Question 1

(a) Submitted sketch in zip.

(b) Classify to red if $B_0 + B_1X_1 + B_2X_2 \geq 0$. Classify to blue if $B_0 + B_1X_1 + B_2X_2 < 0$.

$$B_0 = 1, B_1 = -1, B_2 = 1$$

Therefore, classify to red if $1 - X_1 + X_2 \geq 0$ and classify to blue if $1 - X_1 + X_2 < 0$.

(c) The dashed lines on the sketch for 1(a) indicate the maximal margin hyperplane and are on either side of the hyperplane.

(d) Support vectors are: (2,0), (4,2), (2,2), (4,4)

(e) A slight movement at (4,0) would not affect the maximal margin hyperplane as it is not a support vector. The hyperplane will remain unchanged as long as the change does not cross the margin and become a support vector.

(f) Submitted sketch in zip.

(g) Submitted sketch in zip.

Question 2

(a) The shape of X is $(4, 2)$ as there are 4 samples and 2 features. The shape of y is 4 as it represents the labels for the 4 samples. Bonus: This is an XOR gate. If input bits are the same, the output is false.

X1	X2	Y
1	1	1
-1	-1	1
1	-1	-1
-1	1	-1

Table 1: Data points and classifications.

(b) Sketch submitted in zip.
The points form the vertices of a square and are not linearly separable. No straight line can be drawn to separate the positive and negative examples.

(c) The positive example gets transformed [1, 1, 1] and [-1, -1, 1]. The negative example gets transformed [1, -1, -1] and [-1, 1, -1]
In this 3D space, these points are linearly separable.
Matplotlib sketch provided in zip.

(d) The margin size is the distance between the nearest point(s) of any class and the decision boundary. Without specific linear classifiers parameter, calculating the numerical value of this margin isn't straightforward.

All four points could be considered support vectors as they are all closest to the decision boundary.