Task 1:

1.

The margin vector is the separation line going through a graph of data points and its margins. The support vectors are the data points that are on the edge of the margin formed by the margin vector.

2.

SMVs add a slack variable to account for misclassifications

3.

Kernels are models that allow lower dimensional data to be mapped to higher dimensions to allow for non-linear SVMs to help make more accurate graphs. They can also allow data that is otherwise impossible to separate to be put in a context in which it can be separated.

4.

Kernels allow more complex feature vectors to be created to add more ways to define and separate data to come to more accurate conclusions and models.

Task 2:

1.

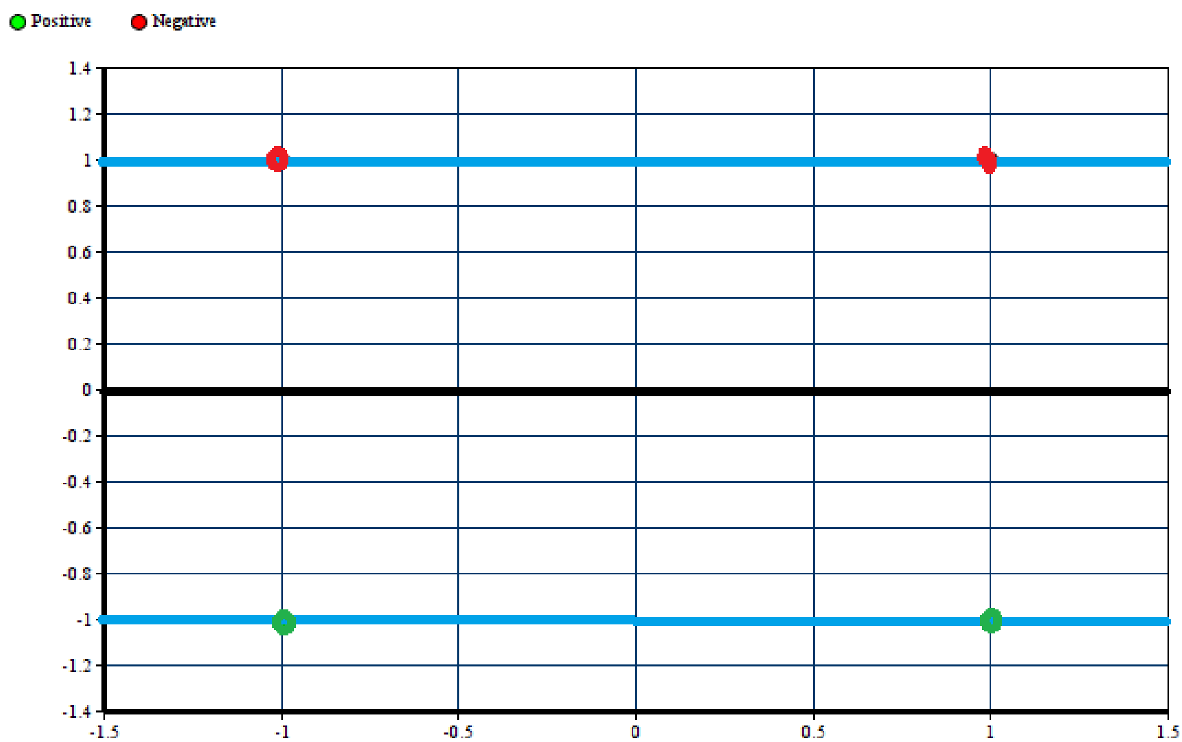
In notebook Homework 3

2.

Class	X_1	X_2	X_1X_2
-	-1	-1	1
+	-1	1	-1
+	1	-1	-1
-	1	1	1

The black line is the separator line. The blue lines are the margins. All points in this case are support vectors. The margin is size 1 spanning from -1 to 1 while the separator line is the x axis.

X_1 is graphed to the x axis and $X_1X_2$is graphed to the y axis



Task 3:

$$(x_1 - a)^2 + (x_2 - b)^2 - r^2 = 0$$

$$x_1{}^2 - 2ax_1 + a^2 + x_2{}^2 - 2bx_2 + b^2 - r^2 = 0$$

$$x_1{}^2 + x_2{}^2 - 2ax_1 - 2bx_2 + a^2 + b^2 - r^2 = 0$$

$$\begin{bmatrix} 1 & 1 & -2a & -2b \end{bmatrix} \begin{bmatrix} x_1{}^2 \\ x_2{}^2 \\ x_1 \\ x_2 \end{bmatrix} + (a^2 + b^2 - r^2) = 0$$

Task 4:

$$c(x_1 - a)^2 + d(x_2 - b)^2 - 1 = 0$$

$$cx_1{}^2 - 2acx_1 + ca^2 + dx_2{}^2 - 2bdx_2 + db^2 - 1 = 0$$

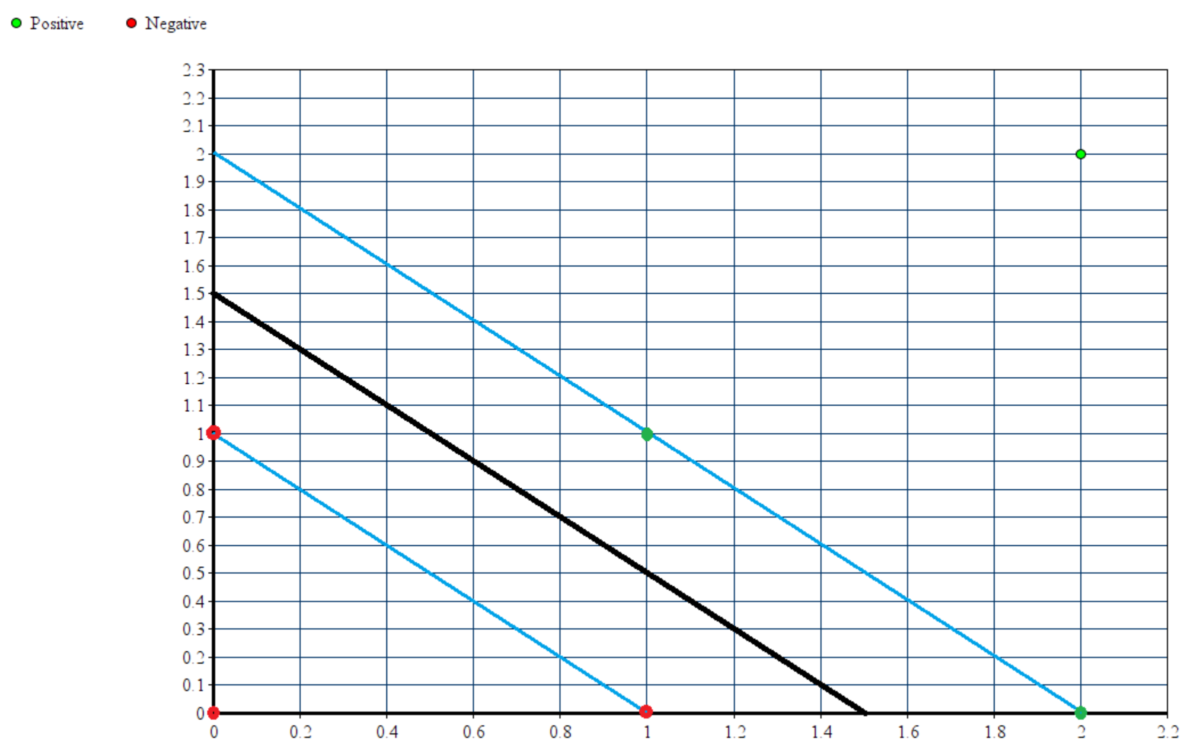
$$c{x_1}^2 + d{x_2}^2 - 2acx_1 - 2bdx_2 + ca^2 + db^2 - 1 = 0$$

$$\begin{bmatrix} 0 & -2ac & -2bd & c & d & 0 \end{bmatrix} \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ {x_1}^2 \\ {x_2}^2 \\ x_1x_2 \end{bmatrix} + (ca^2 + db^2 - 1) = 0$$

Task 5:

a.

Since the decision boundary intends to find a line that has all of one type of answer on one side and the other type on the other, I created this graph with the black line representing the decision boundary and the blue lines as the margins. The points on the blue lines are the support vectors



These classes are linearly separable because a line can be drawn between the classes and one class on each side of the line b.

Given $h(x) = w_1 x_1 + w_2 x_2 + b > 0$

$$1x_2 > -1x_1 + \frac{3}{2}$$

$$x_2 + x_1 - \frac{3}{2} > 0$$

$$w_1 = 1, w_2 = 1, b = -\frac{3}{2}$$

The slope of the MMHP is -1 and passes through (0, 3/2)

Task 6:

a.

These classes are not linearly separable. There is no way to draw a straight line that isolate the positive and negative classes

b.

Class	x	1	$\sqrt{2}x$	x^2
+	0	1	0	0
-	1	1	$\sqrt{2}$	1
-	-1	1	$-\sqrt{2}$	1

considering how the data translates, we do not need three dimensions since everything is the same in the 1 dimension. There would not be any information gleaned from adding a third dimension. Therefore, we can isolate it to two dimensions [$\sqrt{2}x$,x^2]

this would give us the points + (0,0), - ($\sqrt{2}$,1), - (-$\sqrt{2}$,1) In this case, we can find a hyperplane of y=0.5. with a line like that, we can have a margin of 0.5 and having all three points as supporting vectors.

Task 7:

```
Linear:  [0.80446927 0.80337079 0.78651685 0.75280899 0.78651685]

Quadratic:  [0.74860335 0.79775281 0.76966292 0.75280899 0.78651685]

RBF:  [0.73743017 0.7752809  0.78089888 0.75842697 0.76966292]
```

From these results, linear seems the best one. No matter how I changed gamma or C, Linear always gave the highest results, all the time and only when gamma was really small were the others able to become the same as Linear, but never greater. So based upon my experimentation, linear seems the best.