

Derivation of the equations

The mutual information can be calculated as,

$$I(W;S) = H(S) - H(S|W) = H(S) - \sum P(w) \cdot H(S|W=w) = H(S) - \int_{-q}^q f_w(w) \cdot H(S|W=w) dw \quad (1)$$

with

$$\begin{aligned} H(S|W=w) &= \sum_{s_i \in S} P(s_i|w) \log \frac{1}{P(s_i|w)} \\ &= P(S=1|W=w) \log \frac{1}{P(S=1|W=w)} + P(S=0|W=w) \log \frac{1}{P(S=0|W=w)} \end{aligned} \quad (2)$$

We have $p_{w0} = P(S=0|W=w)$ and $p_{w1} = P(S=1|W=w)$. Therefore,

$$H(S|W=w) = -(p_{w1} \log p_{w1} + p_{w0} \log p_{w0}) \quad (3)$$

Substituting (2) and (3) into (1), it gives,

$$\begin{aligned} I(W;S) &= H(S) + \int_{-q}^q (p_{w1} \log p_{w1} + p_{w0} \log p_{w0}) f_w(w) dw \\ &= H(S) + \int_{-q}^q (p_{w1} \log p_{w1} + p_{w0} \log(1-p_{w1})) f_w(w) dw \end{aligned}$$

(Instead of $I(W;S) = H(S) + \int_{-q}^q (p_{w1} \log p_{w1} + p_{w0} \log(1-p_{w0})) f_w(w) dw$ given in the paper).

I think the paper has a typo, for this result is in accord with equations presented in its rest part.