Ayberk Aydin

Task 1:

Q1: The margin of a linear classifier is the width that the boundary could be increased by before hitting a data point on two sides. Support vectors are datapoints that the margin pushes up against.

Q2: When data is not linearly separable, the original input space can always be mapped to some higher-dimensional feature space where the data is linearly separable.

Q3: A kernel is a function that directly obtains the inner product between the images of all pairs of data in the feature space to avoid the computationally expensive method of feature mapping.

Q4: A kernel is related to feature vectors in that it takes them as input (while they are represented in the original, lower dimensional space) and returns the inner product between their transformed versions in the higher dimensional space.

Task 3:

Expanding we get

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

which implies that the weights are $(-2a, -2b, 1, 1)$ and the intercept is $a^2 + b^2 - r^2$. Therefore, every circular region is linearly separable from the rest of the plane in the given feature space.

Task 4:

Expanding we get

$$cx_1^2 - 2acx_1 + a^2c + dx_2^2 - 2bdx_2 + b^2d - 1 = 0$$

which implies that the weights are $(-1, -2ac, -2bd, c, d, 0)$ and the intercept is $a^2c + b^2d$. Thus, SVMs with the given kernel can separate any elliptic region from the rest of the plane.

Task 6:

a) No, they are not linearly separable in 1-D.

b) The given mapping produces the following points in 3-D, respectively: $(1, 0, 0)$, $(1, -\sqrt{2}, 1)$, and $(1, \sqrt{2}, 1)$. Now the classes are linearly separable. One possible separating hyperplane (I believe it is also the maximum margin hyperplane) is given by $z - 0.5 = 0$, where z corresponds to the third dimension of the mapping (x^2). This is the parallel, horizontal plane 0.5 units above the x-y plane, which creates a margin of 0.5 between the positive and negative datapoints.

Associated code can be found here.