



$$\sigma = \text{match}(\hat{y}^{main}, y)$$

$$\mathcal{L}^{main} = \mathcal{L}(\hat{y}^{main}|_{\sigma}, y) \quad \mathcal{L}^{aux} = \mathcal{L}(\hat{y}^{aux}|_{\sigma}, y)$$

$$C_1: \{q \in \mathcal{D}_q | s(q) > \tau_q\} \quad C_2: \text{TopK}(\{q \in \mathcal{D}_q | s(q) < \tau_q\}, N_{tr})$$