Ⓟ Demystifying CL

Demystifying Contrastive Learning

Contrastive loss.

$$\mathcal{L}(D, D^+) = -\sum_{(x,x^+)\in D^+} \frac{\exp\left(f(x)^T f(x^+)/\tau\right)}{\exp\left[f(x)^T f(x^+)/\tau\right] + \sum_{\substack{x^- \in \mathcal{D} \\ x, x^- \notin \mathcal{D}^+}} \exp\left(f(x)^T f(x^-)/\tau\right)}$$

Measuring Invariance;

transformation t

invariant function h iff. $\underbrace{h(x) = h(t(x))}$

Formal Notion iff $y(x) = \overset{\longrightarrow \text{label of image } t(x)\ (x)}{y(t(x))}$; where, $t : \mathcal{X} \to \mathcal{X}$

the $\underbrace{h^*(x) = h^*(t(x))}_{\Downarrow}$

invariant for $t(x)$ & label (y)

Definition of firing unit

$h(x) \in \mathbb{R}^n$; fire if $s_i h_i(x) > t_i$; $s_i \in \{-1, 1\}$

$f_i(x) = \mathbb{1}\left(s_i h_i(x) > t_i\right)$; $f(x) \in \mathbb{R}^n$

ith neuron

global firing rate, $G(i) : \mathbb{E}\left[f_i(x)\right]$ // t_i dependency.

t_i chosen such that $\underbrace{G(i) = \frac{1}{|y|}}_{\Downarrow}$ → no of class.

we want

↳ number of firing unit

one class → one $\underbrace{\text{section firing}}$;

equal parts

(P) Demystifying CL

Local trajectory: $T(x) = \{ t(x, \gamma) \mid \forall \gamma \}$ // set of transformed version of x image.

Local firing rate is defined as below

$$\underbrace{L_y(\overset{\text{ith neuron}}{i})}_{\Downarrow\ \text{fraction of time } i \text{ neuron fires.}} = \underbrace{\frac{1}{|\mathcal{X}_y|}}_{\text{Avg}} \underbrace{\sum_{z \in \mathcal{X}_y}}_{\text{All image in } \mathcal{X}_y} \underbrace{\frac{1}{|T(z)|}}_{\text{Avg}} \underbrace{\sum_{x \in T(z)} f_i(x)}_{\text{measuring local firing for } x \text{ \& their transformation.}} \quad // \ \mathcal{X}_y = \{ x \mid x \in \mathcal{X}, y(x) = y \}$$

Target conditioned invariance $I_y(i) = \dfrac{L_y(i)}{G(i)}$ // $\to$ find (top-k) neurons.

Representation Invariance Score: (RIS):

commonalities in top k neurons for each classes.