

Problem 1: Support Vector Machines

Instructions:

- 1. Please use this q1.ipynb file to complete SVM using cvxopt
- 2. You may create new cells for discussions or visualizations

```
In [3]: # Import modules
import numpy as np
import matplotlib.pyplot as plt
from cvxopt import matrix, solvers
```

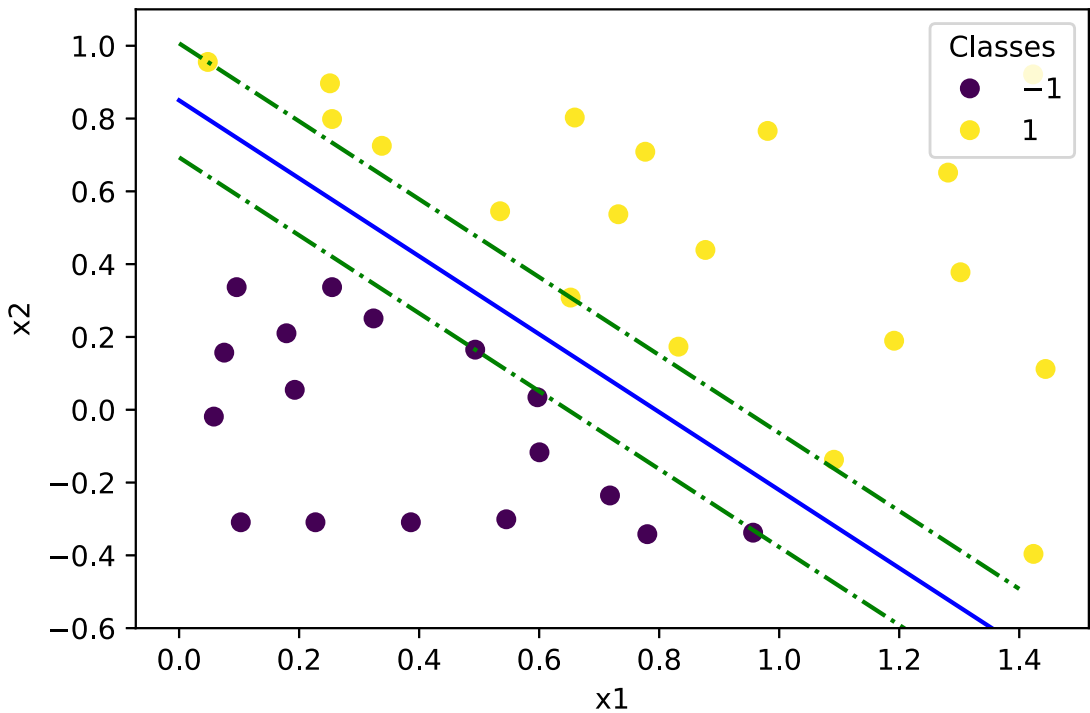
a): Linearly Separable Dataset

```
In [11]: data = np.loadtxt('clean_lin.txt', delimiter='\t')
x = data[:, 0:2]
y = data[:, 2].reshape(-1,1)
n = x.shape[0]
Q = matrix(np.eye((3)),tc = "d")
Q[2,2] = 0
p = matrix(np.zeros((3,1)),tc = "d")
h = matrix(-np.ones((n,1)),tc = "d")
ones = np.ones((n)).reshape(-1,1)
G1 = (-y) * np.hstack((x,ones))
G = matrix(G1,tc = "d")
solvers.options['show_progress'] = False
sol = solvers.qp(Q, p, G, h)
print(sol['x'])
z = sol['x']

# Plot points
x_point = np.linspace(0,1.4,100)
y_point = (-z[2]-z[0]* x_point) / z[1]
y_point1 = (1-z[2]-z[0]* x_point) / z[1]
y_point2 = (-1-z[2]-z[0]* x_point) / z[1]
plt.figure(1)
scatter = plt.scatter(x[:,0],x[:,1],c = y)
plt.xlabel("x1")
plt.ylabel("x2")
legend1 = plt.legend(*scatter.legend_elements(),loc="upper right", title="Classes")
plt.plot(x_point,y_point,c = "b")
plt.plot(x_point,y_point1,c = "green",linestyle="--")
plt.plot(x_point,y_point2,c = "green",linestyle="--")
plt.ylim(-0.6,1.1)
```

[6.83e+00]
[6.38e+00]
[-5.43e+00]

Out[11]: (-0.6, 1.1)



```
In [14]: print("Q matrix:",Q)
print("p matrix:",p)
print("G matrix:",G)
print("h matrix:",h)
```

[illegible]

Page 2 of 18

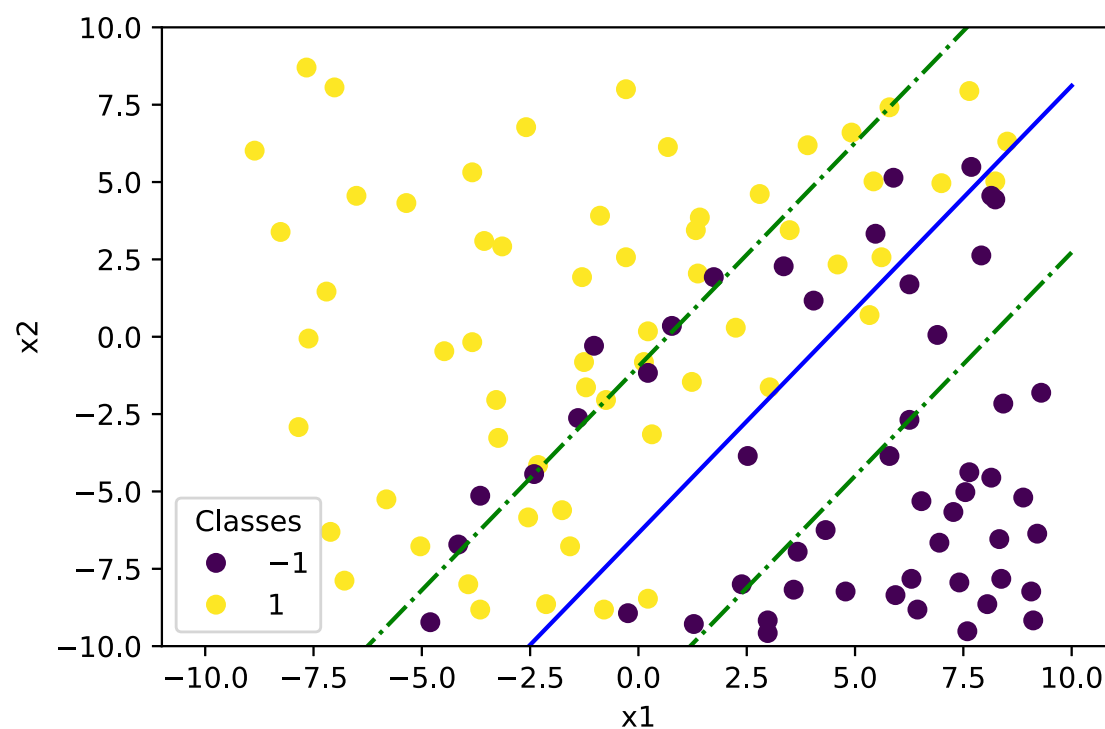
```

In [15]: # Load the data set that is not linearly separable
data = np.loadtxt('dirty_nonlin.txt', delimiter='\t')
x = data[:, 0:2]
y = data[:, 2].reshape(-1,1)
n = x.shape[0]
Q = np.zeros((n+3,n+3))
Q[0,0] = 1
Q[1,1] = 1
Q = matrix(Q,tc = "d")
p = np.ones((n+3,1))
p[0:3,:] = 0
p = matrix(0.05*p,tc = "d")
h = -np.ones((2*n,1))
h[n:2*n,:] = 0
h = matrix(h,tc = "d")
ones = np.ones((n)).reshape(-1,1)
G1 = (-y) * (np.hstack((x,ones)))
G2 = -np.eye((n))
G3 = np.zeros((n,3))
Ga = np.hstack((G1,G2))
Gb = np.hstack((G3,G2))
G = np.vstack((Ga,Gb))
G = matrix(G,tc = "d")
sol = solvers.qp(Q, p, G, h)
#print(sol['x'])
z = sol['x']

# Plot points
x_point = np.linspace(-10,10,200)
y_point = (-z[2]-z[0]* x_point) / z[1]
y_point1 = (1-z[2]-z[0]* x_point) / z[1]
y_point2 = (-1-z[2]-z[0]* x_point) / z[1]
plt.figure(1)
scatter = plt.scatter(x[:,0],x[:,1],c = y)
plt.xlabel("x1")
plt.ylabel("x2")
legend1 = plt.legend(*scatter.legend_elements(),loc="lower left", title="Classes")
plt.plot(x_point,y_point,c = "b")
plt.plot(x_point,y_point1,c = "green",linestyle="-.")
plt.plot(x_point,y_point2,c = "green",linestyle="-.")
plt.ylim(-10,10)

```

Out[15]: (-10.0, 10.0)



```

In [16]: print("Q matrix:",Q)
print("p matrix:",p)
print("G matrix:",G)
print("h matrix:",h)

Q matrix: [ 1.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  1.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]
[ 0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00  0.00e+00 ... ]

```

q/Desktop/AIML/HW6/template util function data/24787 HW6 Q1.html

[illegible]

```
G matrix: [-8.51e+00 -6.31e+00 -1.00e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-8.24e+00 -5.02e+00 -1.00e+00 -0.00e+00 -1.00e+00 -0.00e+00 -0.00e+00 ... ]
[-6.99e+00 -4.96e+00 -1.00e+00 -0.00e+00 -0.00e+00 -1.00e+00 -0.00e+00 ... ]
[-5.61e+00 -2.57e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -1.00e+00 ... ]
[-4.60e+00 -2.34e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-3.49e+00 -3.45e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-1.32e+00 -3.45e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 1.30e+00 -1.93e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 3.15e+00 -2.92e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 3.84e+00 1.75e-01 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 3.28e+00 2.04e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 1.26e+00 8.18e-01 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-2.19e-01 -1.75e-01 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-1.37e+00 -2.04e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 4.48e+00 4.67e-01 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 7.62e+00 5.84e-02 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 7.85e+00 2.92e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 7.11e+00 6.31e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 6.79e+00 7.88e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 5.03e+00 6.77e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 2.55e+00 5.84e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 3.24e+00 3.27e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 1.21e+00 1.64e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-3.11e-01 3.15e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-2.25e+00 -2.92e-01 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-5.33e+00 -7.01e-01 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-3.91e+00 -6.19e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 2.88e-01 -8.00e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 2.59e+00 -6.77e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 7.02e+00 -8.06e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 7.66e+00 -8.70e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 8.86e+00 -6.01e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 8.26e+00 -3.39e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 5.36e+00 -4.32e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 7.20e+00 -1.46e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 6.51e+00 -4.55e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 3.84e+00 -5.31e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 8.87e-01 -3.91e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-1.42e+00 -3.85e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-2.80e+00 -4.61e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-4.92e+00 -6.60e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-5.79e+00 -7.42e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-7.64e+00 -7.94e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-5.43e+00 -5.02e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-1.27e-01 8.18e-01 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-1.23e+00 1.46e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[-3.03e+00 1.64e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 7.49e-01 2.04e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 2.32e+00 4.15e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 1.76e+00 5.61e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 1.58e+00 6.77e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 2.13e+00 8.64e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
[ 3.65e+00 8.82e+00 -1.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 -0.00e+00 ... ]
```

g/Desktop/AIML/HW6/template_util_function_data/24787_HW6_Q1.html

```
h matrix: [-1.00e+00]
[-1.00e+00]
[-1.00e+00]
[-1.00e+00]
[-1.00e+00]
[-1.00e+00]
[-1.00e+00]
[-1.00e+00]
[-1.00e+00]
```

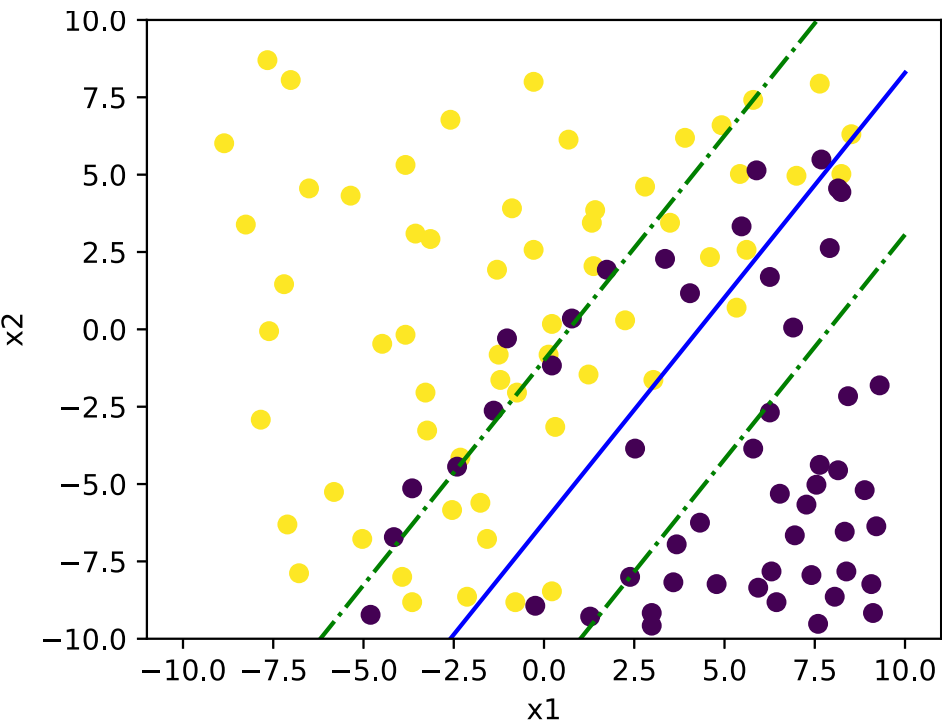


```
[-1.00e+00]
```

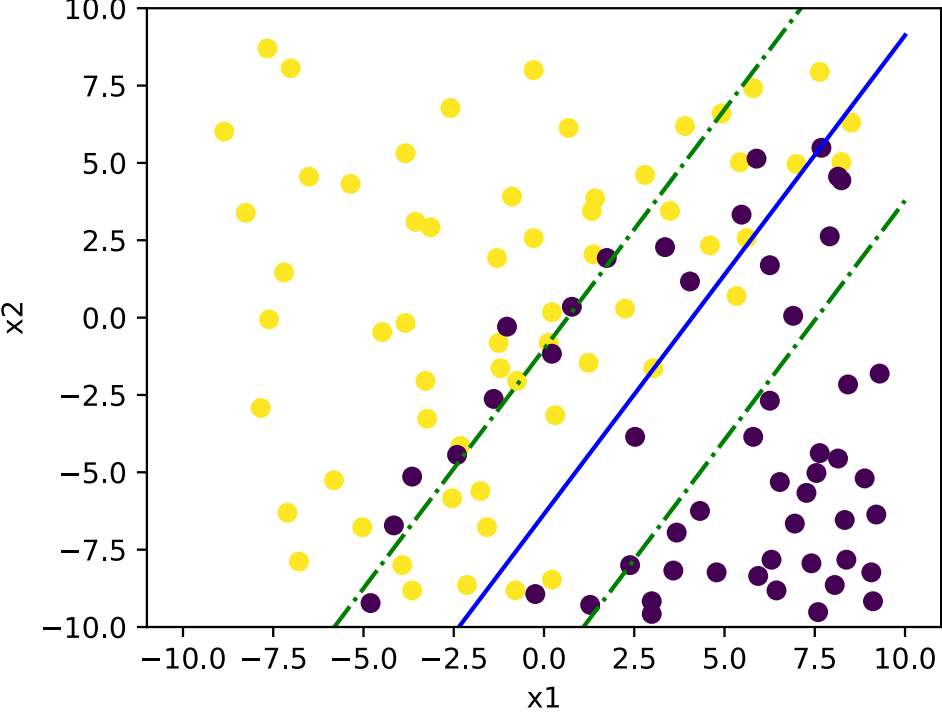
[illegible]

c): Output 4 plots & Explain your observations here:

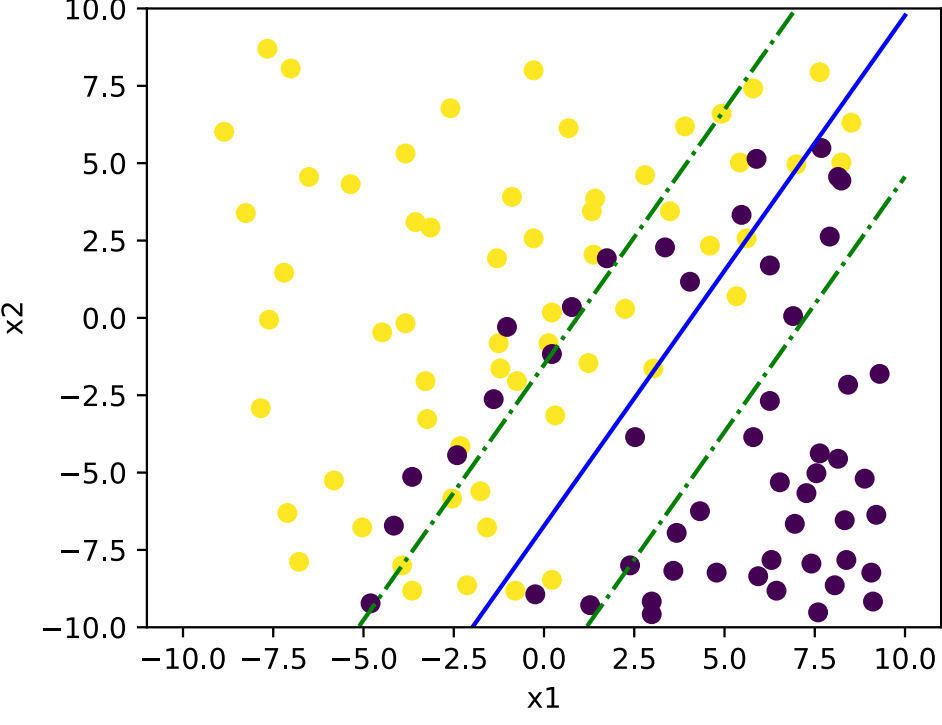
```
fig1 = svm(data,0.1)
fig2 = svm(data,1)
fig3 = svm(data,100)
fig4 = svm(data,1000000)
```

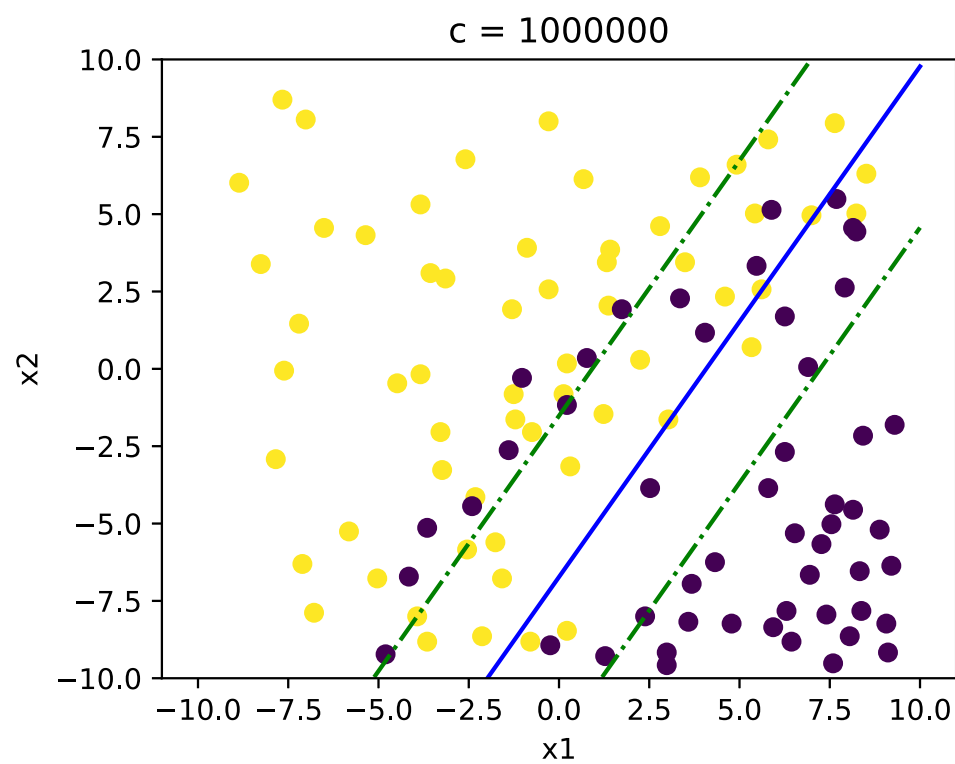



$c = 1$



$c = 100$





I think there are not huge differences between these four figures. Generally, with the increase of C , the constraints are harder to satisfy, the width between margins will become more narrow. However, this is a linearly non-separable case, no matter how hard the constraints are, these classes cannot be linearly separate, which means there aren't big differences between different C values.

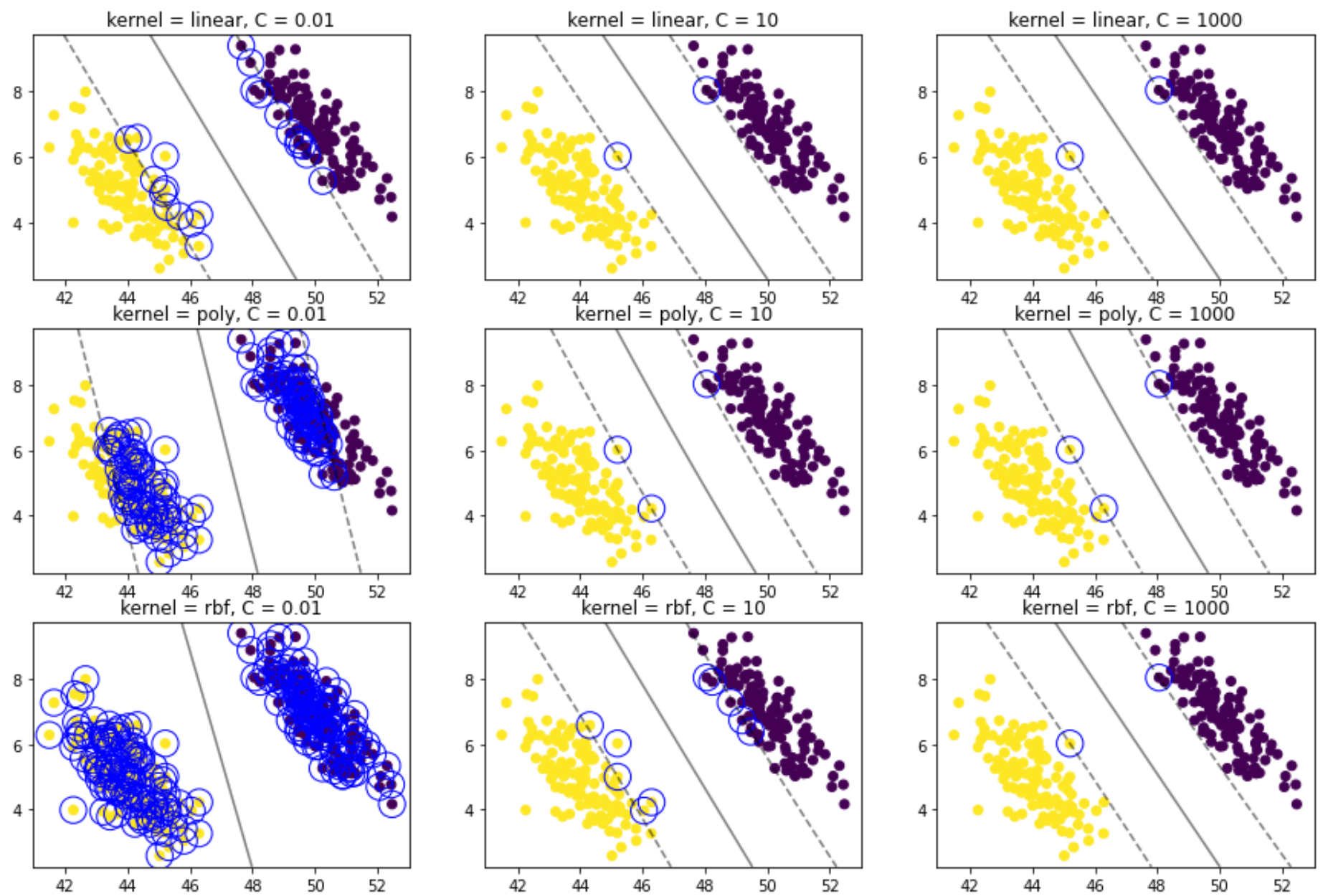
Problem 2: Support Vector Machine Experiments with sklearn

(a)

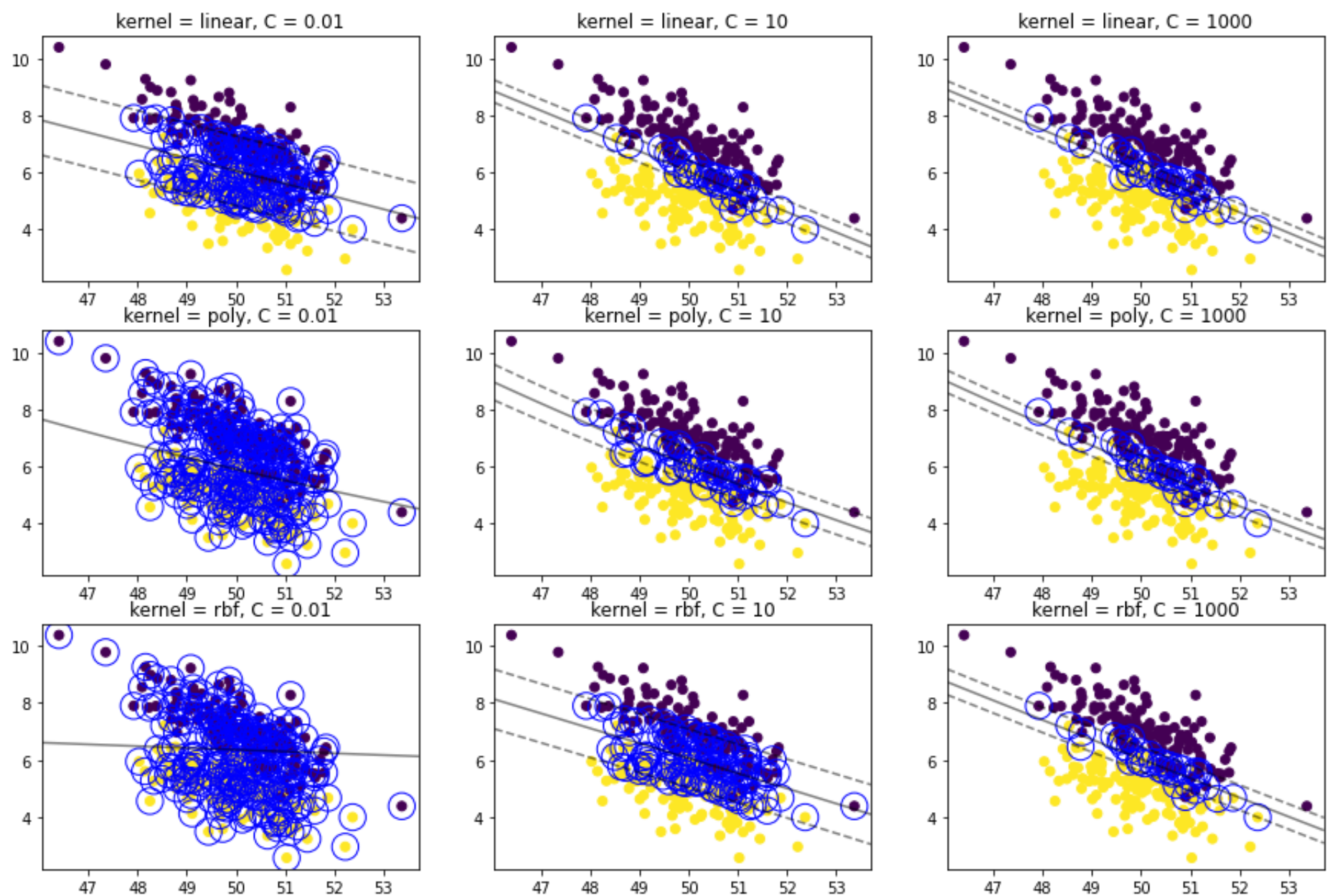
```
In [1]: from sklearn.svm import SVC
        from plot_svc_decision_function import plot_svc_decision_function
```

```
In [3]: def svm(data):
        data = np.loadtxt(data)
        x = data[:, 0:2]
        y = data[:, 2]
        fig, axs = plt.subplots(3,3,figsize = (15,10))
        kernel = ["linear","poly","rbf"]
        C = [0.01,10,1000]
        for i in range(3):
            for j in range(3):
                model = SVC(kernel = kernel[i],C = C[j])
                model.fit(x,y)
                axs[i,j].scatter(x[:,0],x[:,1],c = y)
                axs[i,j].set_title(f"kernel = {kernel[i]}, C = {C[j]}")
                plot_svc_decision_function(model,axs[i,j])
```

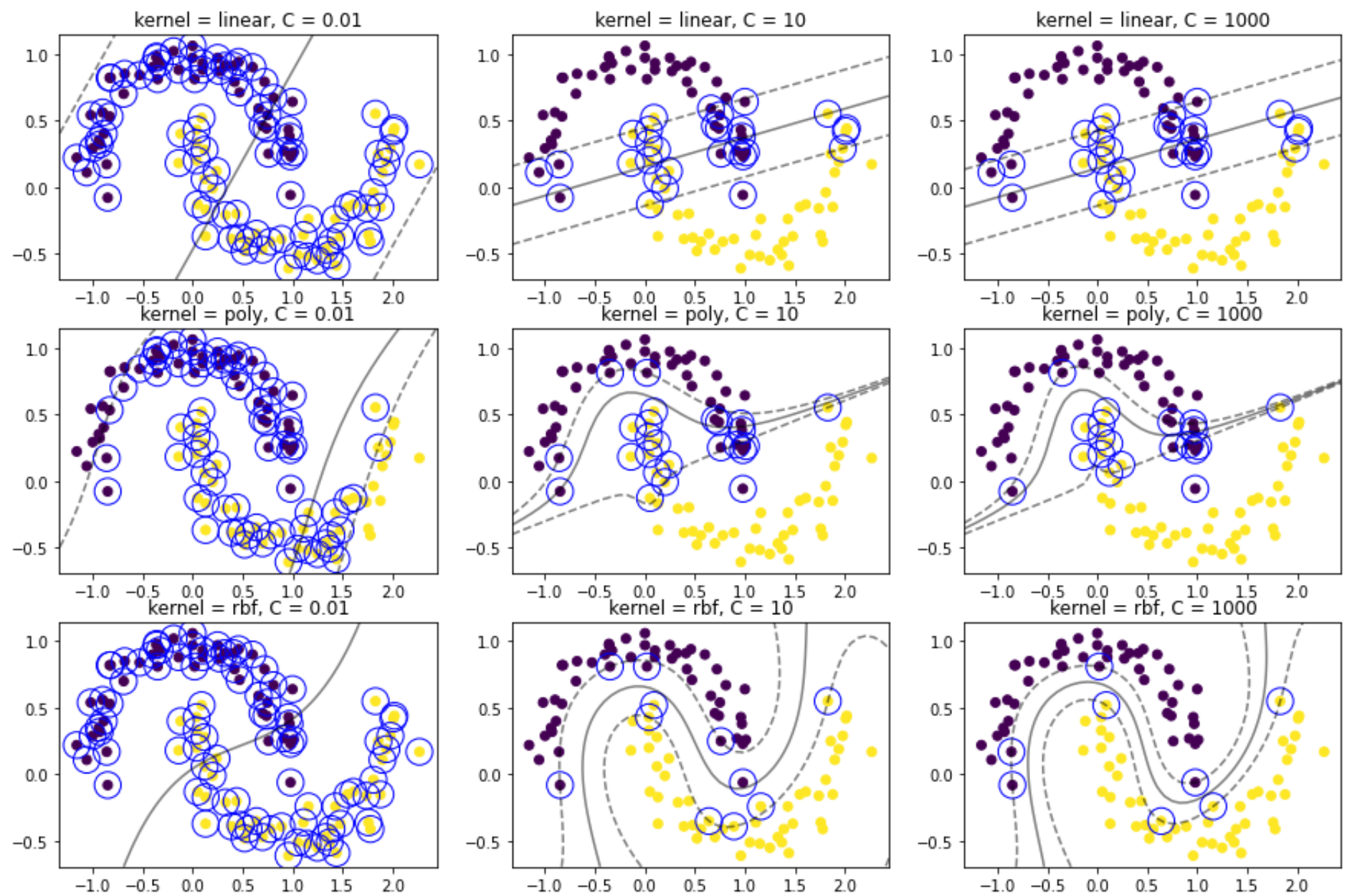
```
In [4]: svm('binary_linsep_yes_n240.txt')
```



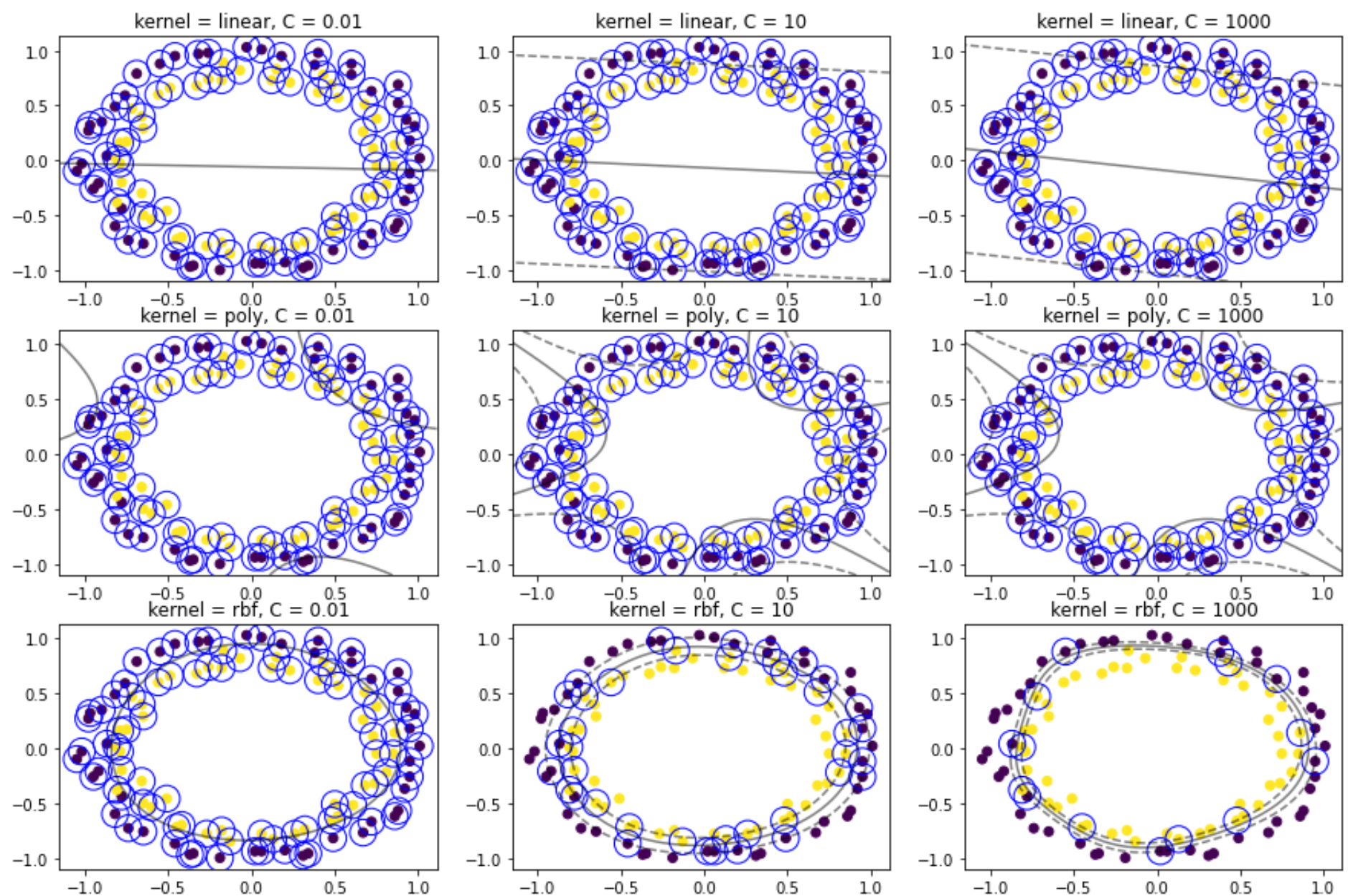
```
In [5]: svm('binary_linsep_no_n240.txt')
```



```
In [6]: svm('binary_moons_linsep_no_n100.txt')
```

```
In [7]: svm('concentriccircles_binary.txt')
```



```
In [8]: data = np.loadtxt('binary_moons_linsep_no_n100.txt')
x = data[:, 0:2]
y = data[:, 2]
model = SVC(kernel = "rbf", C = 1000)
model.fit(x,y)
model.support_vectors_
```

```
Out[8]: array([[ -0.86,  0.17],
               [  0.02,  0.81],
               [  0.98, -0.06],
               [-0.85, -0.08],
               [  0.64, -0.35],
               [  0.09,  0.52],
               [  1.83,  0.55],
               [  1.16, -0.24]])
```

(b)

Kernel:

for the first datasets, linear kernel is better because it takes less time.

for the second datasets, three types of kernel can't linearly separate the datasets. Therefore, I think the simple one is better.

for the third and fourth datasets, RBF did better.

Conclusion: I think we should choose different kernel according to different datasets.

C value:

The larger the C, the greater punishment SVM receives when it commits a misclassification. As a result, the decision boundary will rely on fewer support vectors as the margin becomes narrower.

Moons dataset

I think the model with 5 support vectors is better. Because these two models have identical classification accuracy, more support vectors means it's more expensive to train model. Therefore, less support vectors model is better.

Problem 3: Support Vector Regression

(a)

```
In [246... from sklearn.svm import SVR
data = np.loadtxt('train.txt')
test = np.loadtxt("test.txt")
x = data[:, 0:2]
y = data[:, 2]
test_x = test[:,0:2]
test_y = test[:,2]
mu_x = np.mean(x,axis = 0)
std_x = np.std(x,axis = 0)
x_new = (x - mu_x)/std_x
test_new = (test_x - mu_x)/std_x
model = SVR(kernel = "rbf",C = 1, epsilon = 0.0001)
model.fit(x_new,y)
y_pred = model.predict(test_new)
rmse = np.sqrt(np.mean((y_pred.reshape(-1,1))-test_y.reshape(-1,1))**2,axis = 0))
rmse
```

```
Out[246... array([24.03386007])
```

(b)

```
In [236... def svr(C,epsilon):
    model = SVR(kernel = "rbf",C = C, epsilon = epsilon)
    model.fit(x_new,y)
    y_pred = model.predict(test_new)
    rmse = np.sqrt(np.mean((y_pred.reshape(-1,1))-test_y.reshape(-1,1))**2,axis = 0))
    return rmse
```



```
In [242... print("C = 0.01, epsilon = 0.0001, RMSE = ",svr(0.01,0.0001))
print("C = 0.01, epsilon = 0.01, RMSE = ",svr(0.01,0.01))
print("C = 0.01, epsilon = 1, RMSE = ",svr(0.01,1))
print("C = 10, epsilon = 0.0001, RMSE = ",svr(10,0.0001))
print("C = 10, epsilon = 0.01, RMSE = ",svr(10,0.01))
print("C = 10, epsilon = 1, RMSE = ",svr(10,1))
print("C = 100, epsilon = 0.0001, RMSE = ",svr(100,0.0001))
print("C = 100, epsilon = 0.01, RMSE = ",svr(100,0.01))
print("C = 100, epsilon = 1, RMSE = ",svr(100,1))
```

```
C = 0.01, epsilon = 0.0001, RMSE = [49.86230822]
C = 0.01, epsilon = 0.01, RMSE = [49.86230822]
C = 0.01, epsilon = 1, RMSE = [49.71919847]
C = 10, epsilon = 0.0001, RMSE = [7.2066139]
C = 10, epsilon = 0.01, RMSE = [7.20308351]
C = 10, epsilon = 1, RMSE = [7.19590857]
C = 100, epsilon = 0.0001, RMSE = [4.35615181]
C = 100, epsilon = 0.01, RMSE = [4.35497666]
C = 100, epsilon = 1, RMSE = [4.18971781]
```

C value

From these number, I can figure out when C becomes bigger, RMSE will decrease. But this doesn't mean that bigger C is better. When C is too big, the model might be overfitted.

Epsilon

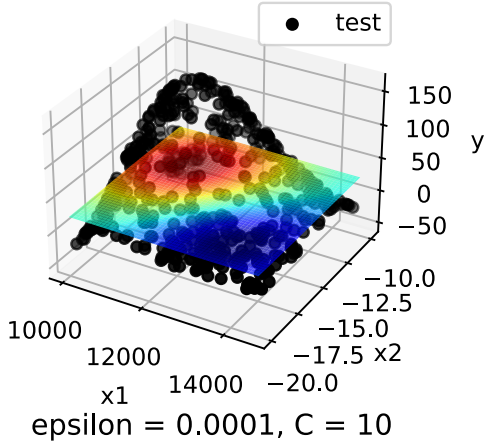
The value of epsilon defines a margin of tolerance where no penalty is given to errors. From numbers above, I can see there is no huge difference between the models which have the same C value.

(c)

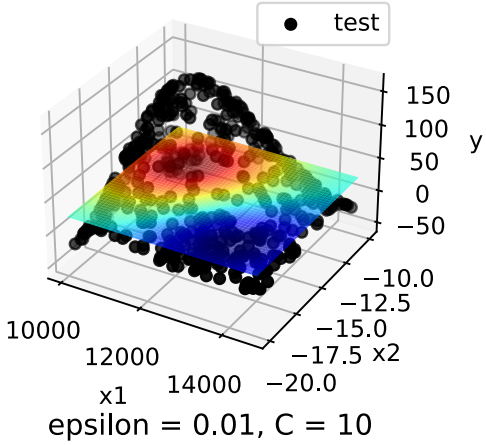
```
In [256... #from mpl_toolkits.mplot3d import Axes3D
x1 = np.linspace(np.min(test_x[:,0]),np.max(test_x[:,0]),100)
x2 = np.linspace(np.min(test_x[:,1]),np.max(test_x[:,1]),100)
x1_new, x2_new = np.meshgrid(x1,x2)
std_x1 = ((x1_new.ravel()).reshape(-1,1)-mu_x[0])/std_x[0]
std_x2 = ((x2_new.ravel()).reshape(-1,1)-mu_x[1])/std_x[1]
X_new = np.hstack((std_x1,std_x2))
```

```
In [255... fig = plt.figure(figsize=(15,10))
epsilon = [0.0001,0.01,1]
C = [0.01,10,100]
for i in range(3):
    for j in range(3):
        model = SVR(kernel = "rbf",C = C[i], epsilon = epsilon[j])
        model.fit(x_new,y)
        y_pred = model.predict(X_new)
        #axs[i,j] = Axes3D(fig)
        ax = fig.add_subplot(3,3,3*i+j+1, projection='3d')
        ax.scatter(test_x[:,0].reshape(-1,1),test_x[:,1].reshape(-1,1),test_y.reshape(-1,1),c='black',label="test")
        ax.plot_surface(x1_new,x2_new,y_pred.reshape(x1_new.shape),cmap='jet',alpha = 0.7)
        ax.set_xlabel="x1",ylabel="x2",zlabel="y"
        ax.legend()
        ax.set_title(f"epsilon = {epsilon[j]}, C = {C[i]}")
```

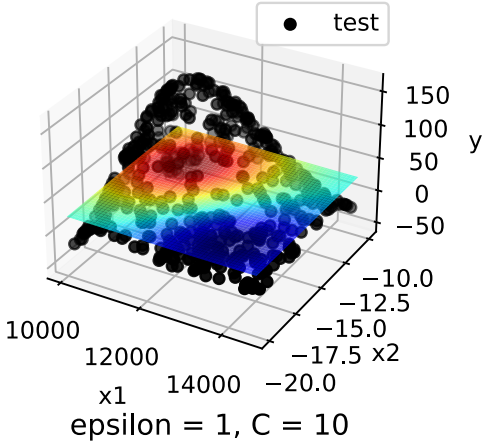
epsilon = 0.0001, C = 0.01



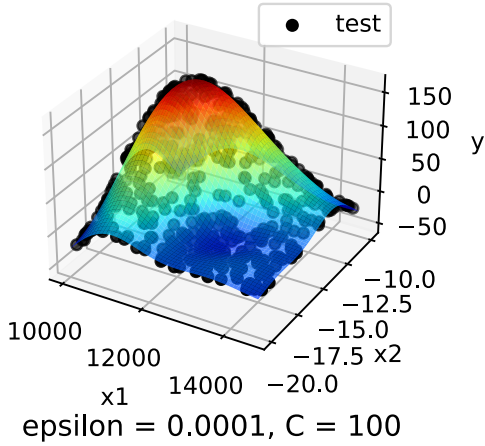
epsilon = 0.01, C = 0.01



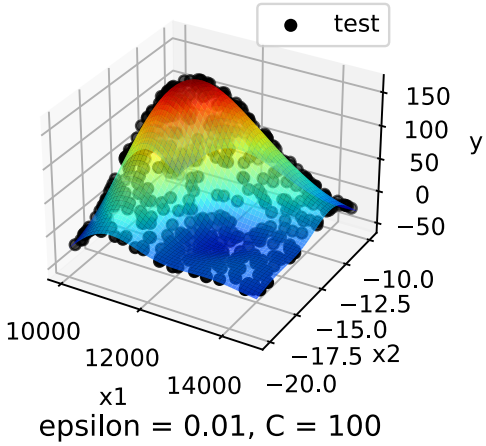
epsilon = 1, C = 0.01



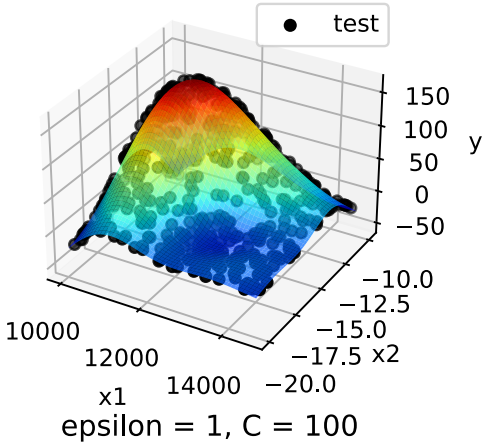
epsilon = 0.0001, C = 10



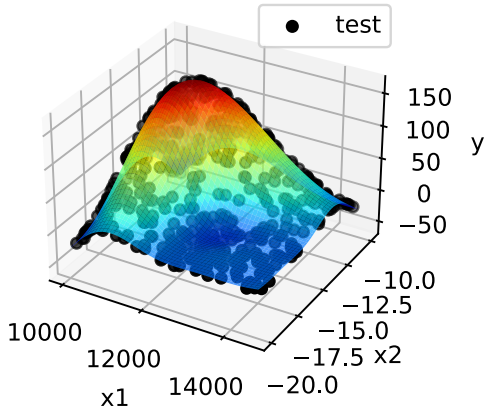
epsilon = 0.01, C = 10



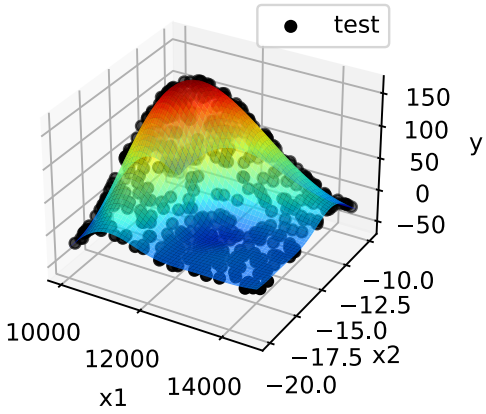
epsilon = 1, C = 10



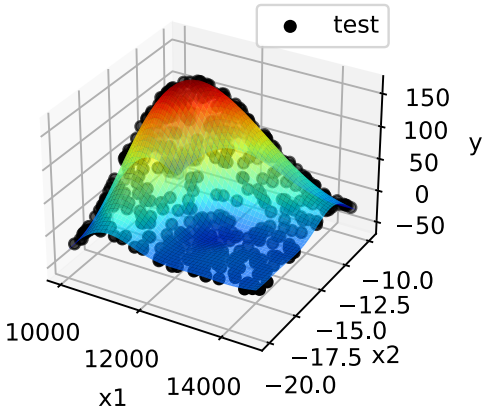
epsilon = 0.0001, C = 100



epsilon = 0.01, C = 100



epsilon = 1, C = 100



In []: