Appendix

Details of the Feedback Coefficient

Let θ be the parameters of the teacher, ξ be the parameters of the student. Let $y_u = f_\theta(x_u)$ be the prediction of unlabeled data. Y_u is the so-called pseudo label sampled from y_u. Let

$$L_{\xi,u} = Loss(f_\xi(x_u), Y_u) \tag{1}$$

We use these pseudo-labeled data to update the student's parameters θ, i.e.,

$$\xi_{n+1} = \xi_n - \eta \nabla_\xi L_{\xi,u} \tag{2}$$

For the student, we expect to decrease its loss on labeled data, and this is achieved through teacher-generated pseudo labels. This means we need to solve

$$\theta = \underset{\theta}{argmin}\, E_{Y_u|\theta}[L_{\xi_{n+1},l}] = \underset{\theta}{argmin}\, E_{Y_u|\theta}[Loss(f_{\xi_{n+1}}(x_l), Y_l)] \tag{3}$$

, where x_l is labeled data and Y_l is the label. Let

$$L_\theta = E_{Y_u \sim\theta}[L_{\xi_{n+1},l}] \tag{4}$$

Since ξ_{n+1} is a function of Y_u sampled from $f_\theta(x_u)$, apparently ξ_{n+1} is also a function of θ. That means we can calculate the gradient of L_θ with respect to θ as

$$\begin{aligned}
\nabla_\theta L_\theta &= \nabla_\theta E_{Y_u\sim\theta}[L_{\xi_{n+1},l}] && (5)\\
&= \nabla_{\xi_{n+1}} L_{\xi_{n+1},l} \cdot \nabla_\theta E_{Y_u\sim\theta}[\xi_{n+1}] && (6)\\
&= \nabla_{\xi_{n+1}} L_{\xi_{n+1},l} \cdot \nabla_\theta E_{Y_u\sim\theta}[\xi_n - \eta\nabla_\xi L_{\xi,u}] && (7)\\
&= -\eta\nabla_{\xi_{n+1}} L_{\xi_{n+1},l} \cdot \nabla_\theta E_{Y_u\sim\theta}[\nabla_\xi L_{\xi,u}] && (8)\\
&= -\eta\nabla_{\xi_{n+1}} L_{\xi_{n+1},l} \cdot \nabla_\theta E_{Y_u\sim\theta}[\nabla_\xi Loss(f_\xi(x_u), Y_u)] && (9)\\
&= \eta\nabla_{\xi_{n+1}} L_{\xi_{n+1},l} \cdot E_{Y_u\sim\theta}[\nabla_\xi Loss(f_\xi(x_u), Y_u) \cdot -\nabla_\theta ln(P(Y_u))] && (10)\\
&= \eta\nabla_{\xi_{n+1}} L_{\xi_{n+1},l} \cdot E_{Y_u\sim\theta}[\nabla_\xi L_{\xi,u} \cdot -\nabla_\theta ln(P(Y_u))] && (11)
\end{aligned}$$

where $P(Y_u)$ is the possibility distribution with respect to $f_\theta(x_u)$. If we have

$$L_{\theta,u} = CrossEntropy(f_\theta(x_u), Y_u) \tag{12}$$

, then

$$\nabla_\theta L_\theta = \eta\nabla_{\xi_{n+1}} L_{\xi_{n+1},l} \cdot \nabla_\xi L_{\xi,u} \cdot \nabla_\theta L_{\theta,u} \tag{13}$$

according to the definition of $CrossEntropy$. Let

$$h = \eta\nabla_{\xi_{n+1}} L_{\xi_{n+1},l} \cdot \nabla_\xi L_{\xi,u} \tag{14}$$

We call h the feedback coefficient as same as MPL.

It is not efficient to calculate h directly, so we use a first-order Taylor series to approximate it. This also gives h a more intuitive explanation.

$$\begin{aligned}
& L_{\xi_n,l} - L_{\xi_{n+1},l} && (15)\\
=\;& L_{\xi_n,l} - L_{\xi_n - \eta\nabla_\xi L_{\xi,u},l} && (16)\\
=\;& L_{\xi_{n+1}+\eta\nabla_\xi L_{\xi,u},l} - L_{\xi_{n+1},l} && (17)\\
\approx\;& L_{\xi_{n+1},l} + \nabla L_{\xi_{n+1}} \cdot \eta\nabla_\xi L_{\xi,u} - L_{\xi_{n+1},l} && (18)\\
=\;& \eta\nabla L_{\xi_{n+1},l}\nabla_\xi L_{\xi,u} && (19)\\
=\;& h && (20)
\end{aligned}$$

As shown above, We use the first order Taylor approximation in equation (18). When $h < 0$, it means that the loss on the label data increases after training with pseudo labels, indicating that the pseudo label is incorrect at this time. Otherwise, if $h > 0$, it means that the pseudo labels reduce the value of loss of labeled data, and the teacher should be encouraged to generate more such pseudo labels.

In the experiment, we notice that what matters is the sign of h rather than its absolute value. Positive values imply positive feedback and encourage teachers to continue with the same pseudo-labeled predictions, while negative values do the opposite and teachers should correct their output. Also, we find it is difficult to estimate h accurately due to the existence of momentum and other randomized regularization methods such as dropout. So we adopt a discrete approach, making the real feedback coefficient $h_{Discrete}$ as a function of h:

$$h_{Discrete} = \begin{cases} 1, h > 0 \\ -0.1, h \leq 0 \end{cases} \tag{21}$$

The rest of the details can be found in the paper.