Towards Accurate Model Selection
in Deep Unsupervised Domain Adaptation

Kaichao You[1] , Ximei Wang[1], Mingsheng Long[1], Michael I. Jordan[2]

[1]School of Software, Tsinghua University
[1]National Engineering Lab for Big Data Software
[2]University of California, Berkeley

Machine Learning and Applications 2019

previously presented at ICML 2019

Validation in Unsupervised Domain Adaptation (UDA): the problem

- Supervised Learning
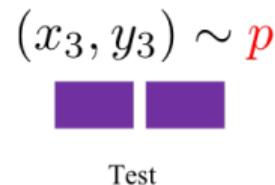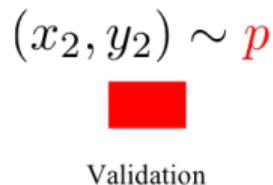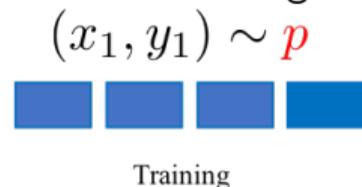
$$(x_1, y_1) \sim p \qquad\qquad (x_2, y_2) \sim p \qquad\qquad (x_3, y_3) \sim p$$


Training

Validation

Test

- Supervised Learning

$(x_1, y_1) \sim p$

$(x_2, y_2) \sim p$

$(x_3, y_3) \sim p$

Training

Validation

Test

- Unsupervised Domain Adaptation

Source Domain

Target Domain

$(x_1, y_1) \sim p$

$(x_2, y_2) \sim q$

(x_1, y_1)

Validation

x_2

(x_2, y_2)

Training

Test
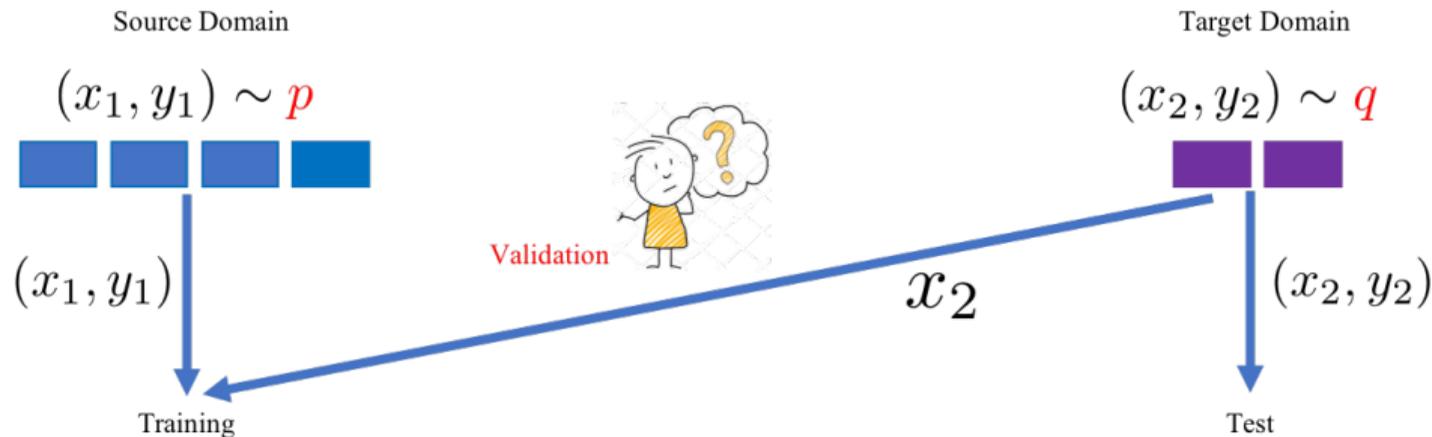
Formalization and Key Insights

- Model Selection: estimate Target Risk $\mathcal{R}(g) = \mathbb{E}_{\mathbf{x} \sim q} \ell(g(\mathbf{x}), y)$

Formalization and Key Insights

- Model Selection: estimate Target Risk $\mathcal{R}(g) = \mathbb{E}_{\mathbf{x} \sim q} \ell(g(\mathbf{x}), y)$
- Covariate Shift Assumption $p(y|\mathbf{x}) = q(y|\mathbf{x})$

Formalization and Key Insights

- Model Selection: estimate Target Risk $\mathcal{R}(g) = \mathbb{E}_{\mathbf{x} \sim q} \ell(g(\mathbf{x}), y)$
- Covariate Shift Assumption $p(y|\mathbf{x}) = q(y|\mathbf{x})$
- Importance Weighted

$$\mathbb{E}_{\mathbf{x} \sim p} w(\mathbf{x}) \ell(g(\mathbf{x}), y) = \mathbb{E}_{\mathbf{x} \sim p} \frac{q(\mathbf{x})}{p(\mathbf{x})} \ell(g(\mathbf{x}), y) = \mathbb{E}_{\mathbf{x} \sim q} \ell(g(\mathbf{x}), y) = \mathcal{R}(g)$$

Formalization and Key Insights

- Model Selection: estimate Target Risk $\mathcal{R}(g) = \mathbb{E}_{\mathbf{x} \sim q} \ell(g(\mathbf{x}), y)$
- Covariate Shift Assumption $p(y|\mathbf{x}) = q(y|\mathbf{x})$
- Importance Weighted

$$\mathbb{E}_{\mathbf{x} \sim p} w(\mathbf{x}) \ell(g(\mathbf{x}), y) = \mathbb{E}_{\mathbf{x} \sim p} \frac{q(\mathbf{x})}{p(\mathbf{x})} \ell(g(\mathbf{x}), y) = \mathbb{E}_{\mathbf{x} \sim q} \ell(g(\mathbf{x}), y) = \mathcal{R}(g)$$

- Embed adapted feature representations

Formalization and Key Insights

- Model Selection: estimate Target Risk $\mathcal{R}(g) = \mathbb{E}_{\mathbf{x} \sim q} \ell(g(\mathbf{x}), y)$
- Covariate Shift Assumption $p(y|\mathbf{x}) = q(y|\mathbf{x})$
- Importance Weighted

$$\mathbb{E}_{\mathbf{x} \sim p} w(\mathbf{x}) \ell(g(\mathbf{x}), y) = \mathbb{E}_{\mathbf{x} \sim p} \frac{q(\mathbf{x})}{p(\mathbf{x})} \ell(g(\mathbf{x}), y) = \mathbb{E}_{\mathbf{x} \sim q} \ell(g(\mathbf{x}), y) = \mathcal{R}(g)$$

- Embed adapted feature representations
- Variance control techniques

Thanks!

Code available at github.com/thuml/Deep-Embedded-Validation