Details of the Incremental Coding Length (ICL)

Let $\boldsymbol{p}$ be the feature bank, the Incremental Coding Length of a feature p_k is defined as below:

$$\begin{aligned}\text{ICL}(p_k) &= \frac{\partial H(\boldsymbol{p})}{\partial p_k} = -\frac{\sum_i \partial p_i \log p_i}{\partial p_k} \\ &= -\frac{\partial p_k \log p_k}{\partial p_k} - \frac{\sum_{j \neq k} \partial p_j \log p_j}{\partial p_k}. \end{aligned} \tag{1}$$

The first term in Eq. 1 equals:

$$\frac{\partial p_k \log p_k}{\partial p_k} = 1 + \log p_k.$$

The second term in Eq. 1 can be further reduced:

$$\begin{aligned} & \frac{\sum_{j \neq k} \partial p_j \log p_j}{\partial p_k} \\ = \; & \lim_{\varepsilon \to 0} \sum_{j \neq k} \frac{1}{\varepsilon} \Big(\frac{p_j}{1+\varepsilon} \log \frac{p_j}{1+\varepsilon} - p_j \log p_j \Big) \\ = \; & \lim_{\varepsilon \to 0} \sum_{j \neq k} \frac{1}{\varepsilon} \Big(\frac{p_j}{1+\varepsilon} \big(\log p_j - \log(1+\varepsilon) \big) - p_j \log p_j \Big) \\ = \; & \lim_{\varepsilon \to 0} \sum_{j \neq k} \frac{1}{\varepsilon} \Big(- \frac{\varepsilon}{1+\varepsilon} p_j \log p_j - \frac{p_j}{1+\varepsilon} \log(1+\varepsilon) \Big) \\ = \; & \sum_{j \neq k} \Big(- p_j \log p_j - p_j \Big) \\ = \; & \sum_{i} \Big(- p_i \log p_i - p_i \Big) + p_k \log p_k + p_k \\ = \; & H(\boldsymbol{p}) - 1 + p_k \log p_k + p_k \end{aligned}$$

Reduce Eq. 2 to Eq. 1, we finally have:

$$\text{ICL}(p_k) = \frac{\partial H(\boldsymbol{p})}{\partial p_k} = -H(\boldsymbol{p}) - p_k - \log p_k - p_k \log p_k. \tag{2}$$