The given objective function is:

$$L(w) = \sum_{i=1}^{N} \left(f(w^T x_i) - y_i \right)^2$$

where $f(t) = \frac{1}{1+\exp(-t)}$ is the logistic function.

The goal is to show that $L(w^*)$ has a minimum at $w = w^*$, where $\left.\frac{\partial L}{\partial w}\right|_{w=w^*} = 0$.

Step 1: Compute the derivative of the logistic function $f(t)$

The logistic function is:

$$f(t) = \frac{1}{1 + \exp(-t)}$$

The derivative of this function with respect to t is:

$$f'(t) = \frac{\exp(-t)}{(1 + \exp(-t))^2} = f(t)(1 - f(t))$$

Step 2: Differentiate the objective function

Now, we compute the gradient of $L(w)$ with respect to w. Expanding $L(w)$:

$$L(w) = \sum_{i=1}^{N} \left(f(w^T x_i) - y_i \right)^2$$

The gradient with respect to w is:

$$\frac{\partial L}{\partial w} = 2 \sum_{i=1}^{N} \left(f(w^T x_i) - y_i \right) \frac{\partial}{\partial w} \left(f(w^T x_i) \right)$$

We already know that:

$$\frac{\partial}{\partial w} f(w^T x_i) = f(w^T x_i)(1 - f(w^T x_i)) x_i$$

Thus, the gradient becomes:

$$\frac{\partial L}{\partial w} = 2 \sum_{i=1}^{N} \left(f(w^T x_i) - y_i \right) f(w^T x_i) \left(1 - f(w^T x_i) \right) x_i$$

Step 3: Evaluate at $w = w^*$

At the optimal point w^*, the gradient must be zero:

$$\frac{\partial L}{\partial w}\bigg|_{w=w^*} = 0$$

This implies that the residual term $(f(w^* x_i) - y_i)$ will be small, balancing out the gradient expression.

Step 4: Second-order condition for a minimum

To confirm that $L(w^*)$ is a minimum, we compute the second derivative (Hessian) of $L(w)$ with respect to w. The Hessian matrix is given by:

$$H_{jk} = \frac{\partial^2 L}{\partial w_j \partial w_k}$$

The gradient is:

$$\frac{\partial L}{\partial w} = 2 \sum_{i=1}^{N} \left(f(w^T x_i) - y_i\right) f(w^T x_i)(1 - f(w^T x_i)) x_i$$

We take the derivative of this gradient to find the Hessian:

$$H_{jk} = 2 \sum_{i=1}^{N} \left[f(w^T x_i)(1 - f(w^T x_i)) x_{ij} x_{ik} - (f(w^T x_i) - y_i) f(w^T x_i)(1 - f(w^T x_i))(1 - 2f(w^T x_i)) x_{ij} x_{ik} \right]$$

This Hessian matrix determines the curvature of the objective function, and since $f(w^T x_i)(1 - f(w^T x_i))$ is always positive, the Hessian is positive semi-definite, ensuring that w^* is a minimum point.

Therefore, $L(w)$ achieves a minimum at $w = w^*$, confirming that $\frac{\partial L}{\partial w}\big|_{w=w^*} = 0$ corresponds to a minimum, not a maximum.