4. Consider the data set with scalar features $x_i$ and binary class labels $y_i = \pm 1$.

| $x_i$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $y_i$ | 1 | -1 | 1 | -1 |

A support vector classifier is of the form

$$\hat{y} = \begin{cases} 1 & z > 0 \\ -1 & z < 0, \end{cases} \qquad z = \sum_i \alpha_i y_i K(x_i, x),$$

where $K(x, x')$ is the radial basis function, $K(x, x') = e^{-\gamma(x-x')^2}$, and $\gamma > 0$ and $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_4]$ are parameters of the classifier.

(a) Use python to plot $z$ vs. $x$ and $\hat{y}$ vs. $x$ when $\gamma = 3$ and $\boldsymbol{\alpha} = [0, 0, 1, 1]$.

(b) Repeat (a) with $\gamma = 0.3$ and $\boldsymbol{\alpha} = [1, 1, 1, 1]$.

**Solution (a,b):**

```
import numpy as np
x = np.array ([0 ,1 ,2 ,3])
y = np.array([1,-1,1,-1])

def plot_kernel(x,y, alpha, gamma):
    x_prime = np.linspace(-2,5,100)
    xpmat, xmat = np.meshgrid(x_prime, x)
    dist = (xpmat-xmat)**2
    kernel = np.exp(-gamma*dist)
    z = (y*alpha).dot(kernel)
    yhat = np.ones_like(z)
    yhat[z<0] = -1
    plt.plot(x_prime ,z)
    plt.plot(x_prime ,yhat)
    plt.grid ()
    plt.legend(["z", "yhat"])
    plt.xlabel('x',fontsize =20)

plt.subplot(1 ,2 ,1)
alpha = np.array([0 ,0 ,1 ,1])
plot_kernel(x,y,alpha,gamma=3)
plt.title('gamma=3')


plt.subplot(1 ,2 ,2)
alpha = np.array([1 ,1 ,1 ,1])
plot_kernel(x,y,alpha,gamma =0.3)
plt.title('gamma=0.3')
plt.savefig('svm_kernel.png')
```

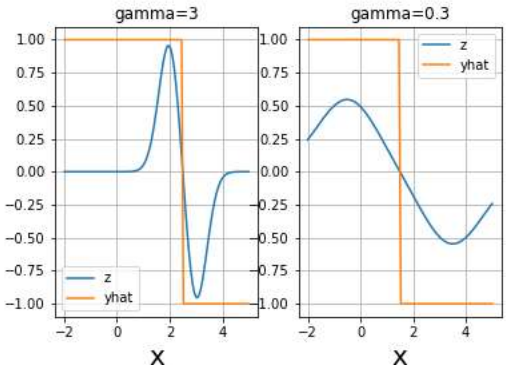


Figure 7: Plots for part a and b

(c) Which classifier makes more errors on the training data.

**Solution c:**

In model 1 where gamma is 3, the model misclassifies points where x = 1.

In model 2, where gamma is 0.3, the model misclassifies point where x = 1 , 2. So the second model makes more mistakes.