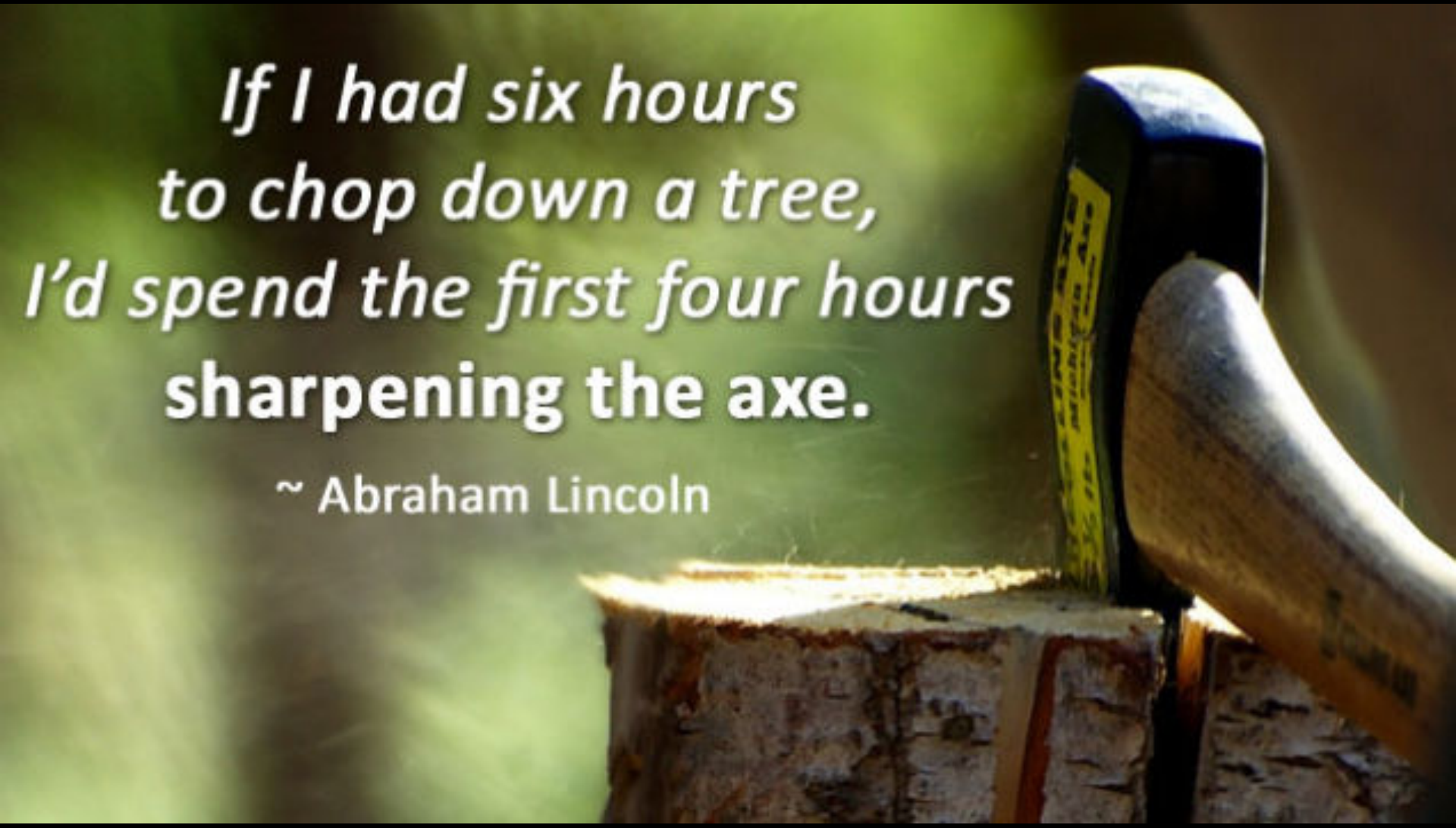
Pessimism is Great for Machine Learning!

Brian Ziebart

bziebart@uic.edu

THE UNIVERSITY OF ILLINOIS AT CHICAGO UIC

This project supports the PhD thesis work of Anqi Liu (aliu33@uic.edu)


If I had six hours
to chop down a tree,
I'd spend the first four hours
sharpening the axe.
~ Abraham Lincoln

Training

M samples

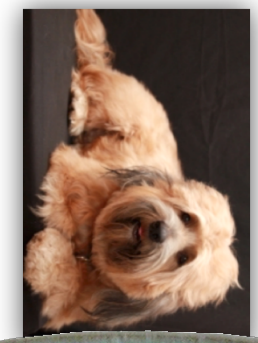
Sample dist.
$\tilde{P}(\mathbf{x}, y)$

Training

Testing

Prediction: $\hat{y}(\mathbf{x})$

Loss: loss($\hat{y}$, y)



Sample dist.
$\tilde{P}(\mathbf{x}, y)$

Testing

Prediction: $\hat{y}(\mathbf{x})$

Loss: $\text{loss}(\hat{y}, y)$

Expected Loss:
$E_P[\text{loss}(\hat{y}(\mathbf{X}), Y)]$

$\hat{y}$ \ y	Dog	Cat	Car
Dog	0	1	1
Cat	1	0	1
Car	1	1	0

$\hat{y}$ \ y	Dog	Cat	Car
Dog	0	1	5
Cat	1	0	5
Car	5	5	0

Sample dist.
$\tilde{P}(\mathbf{x}, y)$

True dist.
$P(\mathbf{x}, y)$

Empirical Risk Minimization

Minimize approximate loss on exact training data

Adversarial Risk Minimization

Minimize exact loss on approximate training data

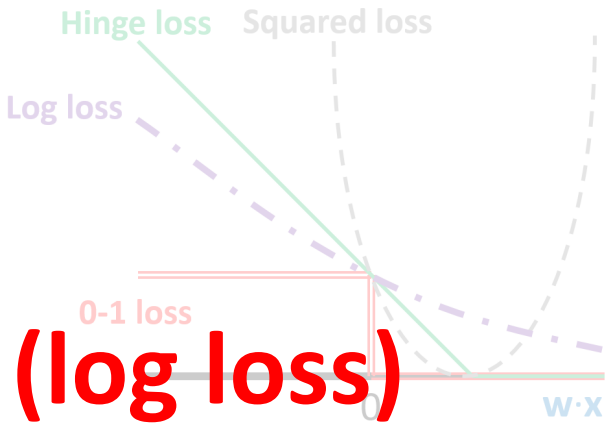
$\hat{y}$ \ y	Dog	Cat	Car
Dog	0	1	5
Cat	1	0	5
Car	5	5	0

Find equilibrium $\hat{P}$, P

[Topsøe 1979, Lanckriet et al. 2002,Grünwald & Dawid 2004, Delage & Ye 2010, Asif et al. 2015, Fathony et al. 2016]

Empirical Risk Minimization

Minimize approximate loss on exact training data



Sometimes equivalent (log loss)

Pessimism is safer/better in general

Adversarial Risk Minimization

Minimize exact loss on approximate training data

$\hat{y}$ \ $\check{y}$	Dog	Cat	Car
Dog	$0+\psi_{dog}$	$1+\psi_{cat}$	$5+\psi_{car}$
Cat	$1+\psi_{dog}$	$0+\psi_{cat}$	$5+\psi_{car}$
Car	$5+\psi_{dog}$	$5+\psi_{cat}$	$0+\psi_{car}$

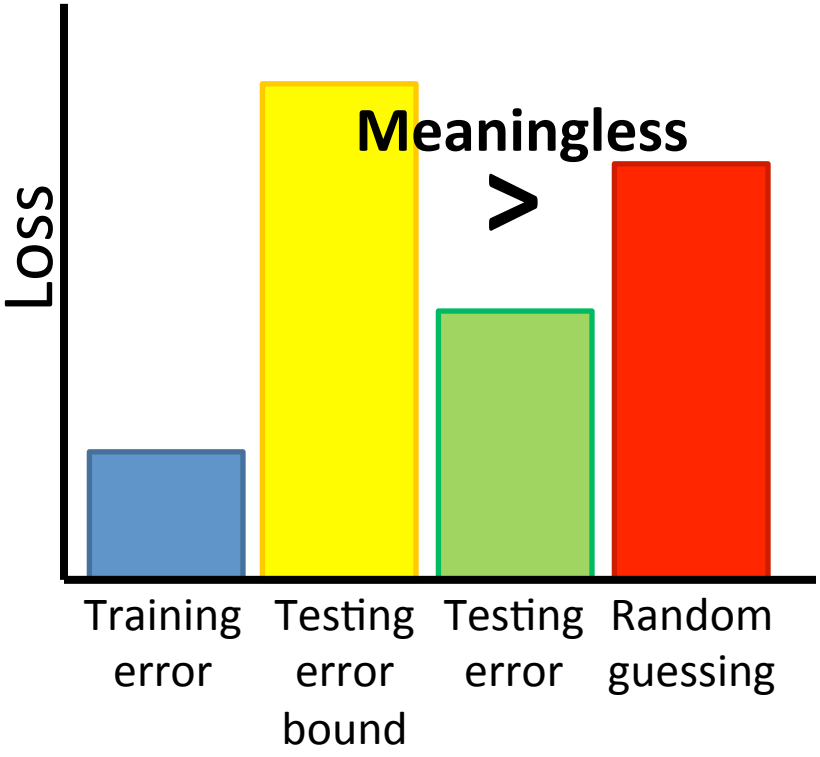
Find equilibrium $\hat{P}$, $\check{P}$

$\psi_{dog} = \mathbf{w}_{dog} \cdot \mathbf{f}(\mathbf{x})$

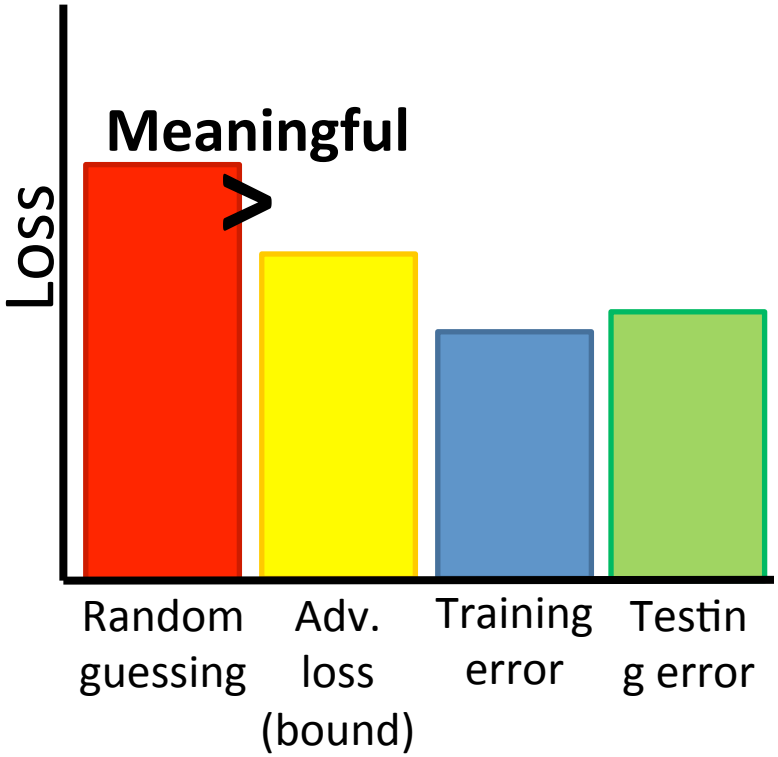
[Topsøe 1979, Lanckriet et al. 2002,Grünwald & Dawid 2004, Delage & Ye 2010, Asif et al. 2015, Fathony et al. 2016]

Meaningful Generalization Bounds

Empirical Risk Minimization

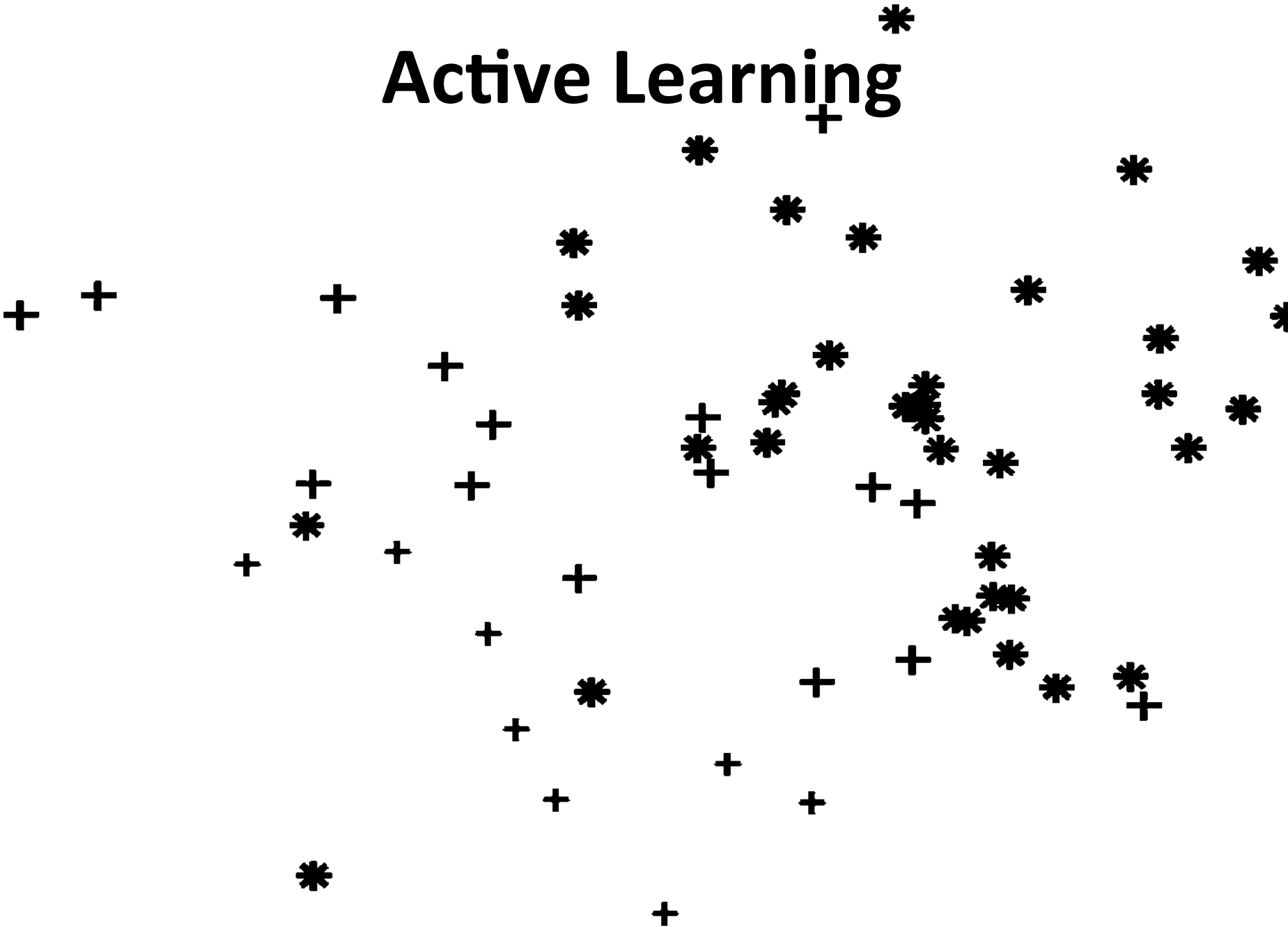


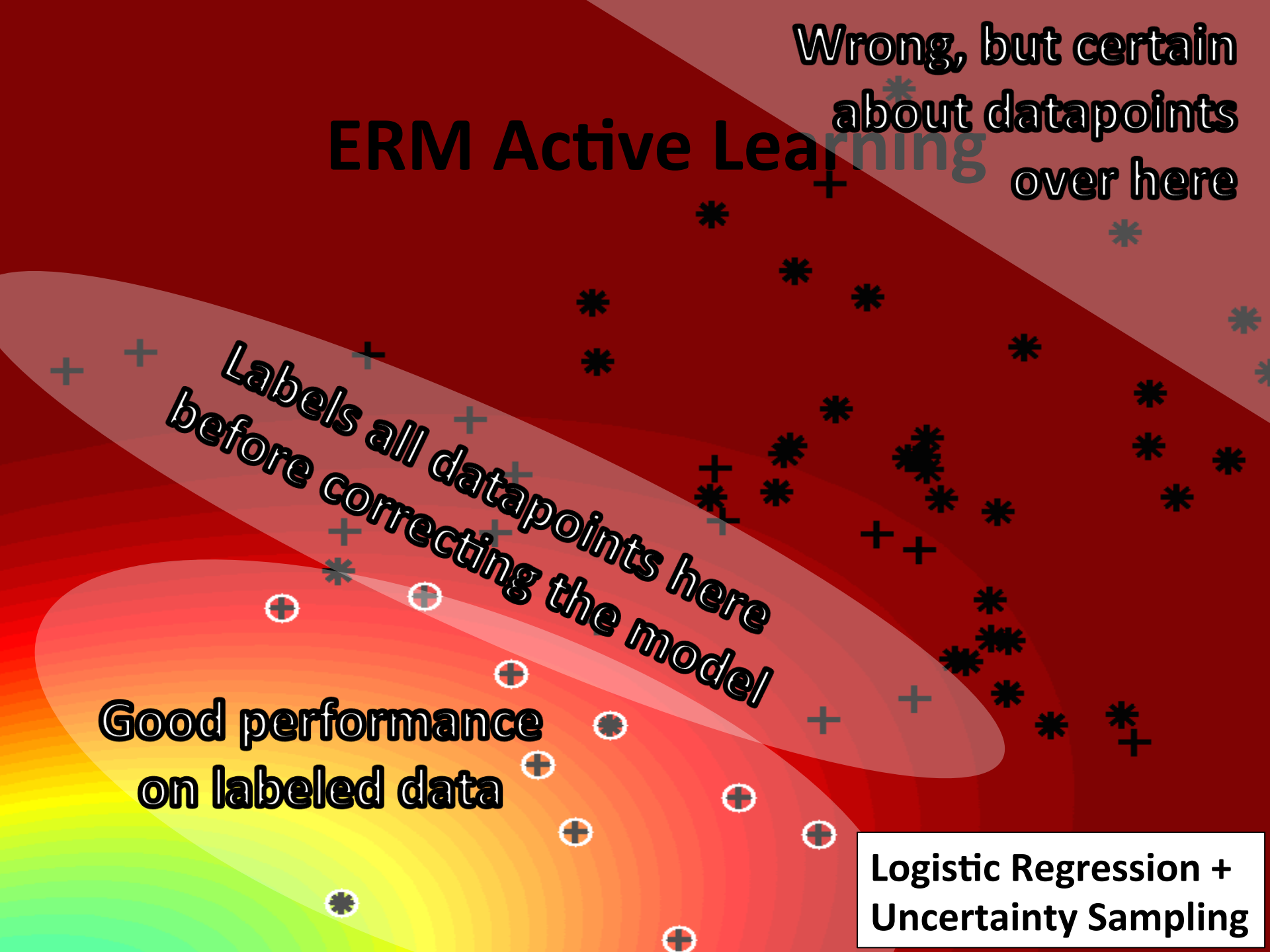
Adversarial Risk Minimization



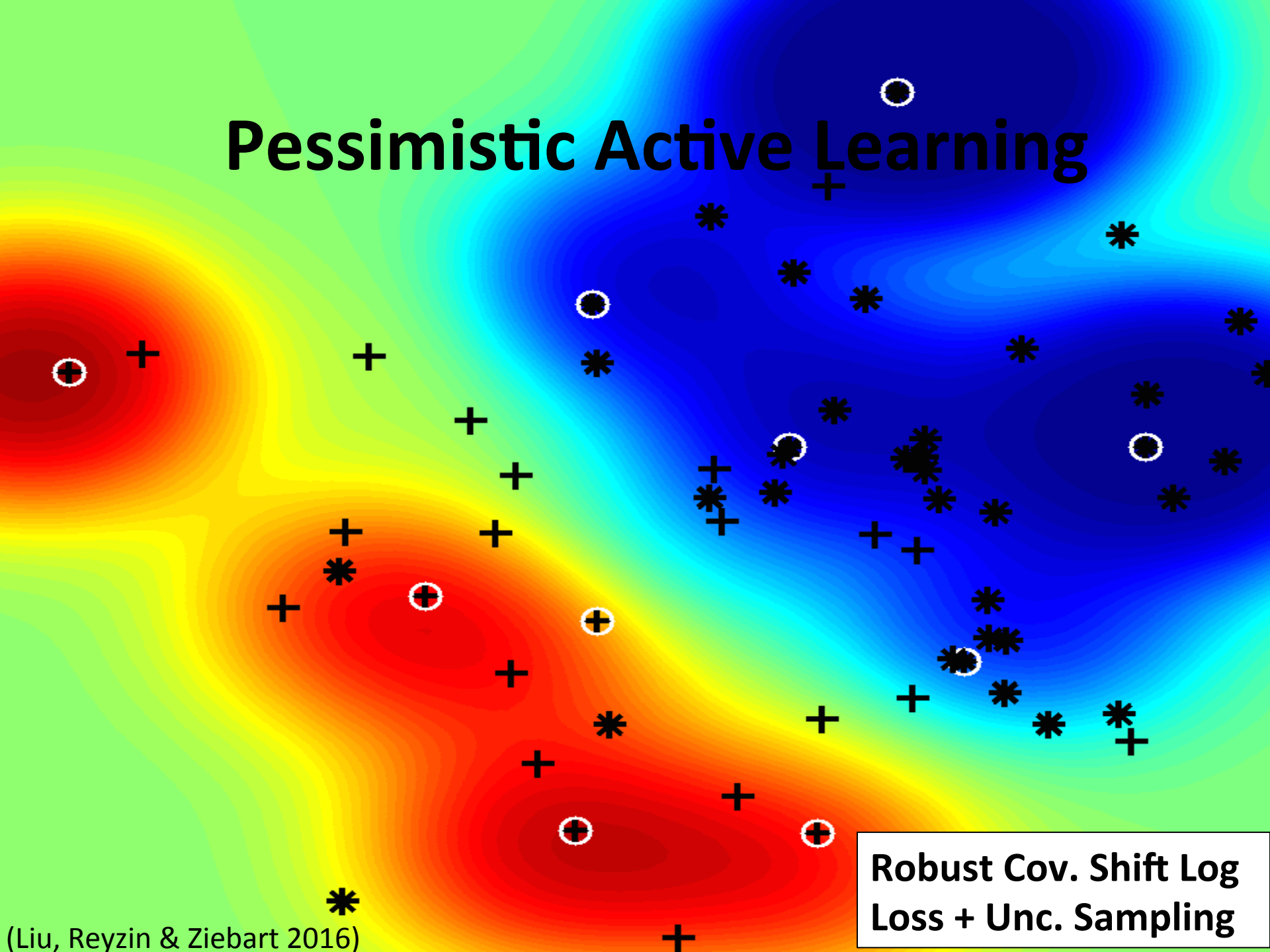
Especially important when training and testing distributions differ (covariate shift) [Liu & Ziebart 2014, Liu et al. 2015, Chen et al. 2016]

Active Learning



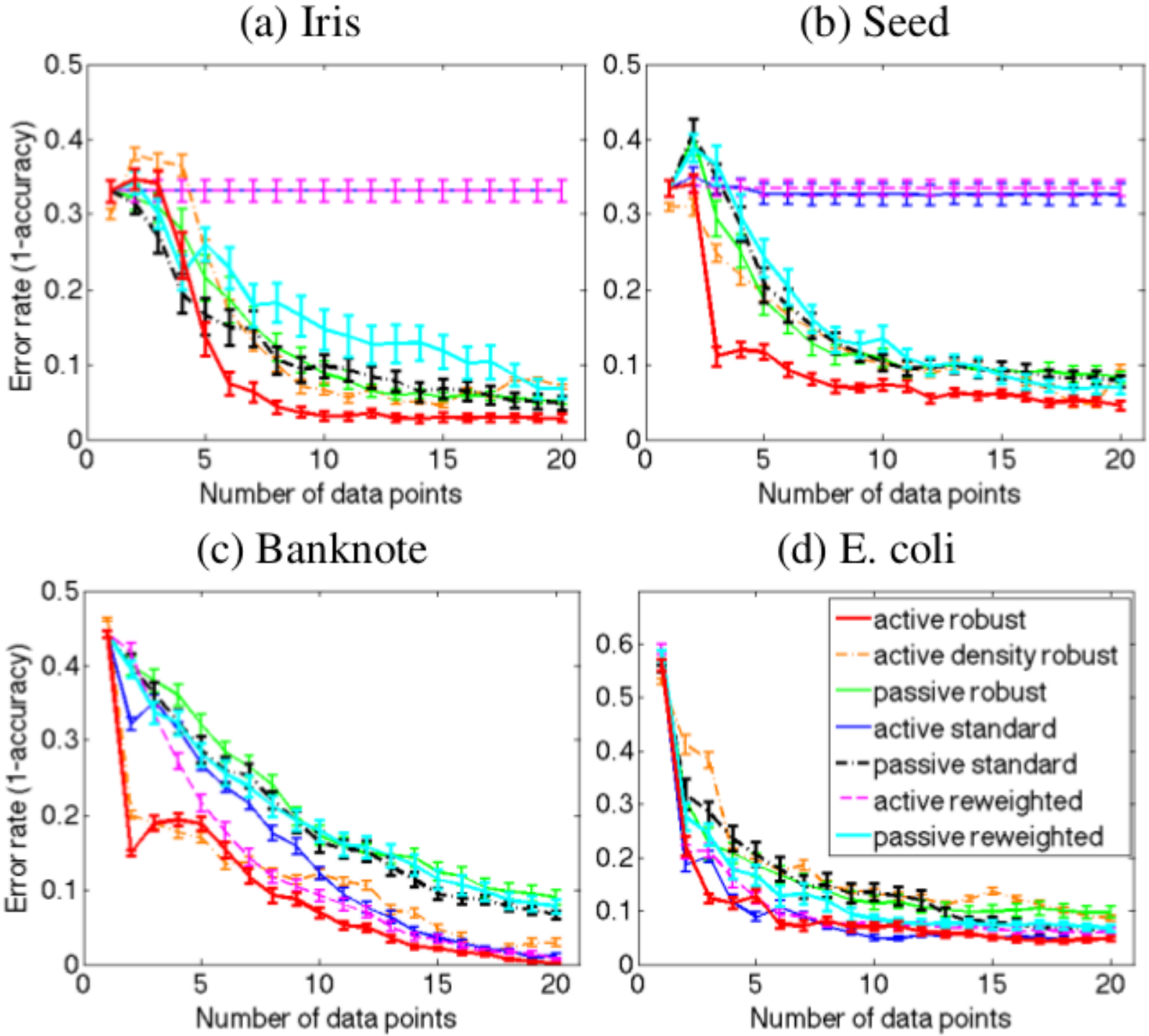

ERM Active Learning
Wrong, but certain about datapoints over here
Labels all datapoints here before correcting the model
Good performance on labeled data
Logistic Regression + Uncertainty Sampling

Pessimistic Active Learning



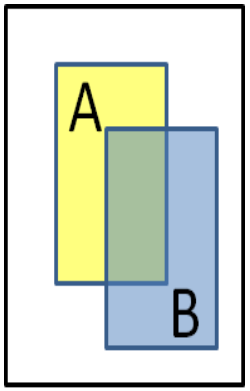
(Liu, Reyzin & Ziebart 2016)

Robust Cov. Shift Log Loss + Unc. Sampling


(a) Iris
(b) Seed
(c) Banknote
(d) E. coli
Error rate (1-accuracy)
Number of data points
active robust
active density robust
passive robust
active standard
passive standard
active reweighted
passive reweighted

Pessimism Better for Specialized Losses

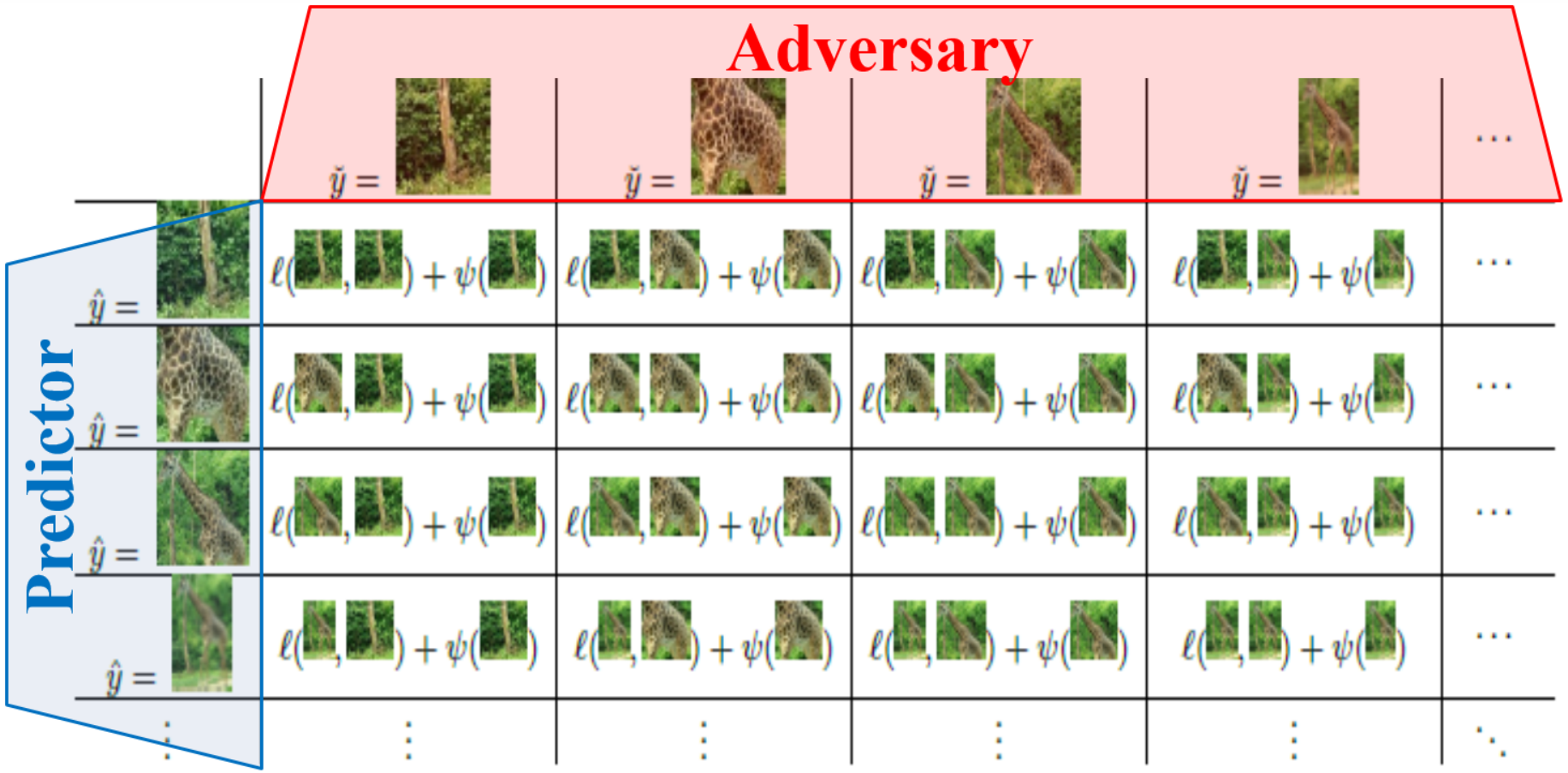
Object localization: performance based on overlap



Thresholded overlap:
Incorrect if overlap < 70%
Correct if overlap ≥ 70%

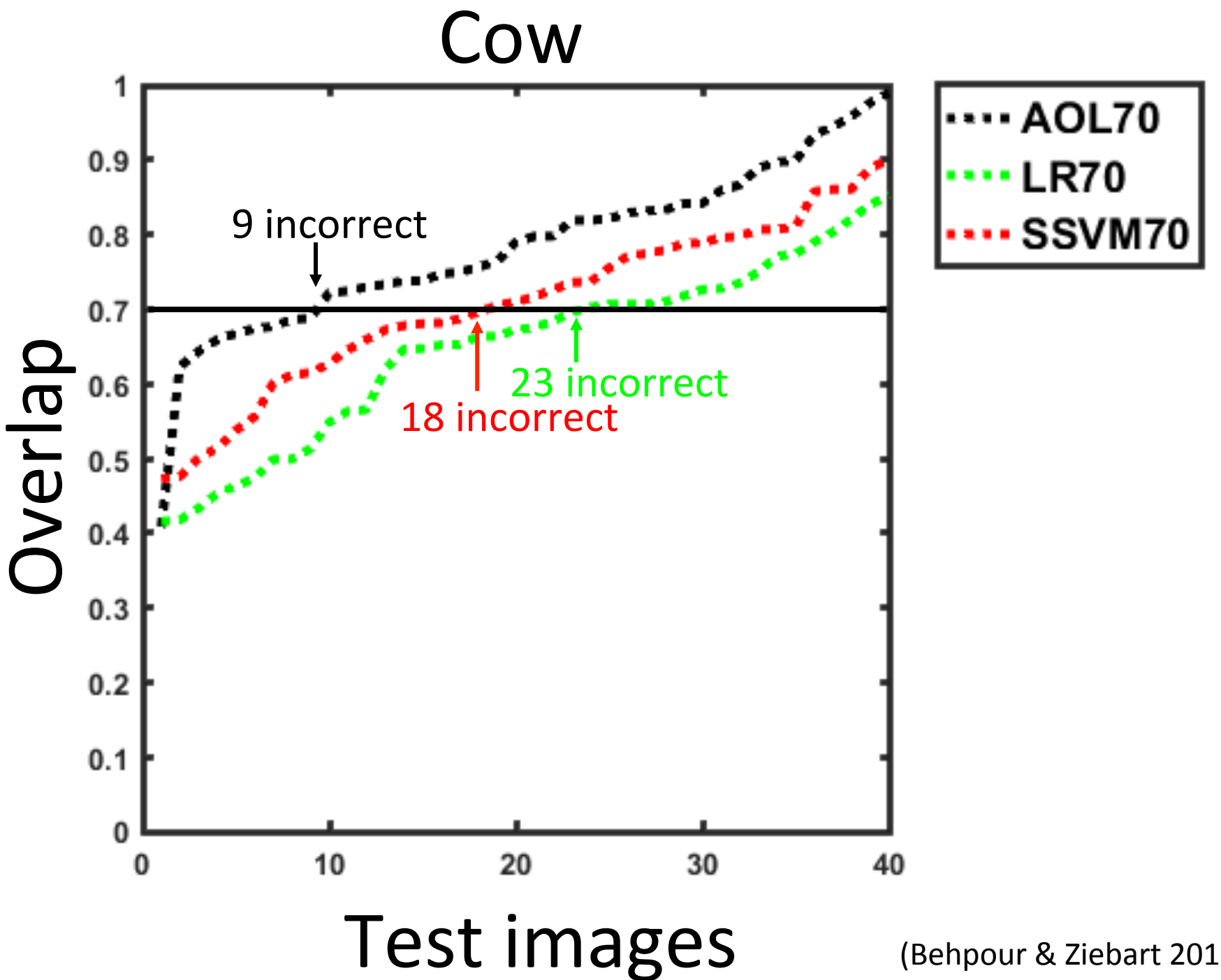
Methods: Bayes act under logistic regression (LR) _
Hinge loss approximation (SSVM)
Adversarial object localization (AOL)

Experimental setup: Features from pre-trained deep network (VGGnet); EdgeBox proposals; ImageNet dataset

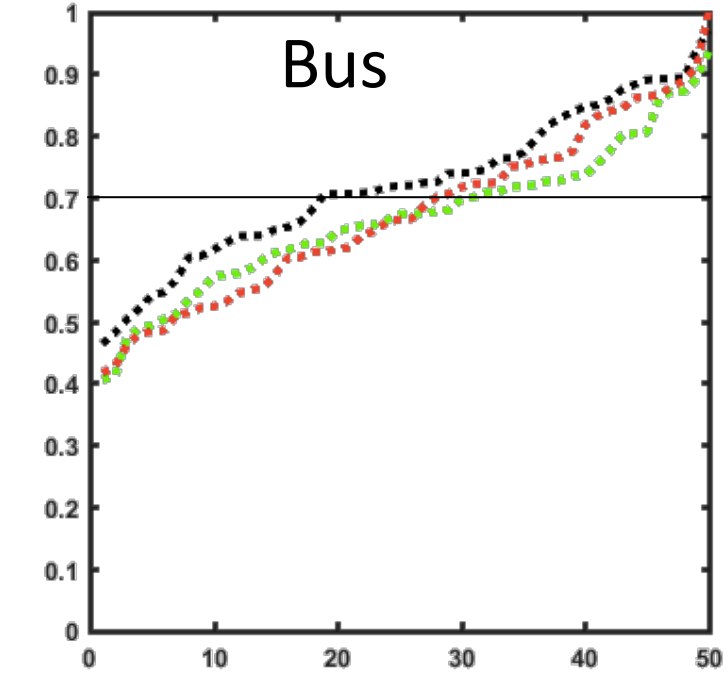


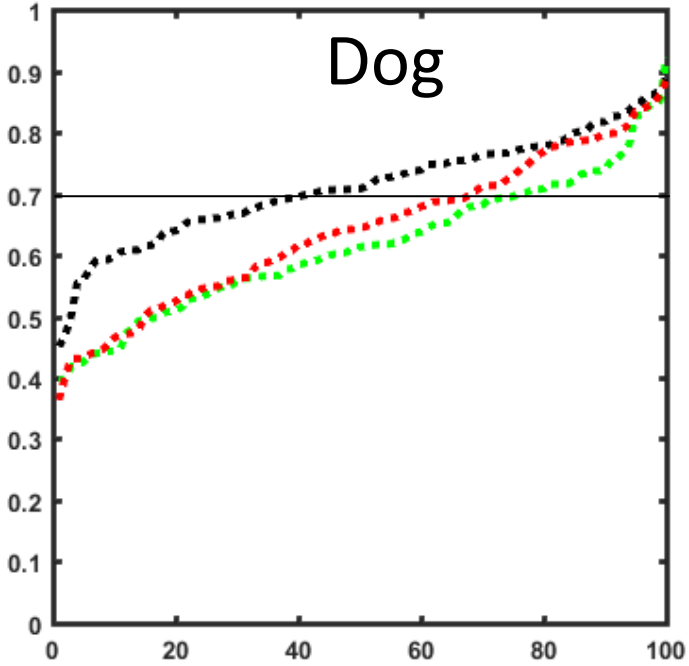
- Prediction game is large (conceptually)
- Efficiently solved using **double oracle method** (McMahan et al. 2003) for constraint generation

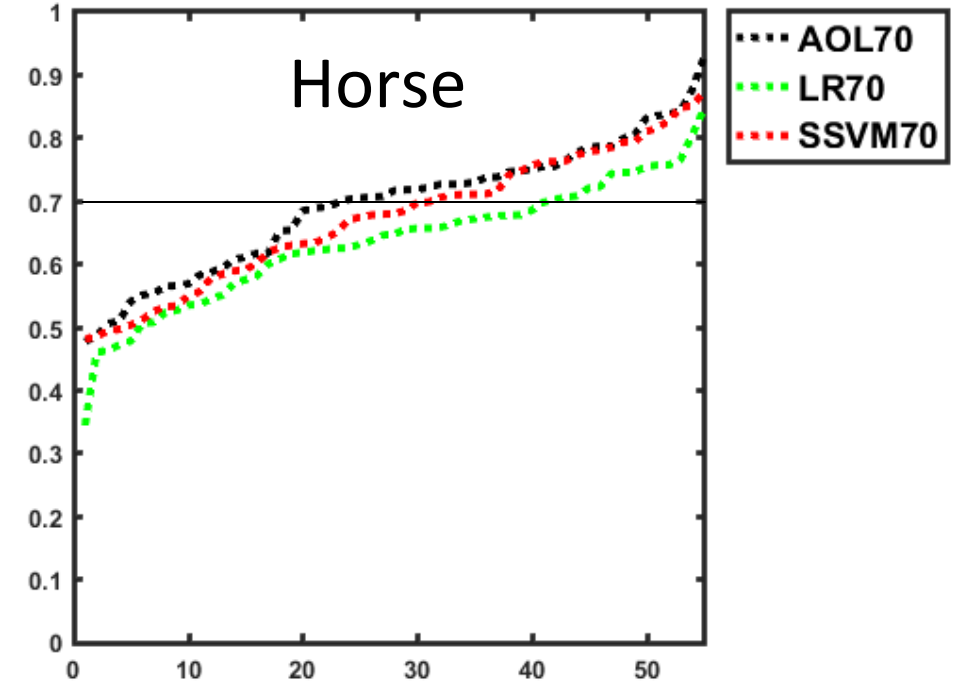
(Behpour & Ziebart 2016)

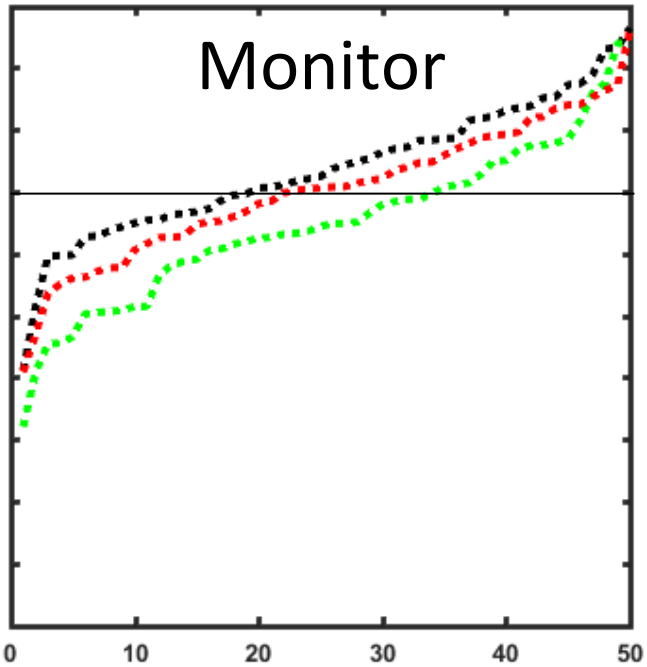


(Behpour & Ziebart 2016)


Bus
AOL70
LR70
SSVM70


Dog
AOL70
LR70
SSVM70


Horse
AOL70
LR70
SSVM70


Monitor
AOL70
LR70
SSVM70

Conclusions

Pessimistic formulation provides benefits:

- In *theory* (**generalization bounds**, **consistency**)
- In *practice* (**custom losses**, **active learning**)

Questions:

- Efficiently solve structured predictions games?
- Effective active learning in high dimensions?
- Stronger theory of robust active learning?

Related papers:

Liu, Ziebart. "*Robust Classification Under Sample Selection Bias*." NIPS 2014.

Liu, Reyzin, Ziebart. "*Shift-Pessimistic Active Learning Using Robust Bias-Aware Prediction*." AAAI 2015.

Asif, Xing, Behpour, Ziebart. "*Adversarial Cost-Sensitive Classification*." UAI 2015.

Wang, Xing, Asif, Ziebart. "*Adversarial Prediction Games for Multivariate Losses*." NIPS 2015.

Li, Asif, Wang, Ziebart. "*Adversarial Sequence Tagging.*" IJCAI 2016.

Chen, Monfort, Liu, Ziebart. "*Robust Covariate Shift Regression*." AISTATS 2016.

Fathony, Liu, Asif, Ziebart. "Adversarial Multiclass Classification: A Risk Minimization Perspective." NIPS 2016.

NRI Award #1227495
IIS Award #1526379