Q1

(a)

This classification model is not linear w.r.t. the input.
Proof (in 1d):
Let $x_1 = 1, x_2 = 2$ be arbitrary inputs, $b = 0$, $W = 1$, and $\tau = 3$. Then:

$$\begin{aligned}
&f(u+v) = \sigma(W \cdot (x_1 + x_2) + b) = \sigma(3) = 1 \\
&f(u) + f(v) = \sigma(Wx_1 + b) + \sigma(Wx_2 + b) = \sigma(1) + \sigma(2) = 0 + 0 = 0 \\
&\Rightarrow f(u+v) \neq f(u) + f(v) \Rightarrow \text{Not linear}
\end{aligned} \qquad \blacksquare$$

(b)

We have 4 samples, each can be correctly or incorrectly classified, thus $\text{acc} \in \{0\%, 25\%, 50\%, 75\%, 100\%\}$. We will prove that $\text{acc} = 100\%$ is impossible, and find a concrete example that shows that 75% accuracy is possible. Thus the maximum accuracy is 75%.
Using the bias trick (first row) we can define:

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}, \mathbf{W} = \begin{bmatrix} b \\ w_1 \\ w_2 \end{bmatrix}$$

Assume by contraditction $\text{acc} = 100\%$:

$$\begin{aligned}
y = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \sigma(WX) &\Longrightarrow \begin{cases} b + w_1 + w_2 > \tau \\ b + w_1 - w_2 < \tau \\ b - w_1 + w_2 < \tau \\ b - w_1 - w_2 > \tau \end{cases} \\
&\Longrightarrow \begin{cases} \cancel{b} + \cancel{w}_1 - w_2 < \cancel{b} + \cancel{w}_1 + w_2 \\ \cancel{b} - \cancel{w}_1 + w_2 < \cancel{b} - \cancel{w}_1 - w_2 \end{cases} \\
&\Longrightarrow \begin{cases} w_2 > 0 \\ w_2 < 0 \end{cases}
\end{aligned}$$

Which is a contradiction. Therefore the maximum accuracy cannot be 100%.

(d)

In general, we'd prefer to use a linear model when:

1. The data is linearly separable.

2.