Contrastive clustering

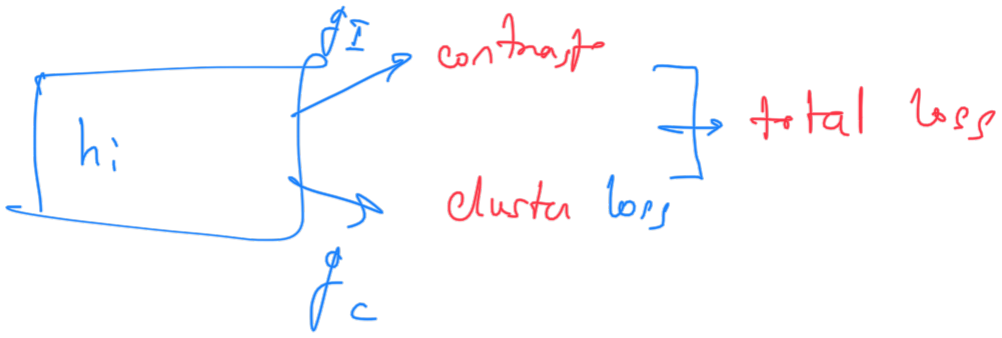
Input: $x_i^a = T^a(x_i)$

$x_i^b = T^b(x_i)$

$h_i^a = f(x_i^a)$

$h_i^b = f(x_i^b)$ } Encoder network.

Multitask setup



$z_i^a = g_I(h_i^a)$

Similarity measurement:

$$s\left(z_i^{k_i}, z_j^{k_2}\right) = \frac{\left(z_i^{k_i}\right)\left(z_j^{k_2}\right)^T}{\|z_i^{k_i}\| \, \|z_j^{k_2}\|}$$

Loss: $l_i^a = -\log \frac{\exp\left(s(\hat{z}_i^a, z_i^b)/\tau\right)}{}$

$$\sum_{i=1}^{N} \left[\exp\left(s\left(z_i^a, z_j^a\right)/\tau_I\right) + \exp\left(s\left(z_i^a, z_j^b\right)/\tau_I\right) \right]$$

[NTXent loss : looks suspicious]

$$l_{ins} = l_i^a + l_i^b$$

cluster label contrastive head

$y^a \in \mathbb{R}^{N \times m}$ // CCI output

$y^a_{n,m}$ ~ Prob of sample n to cluster m

y_i^a denote i colam for y^a

+ve → y_i^b denotec i column

Rest of $2N-2$ are negative

Similar contrastive loss between the column.

$$L_{Class} = \frac{1}{2m} \sum_{i=1}^{m} (\hat{L}_a + \hat{L}_b) - H(y)$$

$$H(y) = -\sum_{i=1}^{m} \left[p(\hat{y}_i^a) \log p(y_i^a) + p(y_i^b) \log p(y_i^b) \right]$$

↗ softmax across the column

Entropy loss

$$L = L_{ins} + L_{cls}$$