1. Consider the data set for four points with features $\mathbf{x}_i = (x_{i1}, x_{i2})$ and binary class labels $y_i = \pm 1$.

x_{i1}	0	1	1	2
x_{i2}	0	0.3	0.7	1
y_i	-1	-1	1	1

   (a) Find a linear classifier that separates the two classes. Your classifier should be of the form

$$\hat{y} = \begin{cases} 1 & \text{if } b + w_1 x_1 + w_2 x_2 > 0 \\ -1 & \text{if } b + w_1 x_1 + w_2 x_2 < 0 \end{cases}$$

   State the intercept b and weights w_1 and w_2 for your classifier. Note there is no unique answer as there are multiple linear classifiers that could separate the classes.

   (b) Find the maximum γ such that

$$y_i(b + w_1 x_{i1} + w_{i2} x_{i2}) \geq \gamma, \text{ for all } i,$$

   for the classifier in part (a)?

   (c) Compute the margin of the classifier

$$m = \frac{\gamma}{\|\mathbf{w}\|}, \quad \|\mathbf{w}\| = \sqrt{w_1^2 + w_2^2}.$$

   (d) Which samples i are on the margin for your classifier?

2. 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

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

   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 $\boldsymbol{\alpha} = [1, 1, 1, 1]$.

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