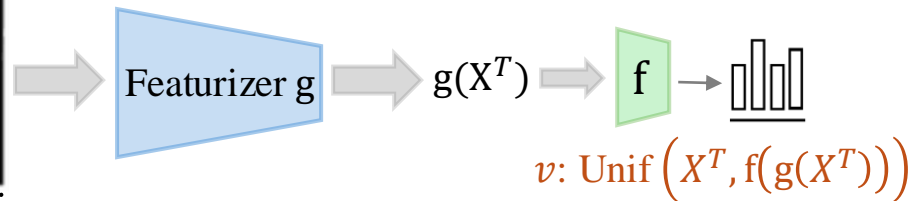


$$\text{CE}(Y^S, f(g(X^S)))$$



$$+ \text{OT}(u, v)$$

Training loss.