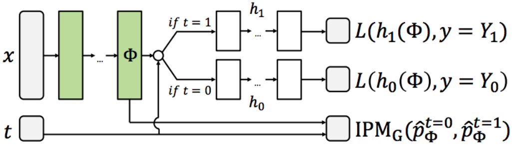
5. Causality, Disentangled Representations and Causal Reinforcement Learning

Causality and Meta-learning

References - *Learning Representations for Counterfactual Inference* and *MetaCI: Meta-learning for Causal Inference ...*

Treatment is imbalanced due to confounding features, i.e. $p(t|x) \neq p(\bar{t}|x)$: So we learn a representation $\phi(x)$ in which $p(t|\phi(x))$ and $p(\bar{t}|\phi(x))$ are close. Goal is to predict average treatment effect $ATE = \mathbb{E}|Y_1(x) - Y_0(x)|$. Note: one of these is 'counterfactual' and always unknown in real data for given x, t!



Batch of n_1 treatments and n_0 non-treatments, minimize loss

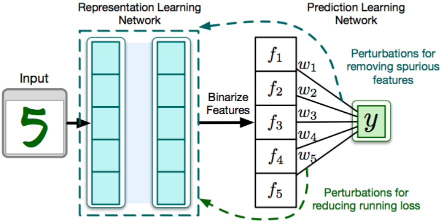
$$\mathcal{L} = \tfrac{1}{n_1} L(h_1(\phi) + \tfrac{1}{n_0} L(h_0(\phi)) + \|f(\phi)\|_2; \quad f_W = \frac{1}{n_0}\sum_{i=0}^{n_0-1} \phi_W(x_i) - \frac{1}{n_1}\sum_{i=n_0}^{n_1} \phi_W(x_i)$$

MetCI applies MAML-like meta-learning across tasks differing in $p(t|x)$ and $p(y|x,t)$, so as to estimate ATE *faster* on new tasks.

Online Meta-learning of Invariant (Causal) Features

References - *Learning Causal Features Online*

Correlation between a causal feature and target should not change as different parts of the source distribution $p(x)$ are experienced. Note: assumption is that $x = o(s)$ is observed from a 'true' set of features s, and $p(y|s)$ does not change over time. Further, learning is online: $w^t = w^{t-1} - \eta \nabla_{w^{t-1}} \mathcal{L}(f^T w^{t-1}, y_t)$.



Track mean u_i and variance v_i of w_i:

if $f_i = 1$**:**

$u_i^t = \alpha u_i^{t-1} + (1-\alpha) w_i$

$v_i^t = \beta v_i^{t-1} + (1-\beta)(w_i^t - u_i^t)(w_i^t - u_i^{t-1})$

Mask f_i using $\sigma(g_i)$; $g_i^0 = 0$ &

$\Delta g_i = -\frac{1}{|v_i|^2}(v_i - \sum_{j=1}^{n} v_j)$ every N steps.

Weights in the representation network are randomly set to $0, 1, -1$ during learning; the change is kept if either overall v or running loss decreases.