Polyloss: Polynomial perspective of classification loss.

Decompose the loss function into

$$\sum_j \alpha_j (1-p_t)^j \; ; \; \alpha_j \in \mathbb{R}^+$$

Special cases: $\alpha_j = \frac{1}{j}$ // cross entropy loss

Cross entropy

$$\mathcal{L}_{CE} = -\log p_t$$

$$= \sum_{j=1}^{\infty} \frac{1}{j} (1-p_t)^j$$

$$-\frac{d\mathcal{L}_{CE}}{dp_t} = \sum_{j=1}^{\infty} (1-p_t)^{j-1}$$

Focal loss:

$$\mathcal{L}_{Fl} = (1-p_t)^\gamma \log p_t$$

$$= \sum^{\infty} \frac{1}{j} (1-p_t)^{j+\gamma}$$

$$\sum_{j=1} \quad j$$

[improve the relative importance of the first term.

$+\delta \rightarrow$

$$-\frac{d L_{CE}}{dP_t} = \sum_{j=1}^{\infty}\left(1+\frac{\gamma}{j}\right)(1-P_t)^{j+\gamma-1}$$

Extra weight

However No control here

→ what if we have control here

[poly loss]

Polynomial loss perturb the Nth polynomial term

perturbed

$$\mathcal{L}_{poly-N} = \overbrace{(\epsilon_1+1)(1-P_t) + \dots + \left(\epsilon_N + \tfrac{1}{N}\right)(1-P_t)^N}$$

$$\underbrace{+\frac{1}{1+N}(1-P_t)^{N+1} + \dots\dots}$$

control

$$= -\log P_t + \sum^{N} \epsilon_i (1-P_t)^j$$

$$j=1$$

$\mathcal{L}_{poly-1} \rightarrow$ just modifies the 1st term