Exam 2021

Question 19: We assume an arbitrary discrete set $\mathcal{H}$ of classifiers and $P(h)$ and $Q(h)$ denote probability masses at h. Note: base 2 logarithms are used

a) Calculate the Kullback-Leibler (KL) divergence between P and Q, when P, Q are two Bernoulli distributions with probability of success p and probability of success q, respectively.

$$D_{KL}(P\|Q) = \sum_h P(h) \log_2 \frac{P(h)}{Q(h)} = \mathbb{E}\left(\log_2 \frac{P(h)}{Q(h)}\right) = \mathbb{E}\left(\log_2 \frac{p^h(1-p)^{1-h}}{q^h(1-q)^{1-h}}\right)$$

$$= \mathbb{E}\left(h \log_2 \frac{p}{q} + (1-h)\log_2 \frac{1-p}{1-q}\right) = p \log_2 \frac{p}{q} + (1-p)\log_2 \frac{1-p}{1-q}$$

b) Under which condition, the KL divergence is symmetric?

Symmetric $D_{KL}(P\|Q) = D_{KL}(Q\|P) \Rightarrow p = 1-q$

(the case $p=q$ results in $D_{KL}(P\|Q) = 0$ and is trivial)

c) If you have to guess the Jensen-Shannon divergence (JSD) between P and Q, will you choose that JSD = 0.5 or JSD = 2? (*Note: base 2 logarithms are used*)

The JSD is lower bounded by 0 and upper bounded by 1 (for base 2 log)

i.e. $0 \leq JSD(P\|Q) \leq 1$

There are many ways to show this; one way is to show

$$JSD(P\|Q) = I(X;Z) = H(X) - H(X|Z)$$

where $X \sim M = \frac{P+Q}{2}$ and Z binary indicator function

i.e. $X \sim P$ if $Z=0$

$X \sim Q$ if $Z=1$

Since $I(X;Z) \leq H(Z) = 1$ (for base 2 logarithm) *actually $I(X;Z) \leq \log(\min(|\mathcal{X}|,|Z|))$

$$0 \leq JSD(P\|Q) \leq 1$$

Note: It was accepted even without proof, that $0 \leq JSD \leq 1$ for base 2 logarithm

Therefore, one should choose that $JSD = 0.5$!

d) If $W = (P+Q)/2$ is approximated by a Bernoulli distribution with probability of success w, find values of w that minimizes and maximizes the JSD when $p = q = 0.5$.

$$JSD = \frac{1}{2} D_{KL}(P\|W) + \frac{1}{2} D_{KL}(Q\|W)$$

$$= \frac{1}{2}\left(p \log_2 \frac{p}{w} + (1-p)\log_2 \frac{1-p}{1-w}\right) + \frac{1}{2}\left(q \log_2 \frac{q}{w} + (1-q)\log_2 \frac{1-q}{1-w}\right)$$

$$= \frac{1}{2}\log_2\left(\frac{0.25}{w(1-w)}\right)$$

$$\min_w JSD = \min_w \frac{1}{2}\log_2 \frac{0.25}{w(1-w)}$$

For $w = 1/2 \longrightarrow JSD = 0$

$$\max_w JSD = \max_w \frac{1}{2}\log_2 \frac{0.25}{w(1-w)}$$

Taking the derivatives gives $w=0$ or $w=1$

However, this does not necessarily give $JSD \leq 1$

The maximal value of JSD is 1, which gives $\frac{1}{2}\log_2 \frac{0.25}{w(1-w)} = 1$

$$\Rightarrow w^2 - w + 1/16 = 0 \Rightarrow w = \frac{1 \pm \sqrt{3}/2}{2}$$