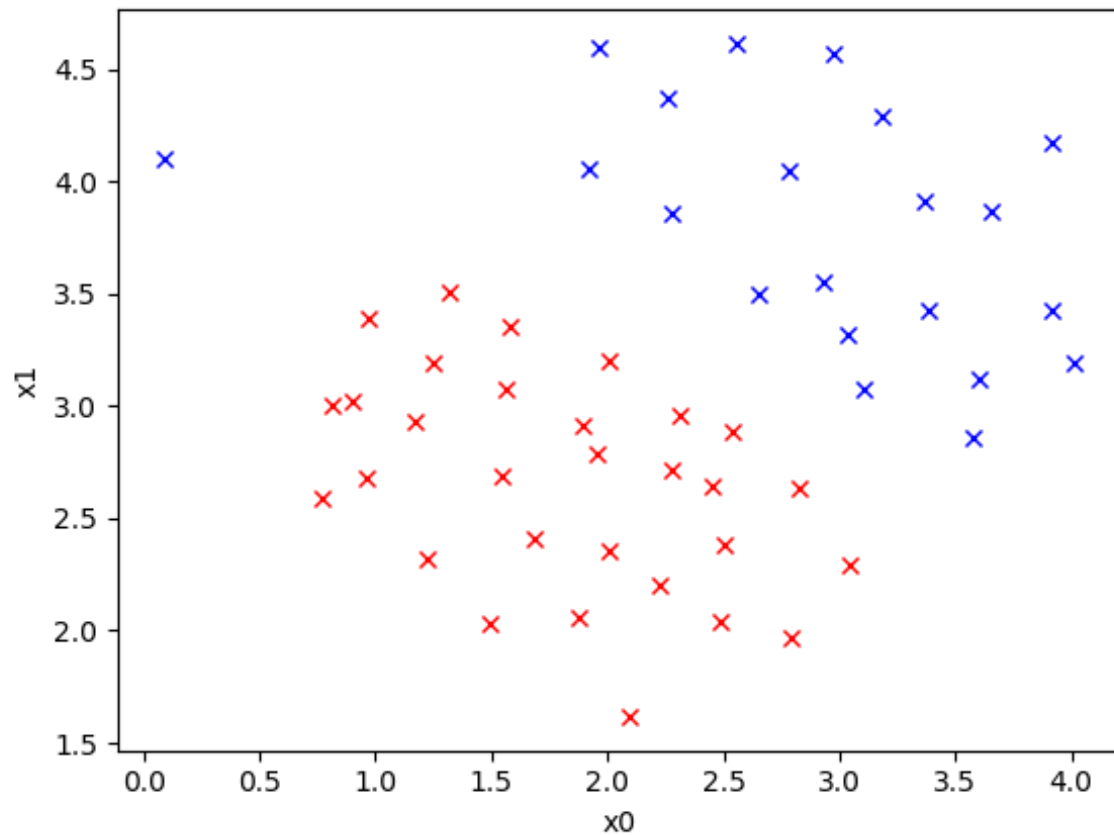
تهیه کننده: حمیدرضا نادمی

بهار ۱۳۹۸

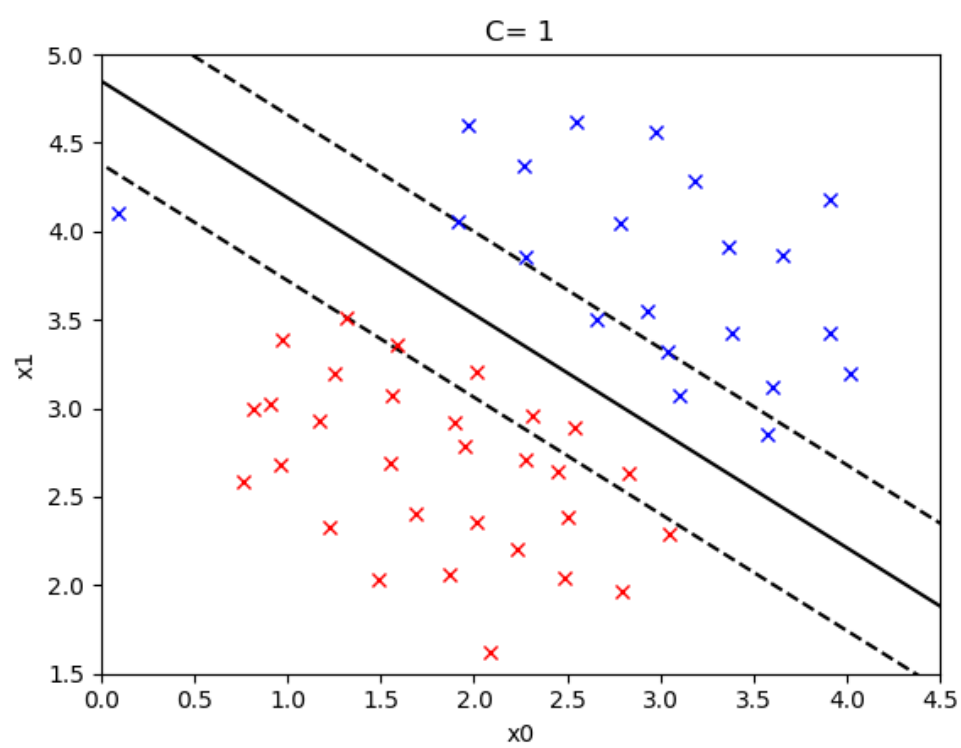
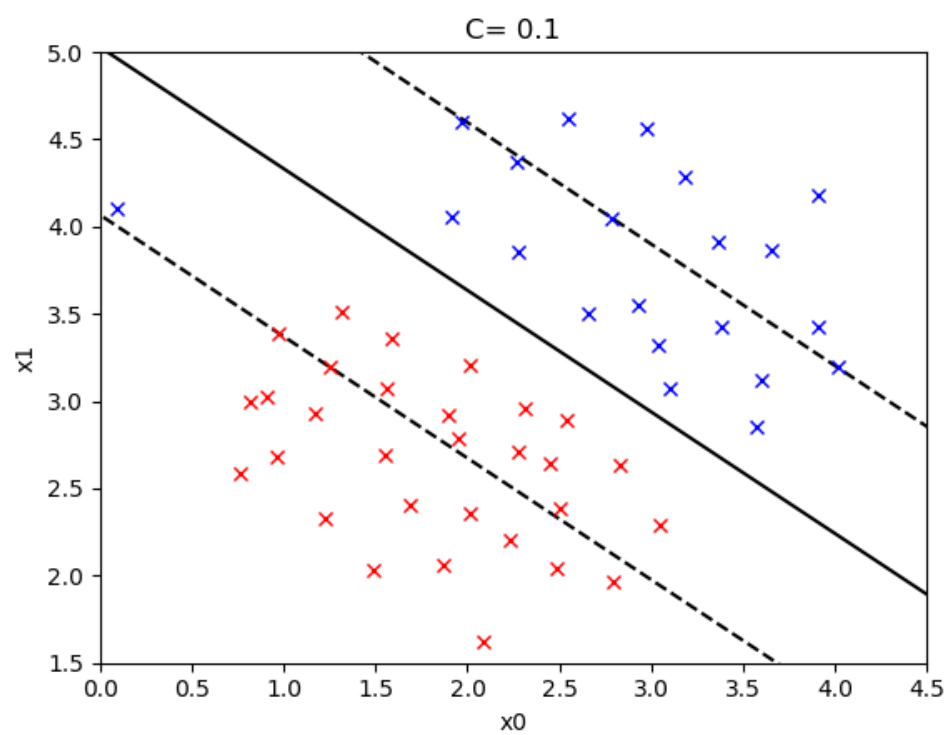
## **PART A. Linear SVM for Two-class Problem**

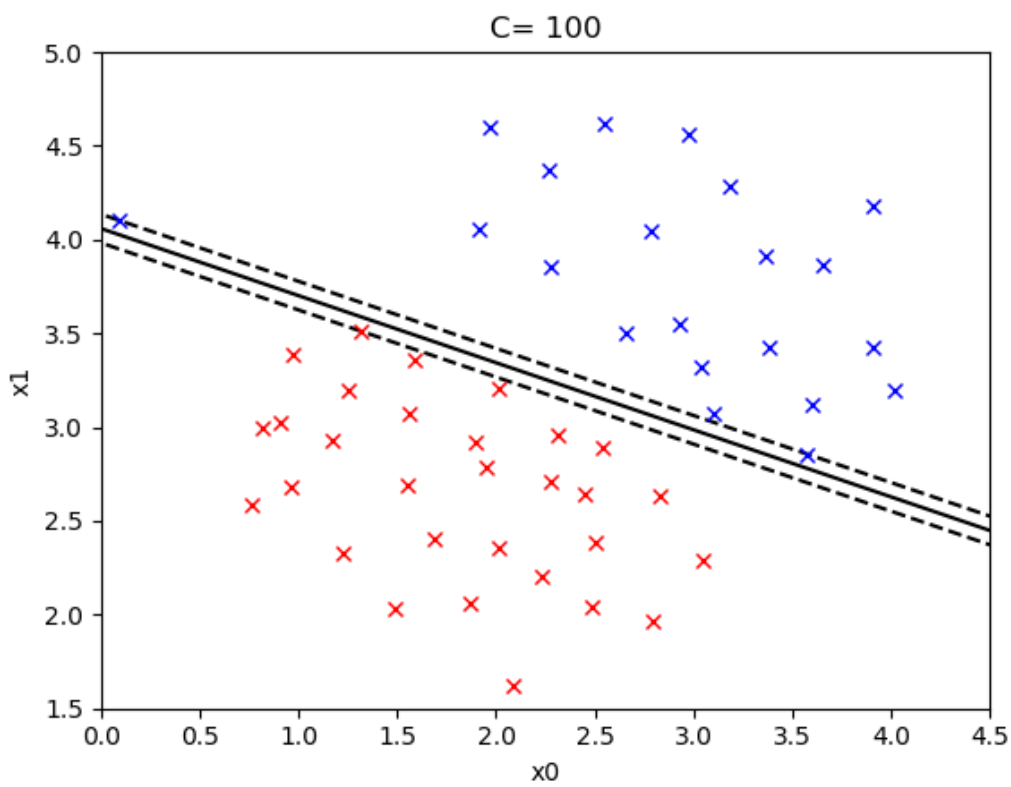
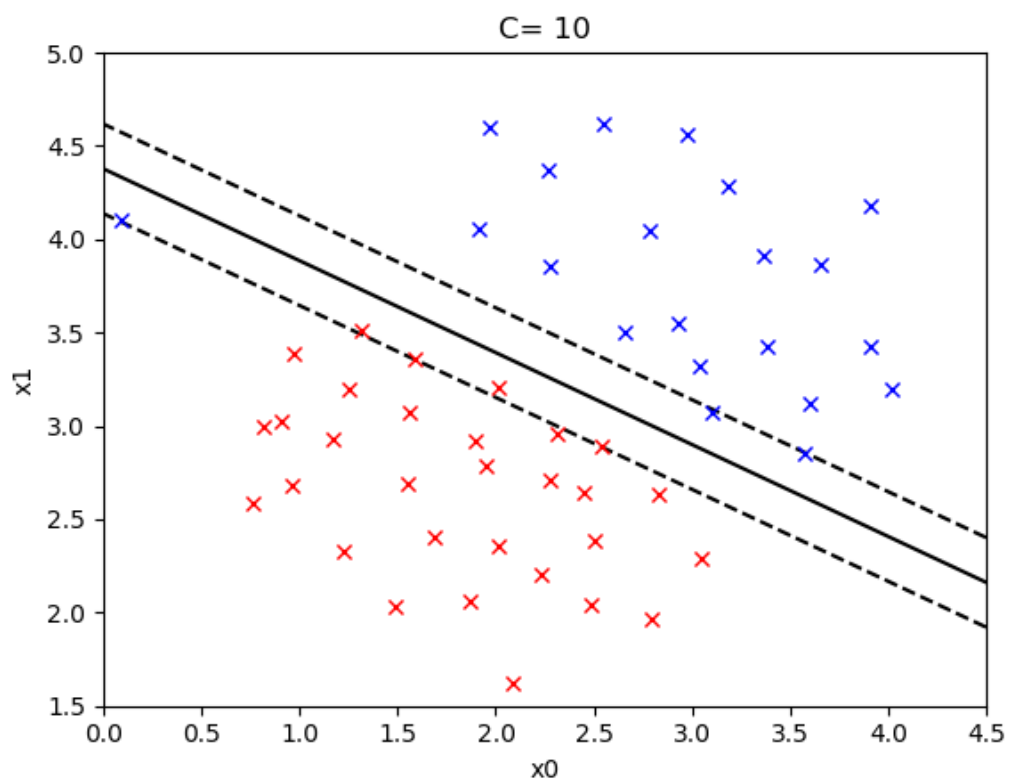
We will begin by a 2D dataset which can be separated by a linear boundary. In this part, you will try different values of the  $C$  parameter of SVM. Informally, the  $C$  parameter is a positive value that controls the penalty for misclassified training examples. A large  $C$  parameter tells the SVM to try to classify all the examples correctly.

1. Load and plot data1.mat to visualizing the dataset.

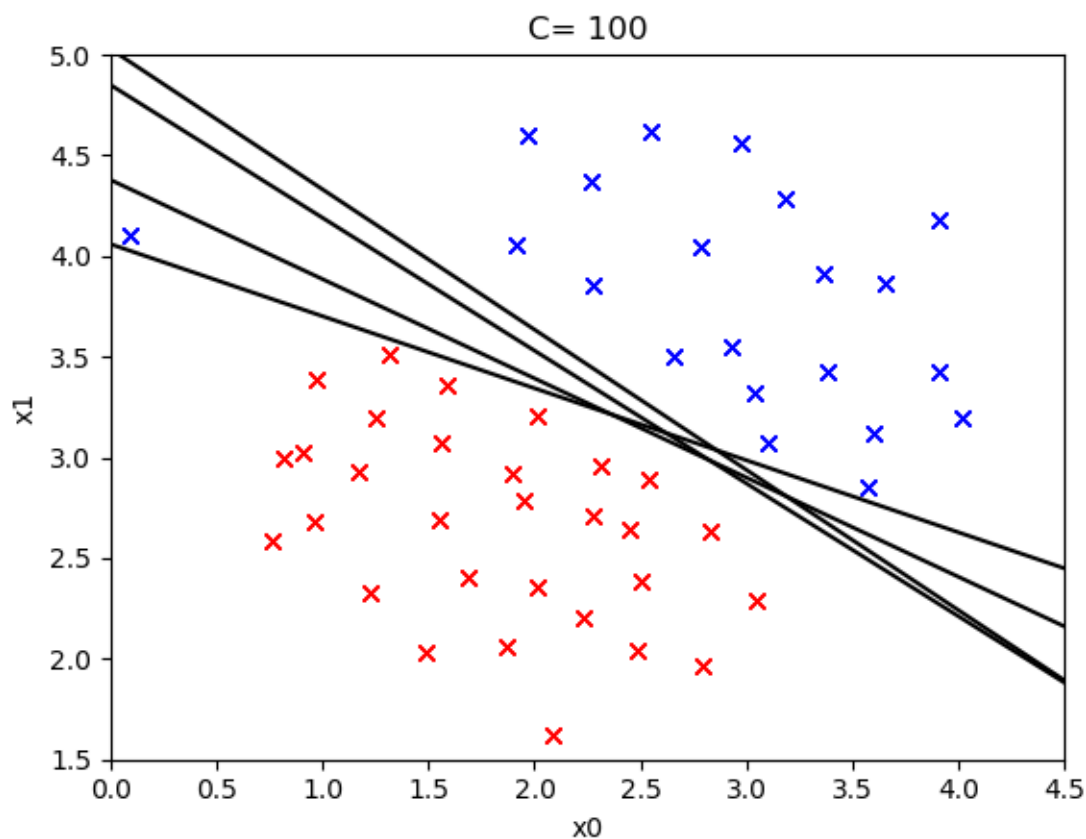


2. Train a linear SVM on the dataset. Try to use different values of  $C$  and see how the decision boundary varies. Use  $C=\{0.1, 1, 10, 100\}$ .





3. Plot different decision boundaries with different C on one figure.



4. Plot training accuracy of the classifier (y-axis) in terms of different values of C (x-axis)

C	Accuracy
0.1	98.039
1	98.039
10	98.039
100	100.0

5. Which value of C is the best value for this dataset? Explain the effect of C in training of SVM.

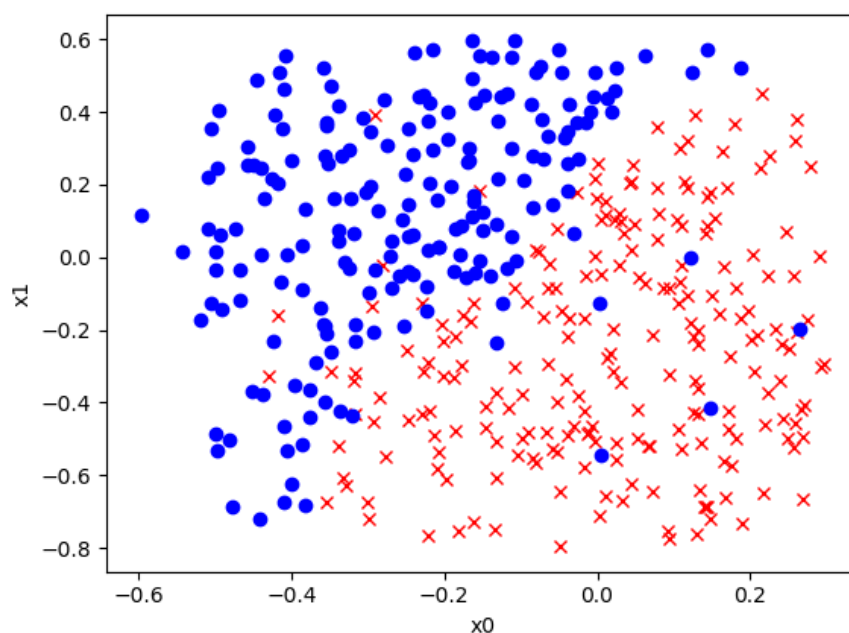
هر چه مقدار C بیشتر باشد، مارجین کوچکتر می شود و دقت مدل بیشتر می شود و Generalization کمتر می شود و overfitting می شود و دقت روی داده های آموزش ۱۰۰ درصد می شود، و برعکس هر چه مقدار C کمتر باشد Margin بزرگتر می شود و مدل به داده های بیشتری اجازه می دهد که در Margin قرار بگیرند و Generalization مدل بیشتر می شود ولی دقت روی داده های آموزش کمتر می شود (۹۸,۰۳۹٪).

6. Notice that there is an outlier positive example + on the far left. As part of this exercise, you will also see how this outlier affects the SVM decision boundary. How SVM is affected by this outlier.

SVM نسبت به outlier مقاوم است.

### PART B. Kernel SVM for Two-class Problem

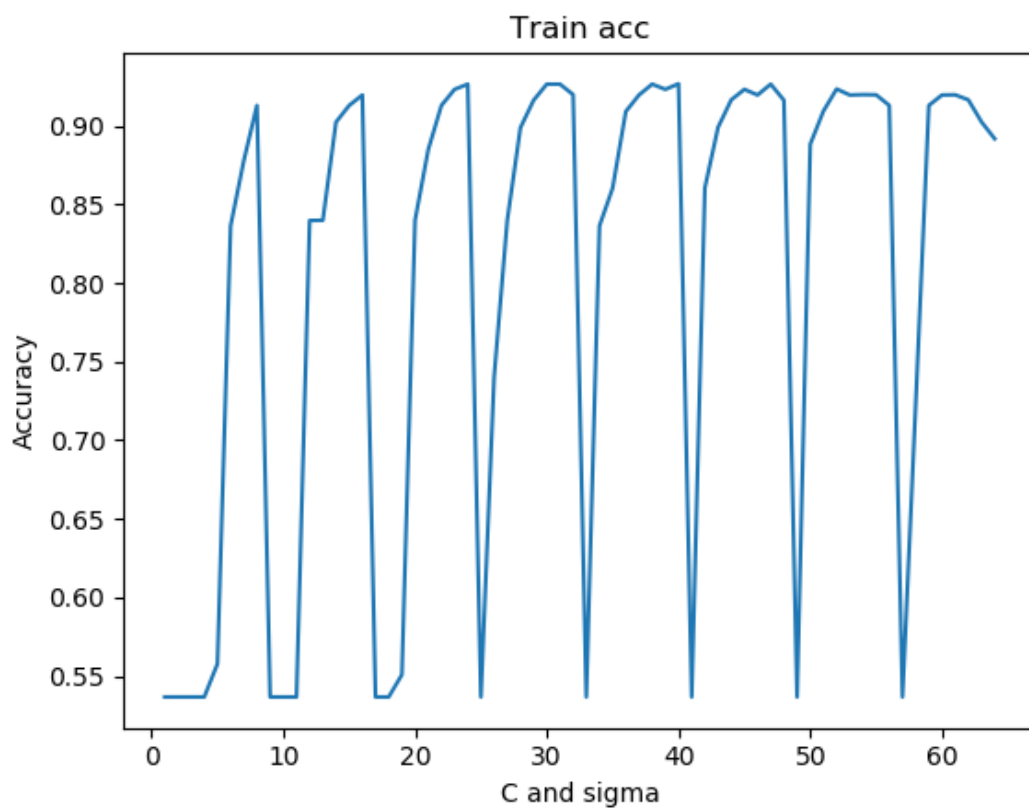
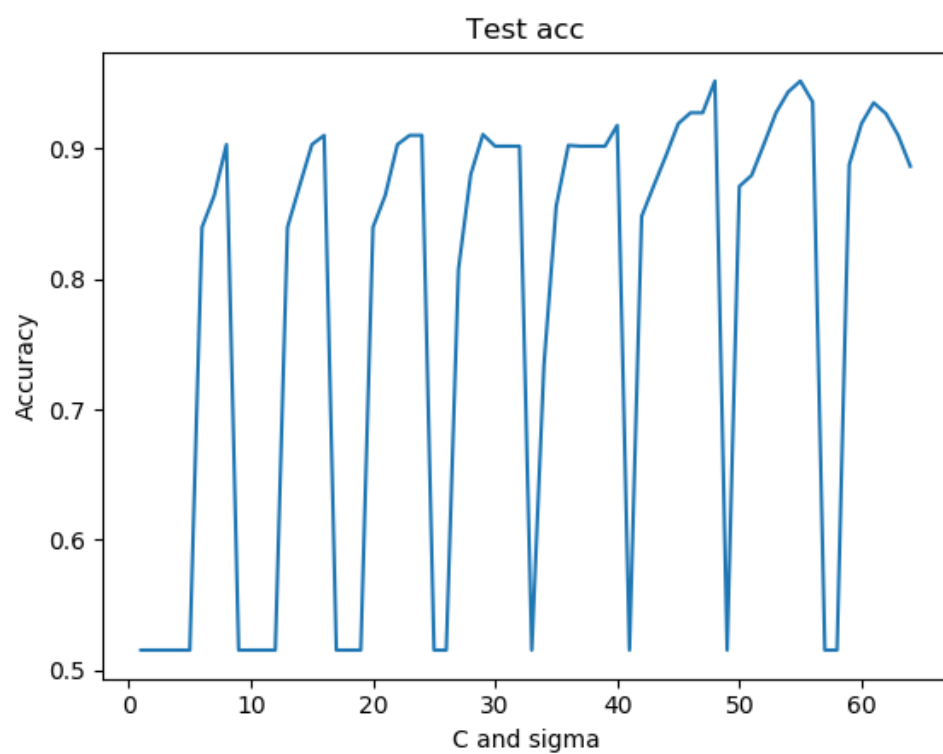
1. Load and plot data2.mat to visualize the dataset. From the figure, you can observe that there is no linear decision boundary that separates the positive and negative examples for this dataset. However, by using the Gaussian kernel with the SVM, you will be able to learn a non-linear decision boundary that can perform reasonably well for the dataset.

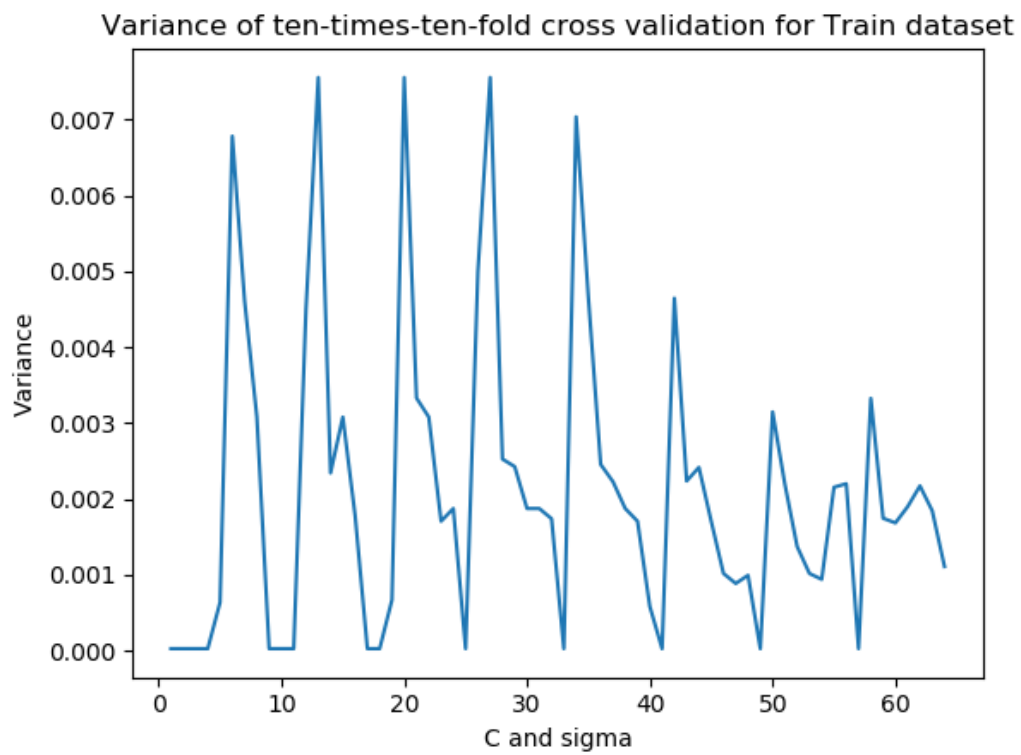
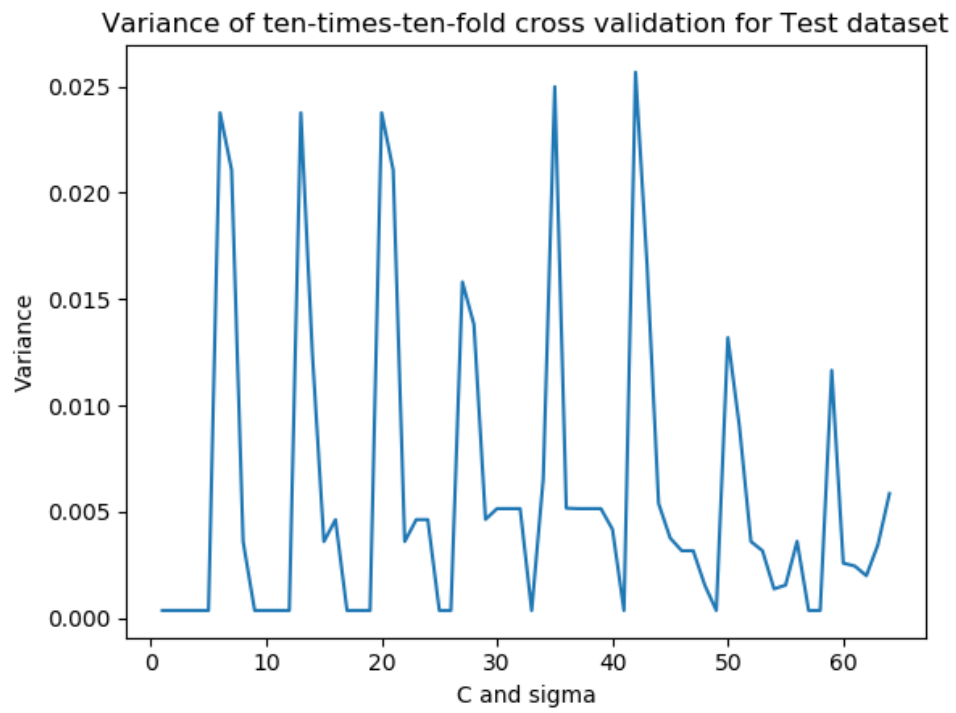


2. Train SVM with Gaussian kernel on this dataset. Find the best value of  $\sigma$  and  $C$  through ten-times-ten fold cross-validation. Note that it is better to test the values in multiplicative steps such as 0.01, 0.04, 0.1, 0.4, 1, 4, 10 and 40. Therefore, you should evaluate 64 different models to select the best model!

Best C	Best Sigma	Accuracy
<b>40</b>	<b>4</b>	<b>0.946</b>

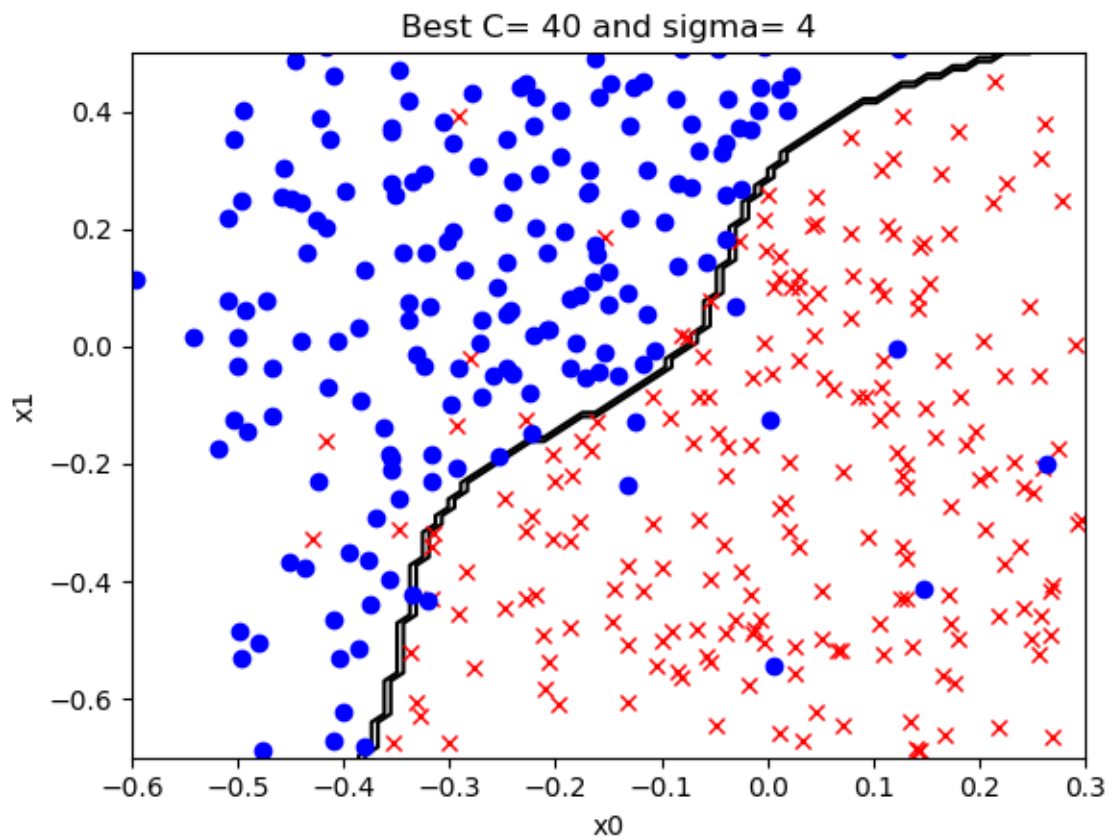
3. Plot the train and test accuracies and their corresponding variances of ten-times-ten-fold cross validation for different values of  $C$  and  $\sigma$ .





4. Plot the data and the decision boundary for the **best** C and .





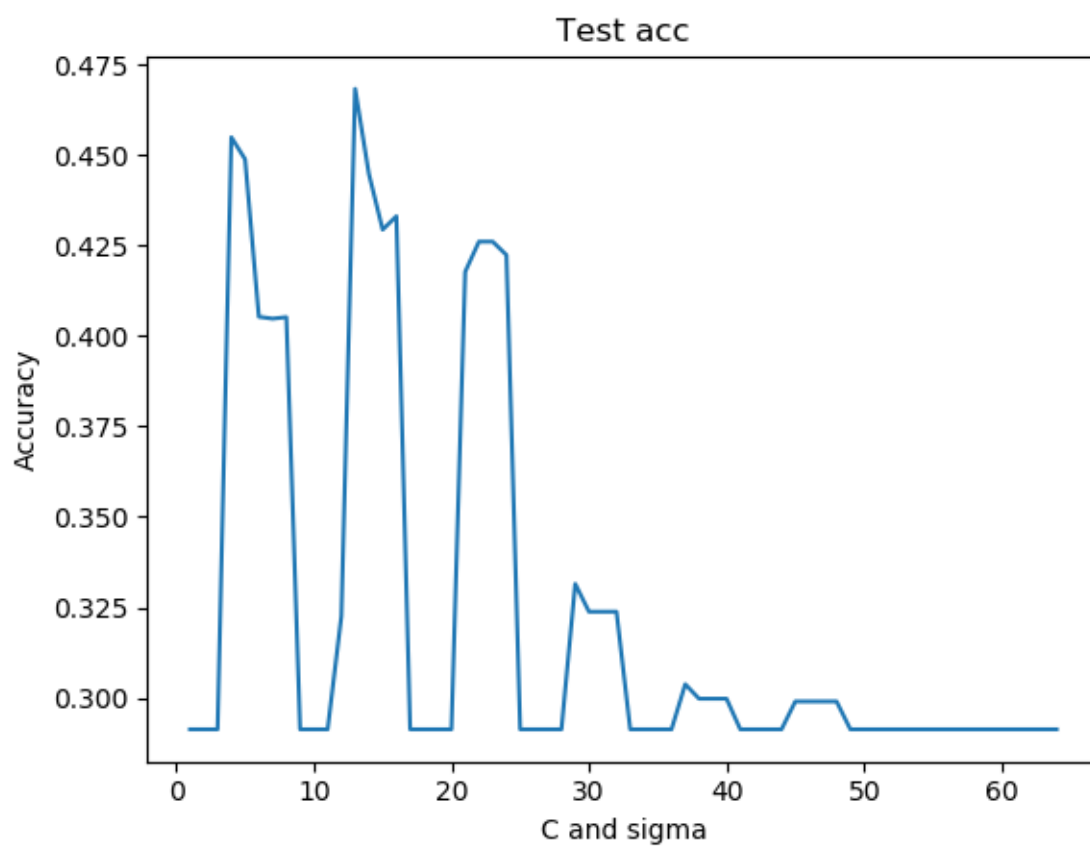
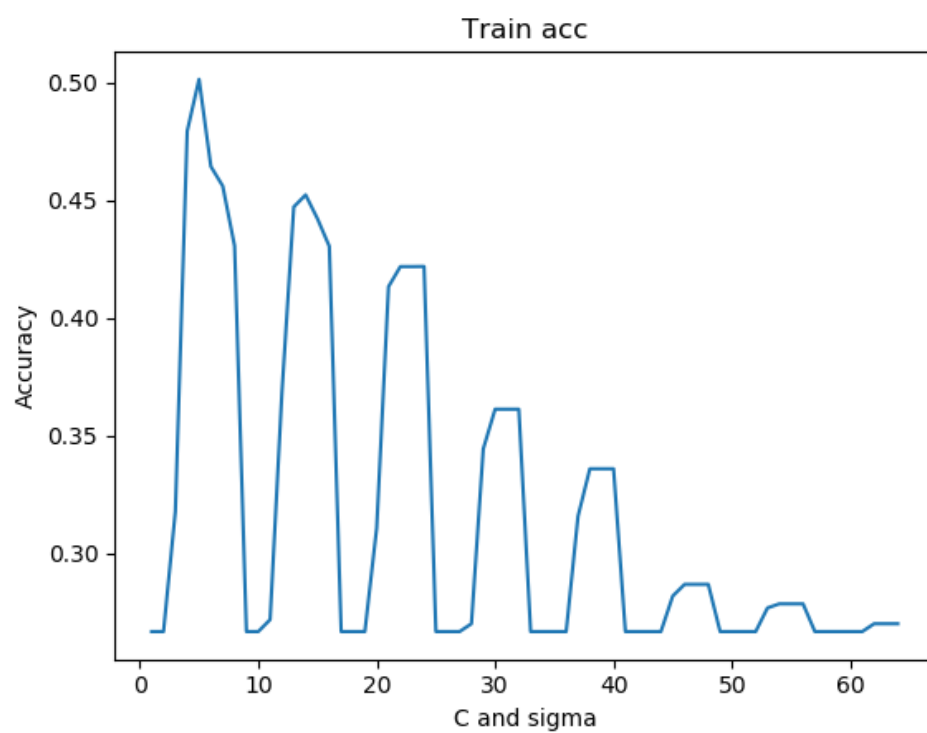
5. Report the test accuracy using the best selected model and discuss about the results.

Test accuracy	<b>0.951</b>
---------------	--------------

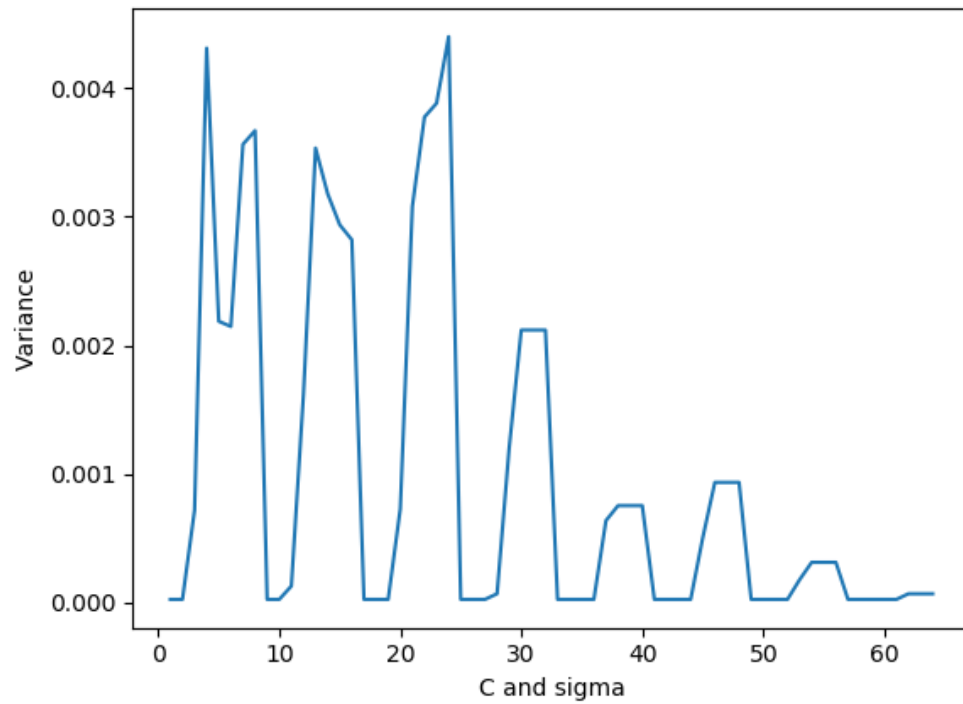
### PART C. Kernel SVM for Multi-class Problem

- a) Determine the best value of  $C$  and  $\sigma$  using ten-times-ten-fold cross validation. Plot the train and test accuracies and their corresponding variances of the ten-times-ten-fold cross validation for different values of  $C$  with and  $\sigma$ .

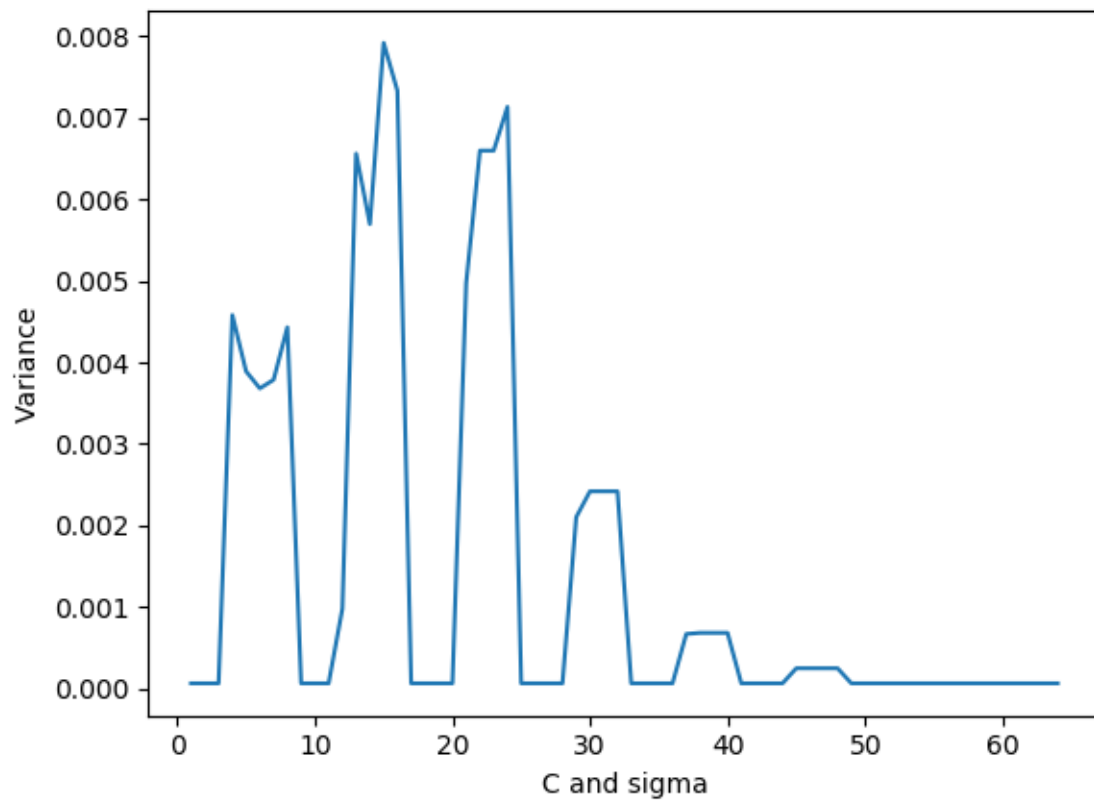
Best C	Best Sigma	Accuracy
<b>1</b>	<b>0.01</b>	<b>0.49</b>



Variance of ten-times-ten-fold cross validation for Train dataset

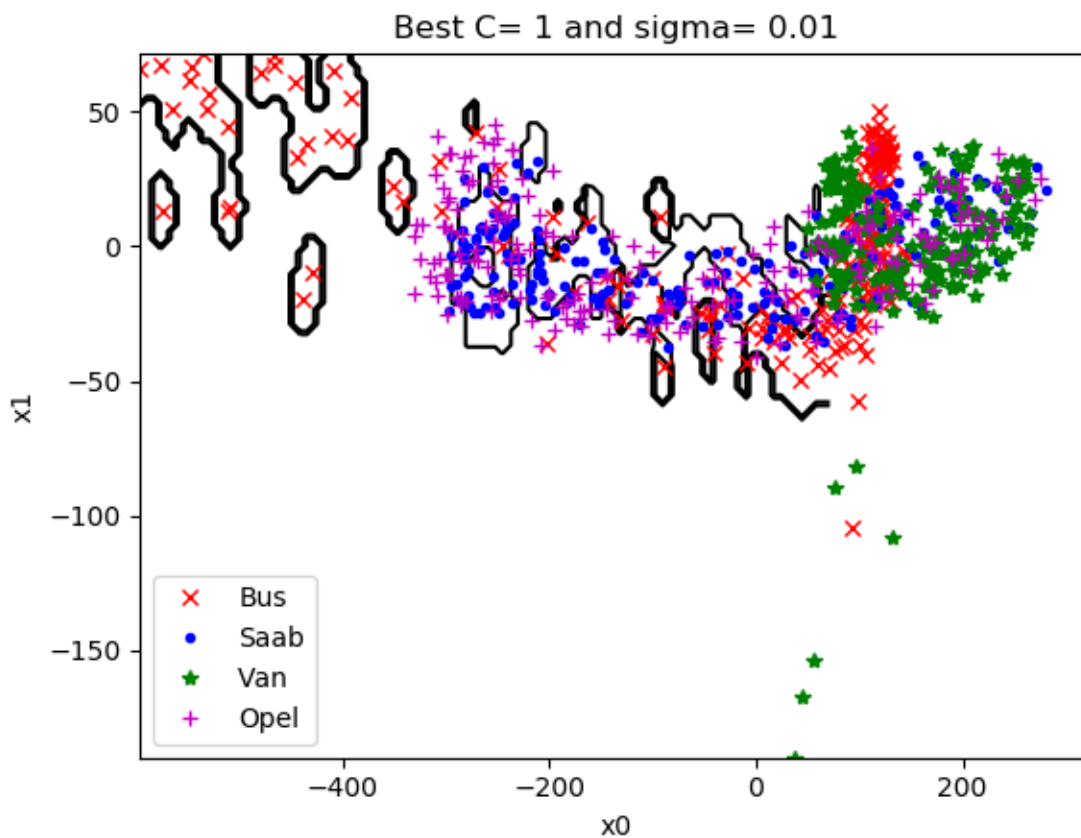


Variance of ten-times-ten-fold cross validation for Test dataset



b) Plot the data and the decision boundary of the **best**  $C$  and  $\sigma$ .

جهت رسم دیتاست در فضای ۲ بعدی با استفاده از PCA ابعاد هر داده را به ۲ کاهش دادیم.



c) Report the test accuracy using the best selected model and Discuss about the results.

Test accuracy	<b>0.448</b>
---------------	--------------

#### **Part D. Conclusion**

Please itemize your findings and your conclusions regarding the overall results of this homework.

- هرچه مقدار  $C$  بزرگتر باشد SVM به داده‌های کمتری اجازه ورود به مارجین را می‌دهد و عمومیت مدل کمتر می‌شود. هرچه مقدار  $C$  کوچکتر باشد داده‌های بیشتری اجازه ورود به مارجین را دارند و عمومیت مدل بیشتر می‌شود
- SVM نسبت به داده نویز و پرت مقاوم است.

- در Kernel-SVM رابطه خاصی بین  $C$  و سیگما وجود ندارد و با کراس ولیدیشن بهترین  $C$  و سیگما رو بدست میاریم
- اگر مقدار  $C$  کمتر از یک باشد با زیاد بودن سیگما خطا بیشتر می شود
- اگر مقدار سیگما بزرگ باشد RBF مانند SVM خطی عمل می کند