Q. $\hat{d}_i = \hat{w}^T x_i + \varepsilon$ $\quad [x_i \longrightarrow$ i^{th} sample from training set $]$

is the equation for linear regression where we know that

$$\varepsilon \sim \text{Laplace}(\mu, \sigma)$$

⇒ given input vectors x_n (training data) and labels y_n we have (for size n dataset)

~~$\varepsilon = \hat{w}x - y$~~

~~$\varepsilon = \hat{w}x - y$ if we consider~~

$l(w) \longrightarrow$ likelihood of w being the ~~right~~ right weights

$$l(w) = \prod_{i=1}^{n} p(w_n)$$

$$= \prod_{i=1}^{n} \frac{1}{2\sigma} \exp\left(-\frac{|\varepsilon_i|}{\sigma}\right)$$

$$= \prod_{i=1}^{n} \frac{1}{2\sigma} \exp\left(\frac{|d_i - w^T x_i|}{\sigma}\right)$$

$$\log(l(w)) = \sum_{i=1}^{N} \frac{-1}{\sigma} |d_i - w^T x_i| + N \log\left(\frac{1}{2\sigma}\right)$$

$$= -\left[\sum_{i=1}^{N} -\frac{1}{\sigma} |d_i - w^T x_i| + N \log(2\sigma)\right]$$

Ans.

The negative of this log likelihood is minimized, hence $-\log(l(w))$ is the desired objective function.