Exercise 1

1. let $f(x) = w^T x + b = 0$

plugging in parameters, we get.

$x_1 - 2x_2 + 3x_3 - 1 = 0$

since this equation $x_1 - 2x_2 + 3x_3 = 1$ defines a flat 2D surface and it is in 3D feature space, it is then a hyperplane.

2. $\nabla_x f(x) = w = n = [1 \ -2 \ 3]^T$

normal vector $n = w$

the normal vector represents where the decision function grows most rapidly.

3.

(a)

$f([1 \ 1 \ 1]^T) = 1 - 2 + 3 - 1 = 1$

$\| [1 \ -2 \ 3]^T \| = \sqrt{1+4+9} = \sqrt{14}$.

$d(x) = \frac{1}{\sqrt{14}}$.

(b) since $d(x) > 0$, the point $x \, [1 \ 1 \ 1]^T$ is positive by the decision boundary classification, which is a prediction.

(c)
The distance represent the margin, if the margin is small, it means distance is small which leads to low confidence and vice versa.

(d) from (a) we have $d(x) = \frac{1}{\sqrt{14}}$ $\|w\| = \sqrt{14}$ $w = [1, -2, 3]^T$

$$\text{Projection} = x - d(x) \frac{w}{\|w\|}$$

$$= [1\ 1\ 1]^T - \frac{1}{\sqrt{14}} \frac{1}{\sqrt{14}} [1\ -2\ 3]^T$$

$$= \left[\frac{13}{14}\ \ \frac{8}{7}\ \ \frac{11}{14}\right]^T$$

(e)
Since orthogonal projection gives the distance from points to decision boundary, the maximal margin classifier use this value to find support vectors and thus define the hyperplane.