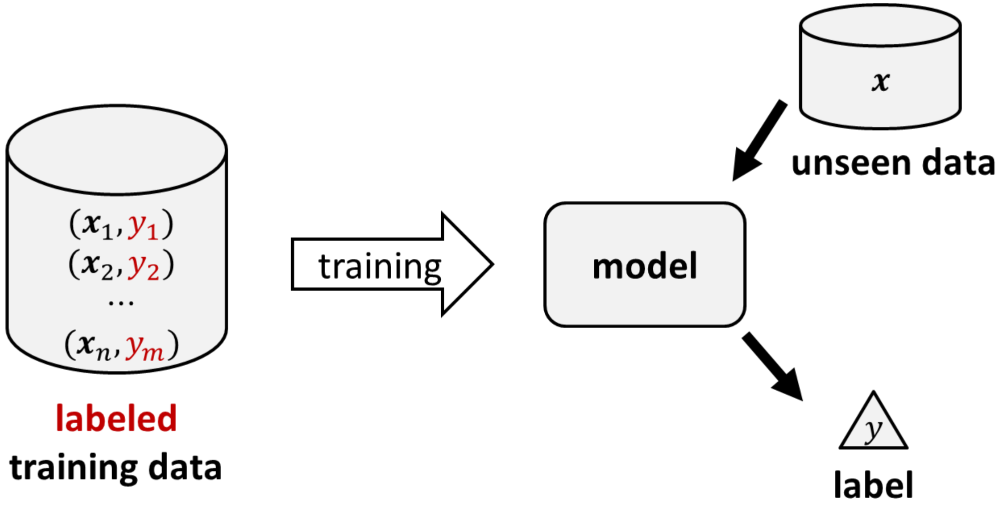
Recent Progress on Active Learning

Sheng-Jun Huang (黄圣君)

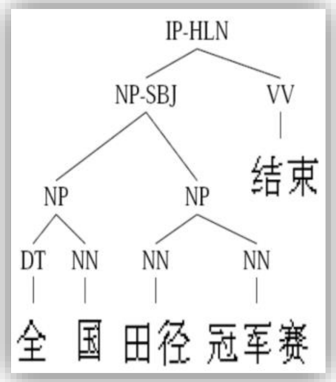
Nanjing University of Aeronautics and Astronautics

2018-4-22 @VALSE

Learning with Fewer Labeled Data



$$R(h) \leq R_{\text{emp}}(h) + \sqrt{\frac{\ln|\mathcal{H}| + \ln\frac{2}{\delta}}{2m}}$$



2 years for 4000 sentences in Penn Chinese Treebank

time consuming

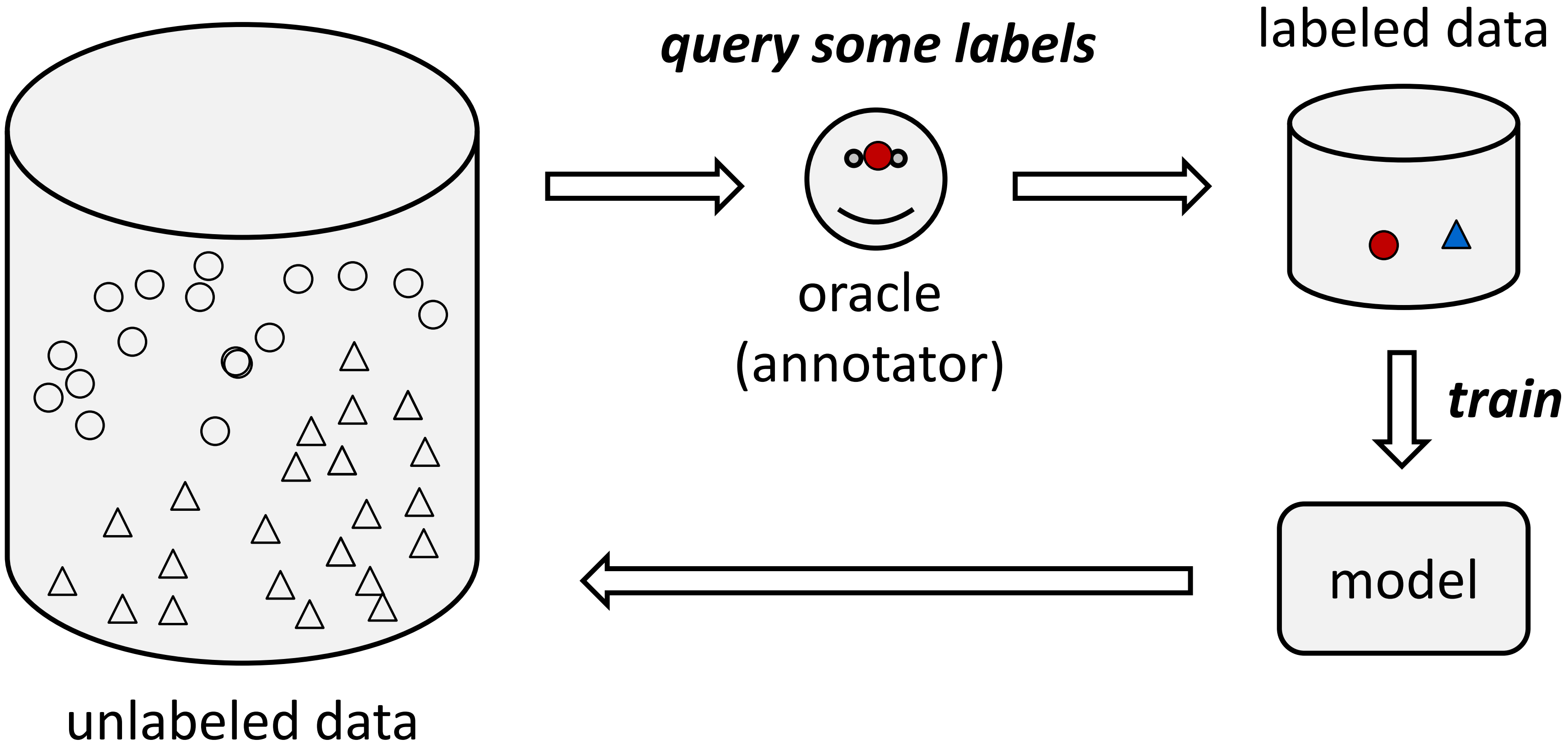
only experts can provide accurate annotations

high expertise

Labeled data is **important** but **expensive**

➔ Can we learn with fewer labeled data?

Active Learning

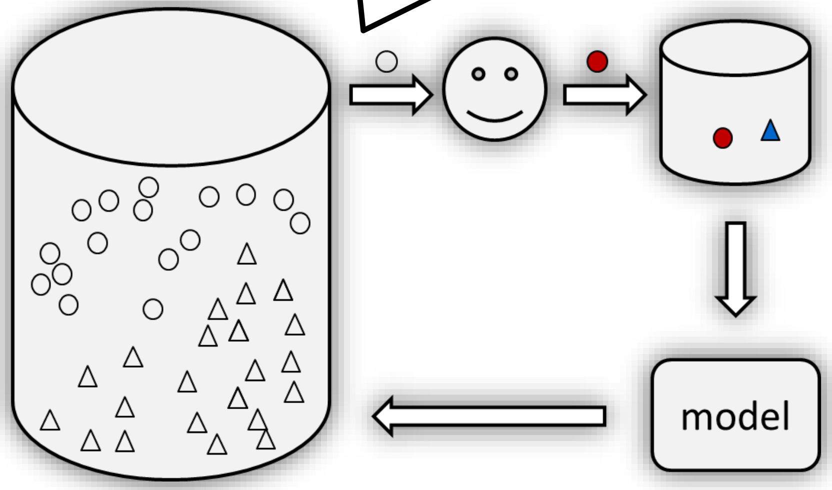


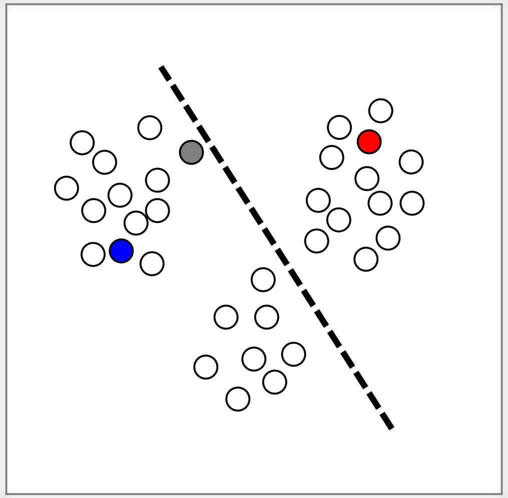
Goal: train an effective model with least labeling cost

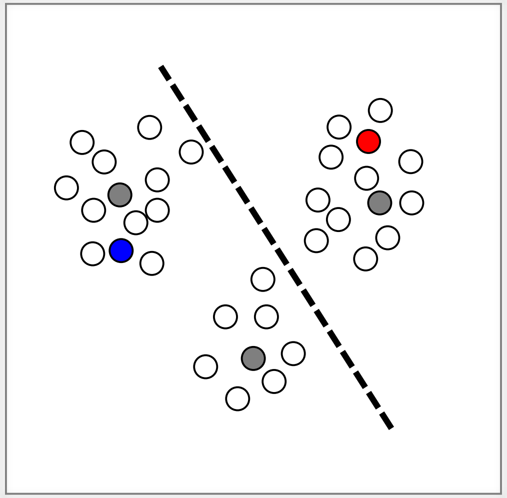
Active Learning

Which instance to select?

- Informative instances
- Representative instances
- Informative & representative instances





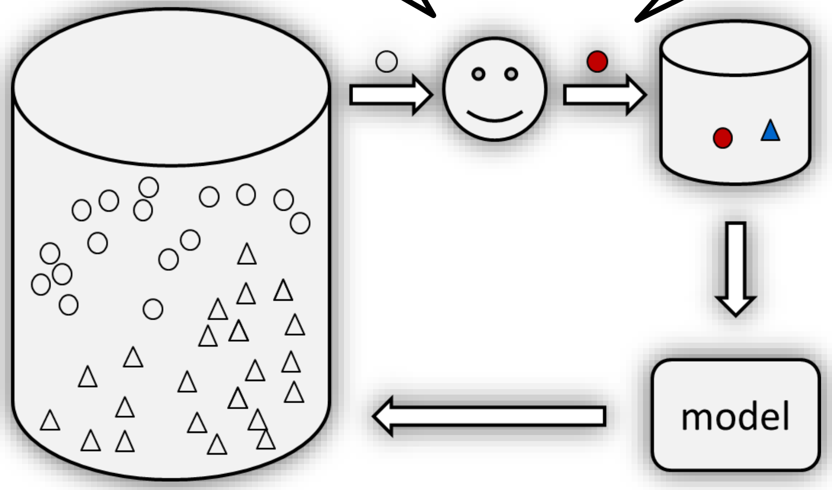


Recent Progress

Weak supervision
The oracle may be noisy or unavailable

Cost sensitive
Care the cost rather than the number

Model dependent
Different models may have diverse needs



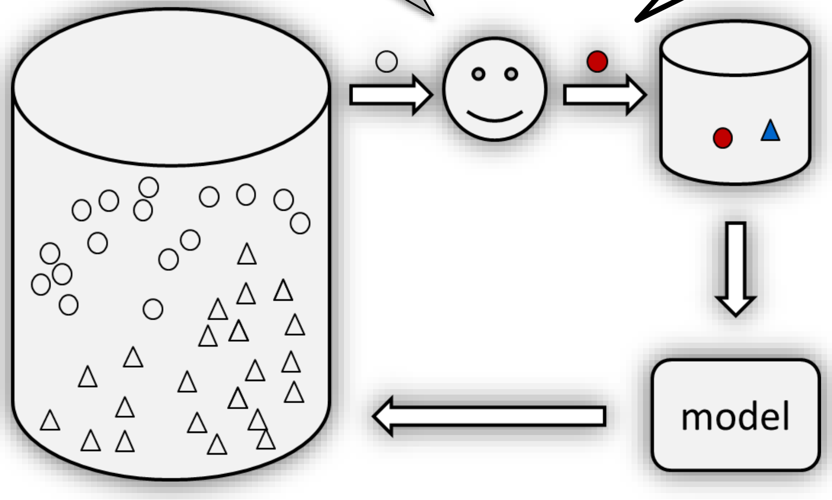
More Practical and More Systematic

Recent Progress

Weak supervision
The oracle may be noisy or unavailable

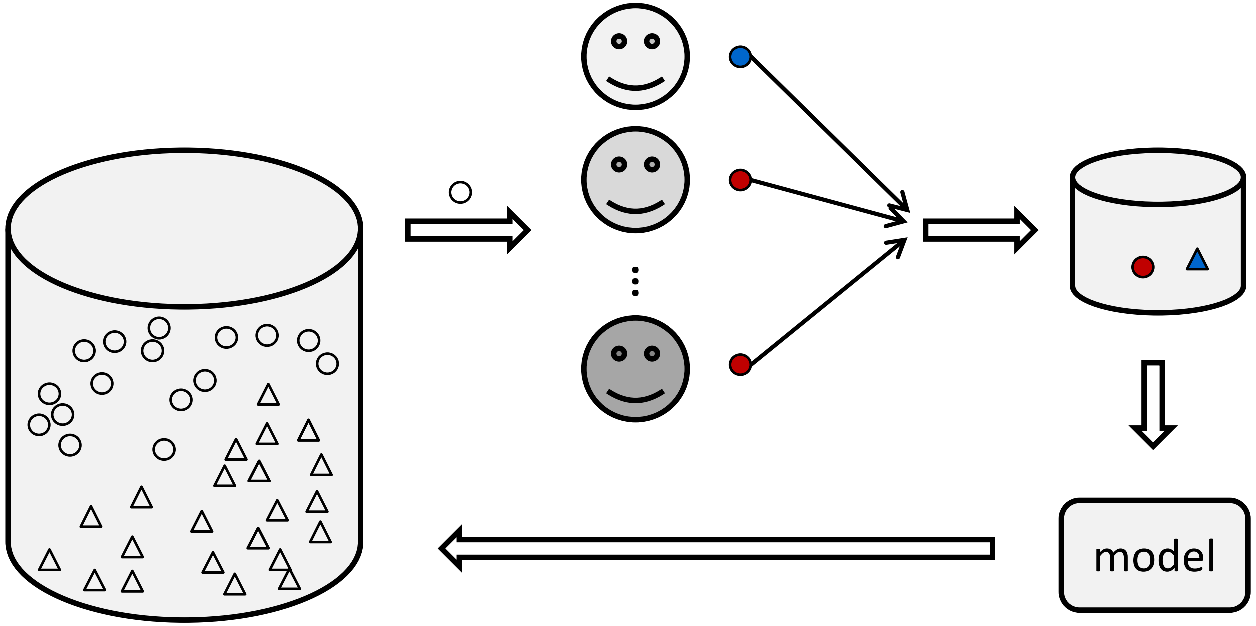
Cost sensitive
Care the cost rather than the number

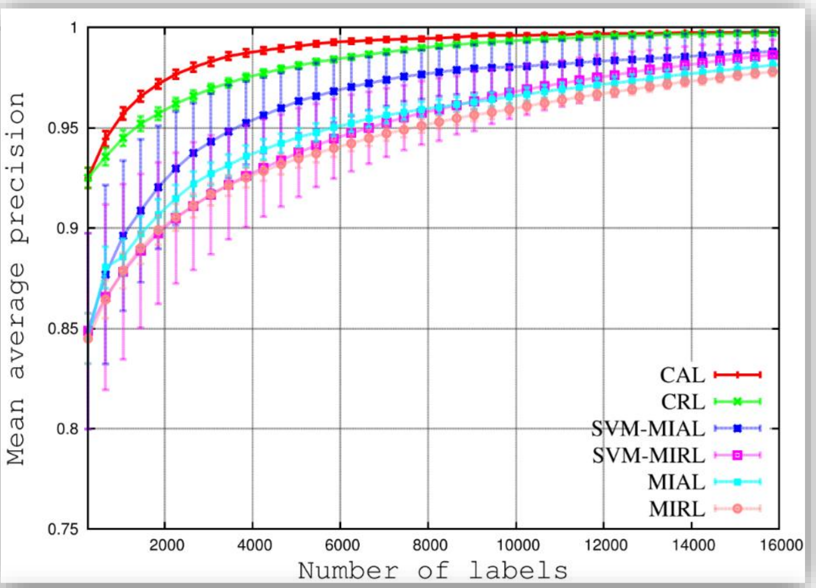
Model dependent
Different models may have diverse needs

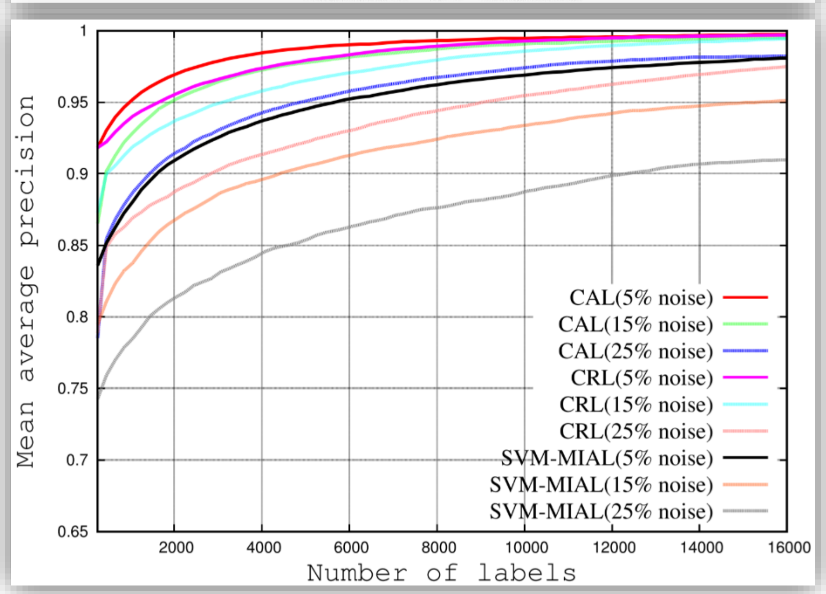


Active learning with Weak Supervision

- Collaborative labeling from crowds
  - Labeler quality estimation
  - Ensemble kernel machine classifier
  - Robust to label noise







Hua et al. Collaborative Active Visual Recognition from Crowds A Distributed Ensemble Approach. PAMI 2018.

Active learning with Weak Supervision

- Pairwise comparison from noisy labelers
  - Leverage both types of oracles
  - Lower querying complexity under different noise conditions

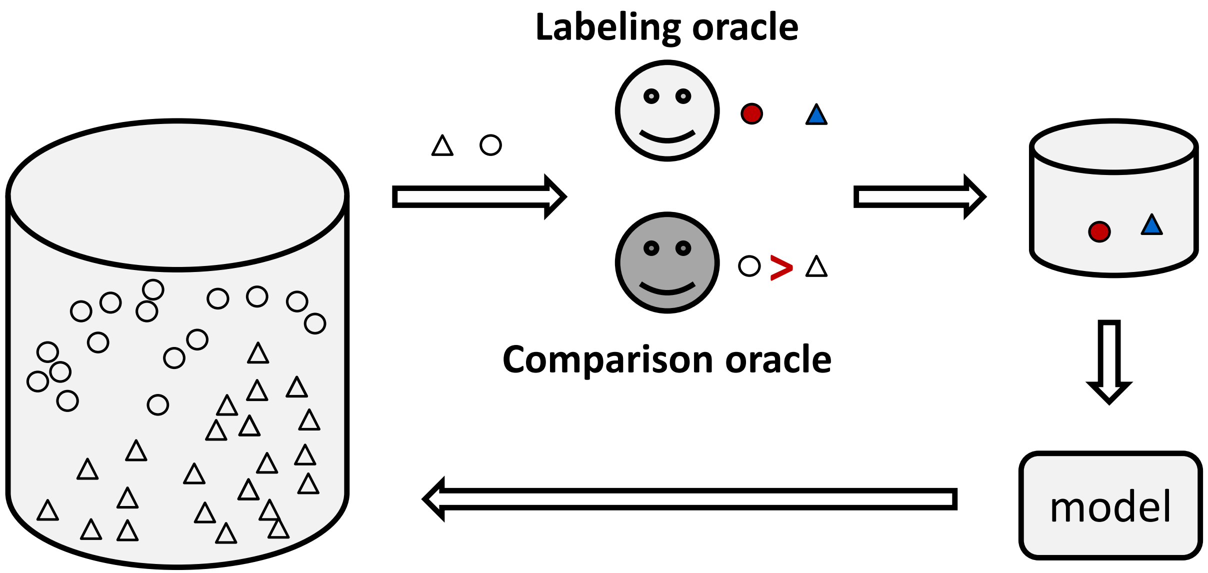


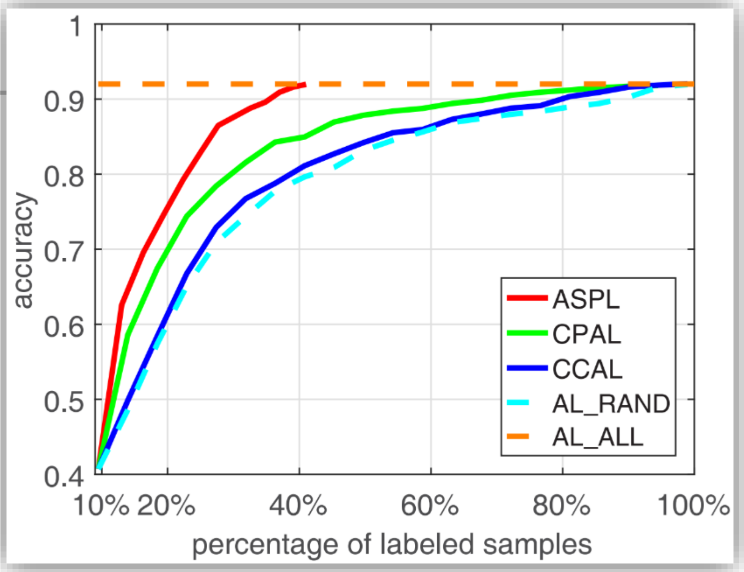
Table 2: Comparison of various methods for learning of halfspaces (Omit

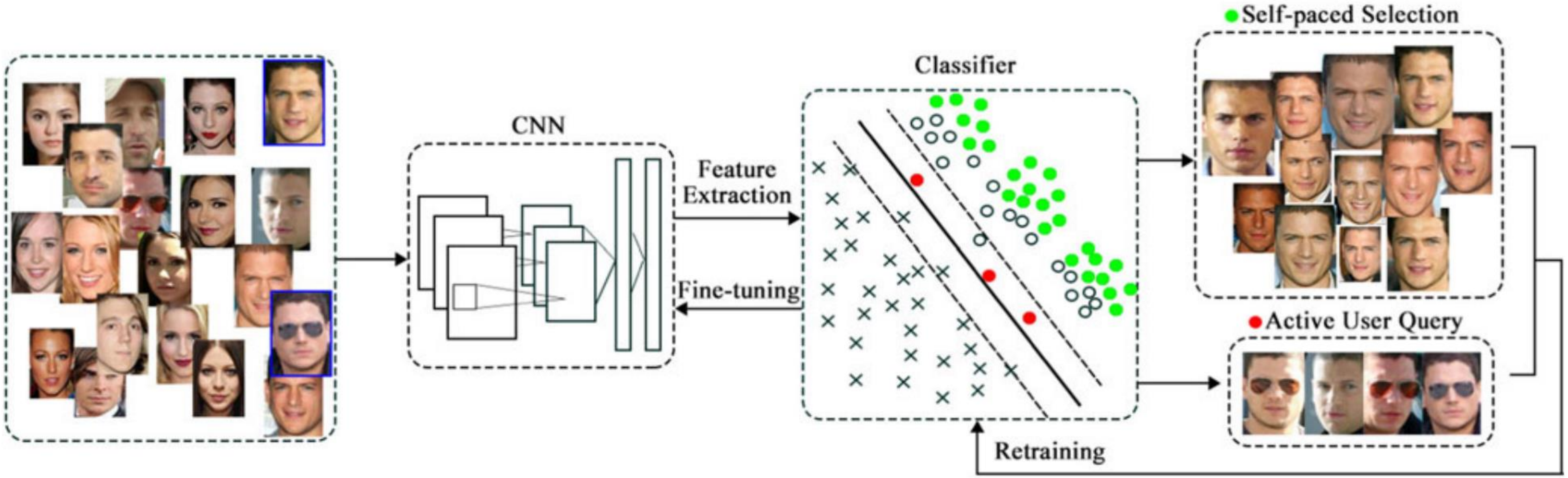
Label Noise	Work	# Label	# Query
Massart	[8]	$\tilde{O}(d)$	$\tilde{O}(d)$
Massart	[5]	$\mathrm{poly}(d)$	$\mathrm{poly}(d)$
Massart	**Ours**	$\tilde{O}(1)$	$\tilde{O}(d)$
Tsybakov (κ)	[19]	$\tilde{O}((\frac{1}{\varepsilon})^{2\kappa-2} d\theta)$	$\tilde{O}((\frac{1}{\varepsilon})^{2\kappa-2} d\theta)$
Tsybakov (κ)	**Ours**	$\tilde{O}\left(\left(\frac{1}{\varepsilon}\right)^{2\kappa-2}\right)$	$\tilde{O}\left(\left(\frac{1}{\varepsilon}\right)^{2\kappa-2}+d\right)$
Adversarial ($\nu = O(\varepsilon)$)	[34]	$\tilde{O}(d)$	$\tilde{O}(d)$
Adversarial ($\nu = O(\varepsilon)$)	[6]	$\tilde{O}(d^2)$	$\tilde{O}(d^2)$
Adversarial ($\nu = O(\varepsilon)$)	**Ours**	$\tilde{O}(1)$	$\tilde{O}(d)$

Xu et al. Noise-Tolerant Interactive Learning Using Pairwise Comparisons. NIPS 2017.

Active learning with Weak Supervision

- Self-paced active learning
  - Self-annotation for high-confident instances
  - Oracle annotation for low-confident instances

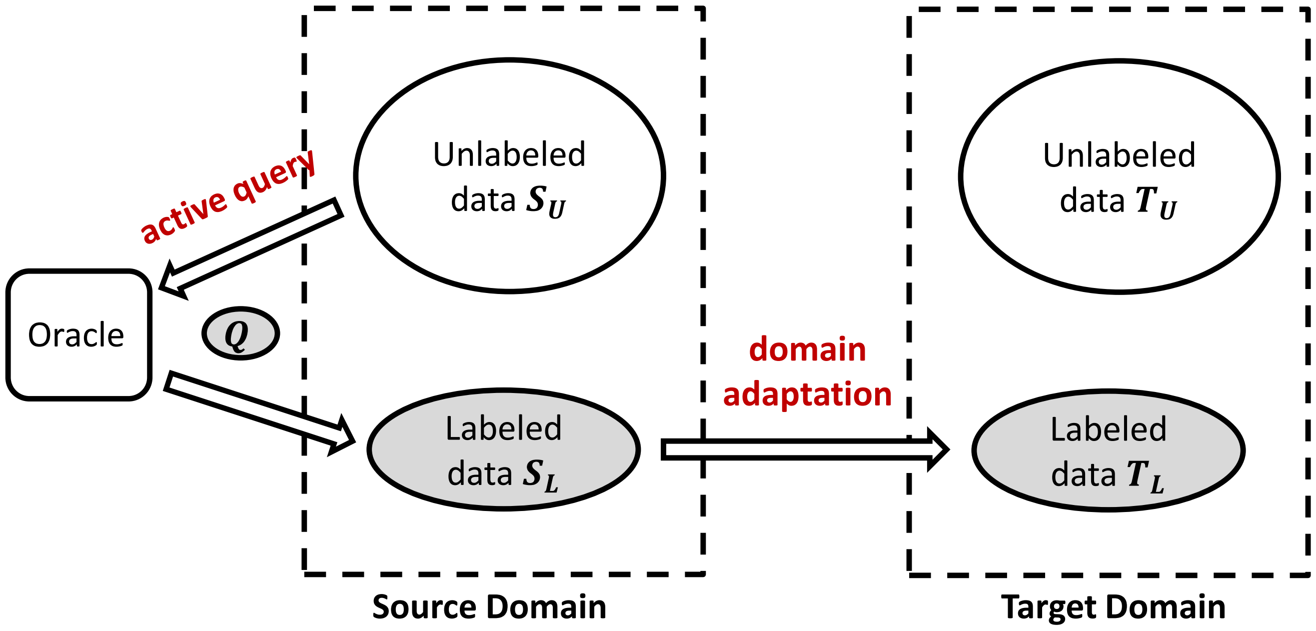


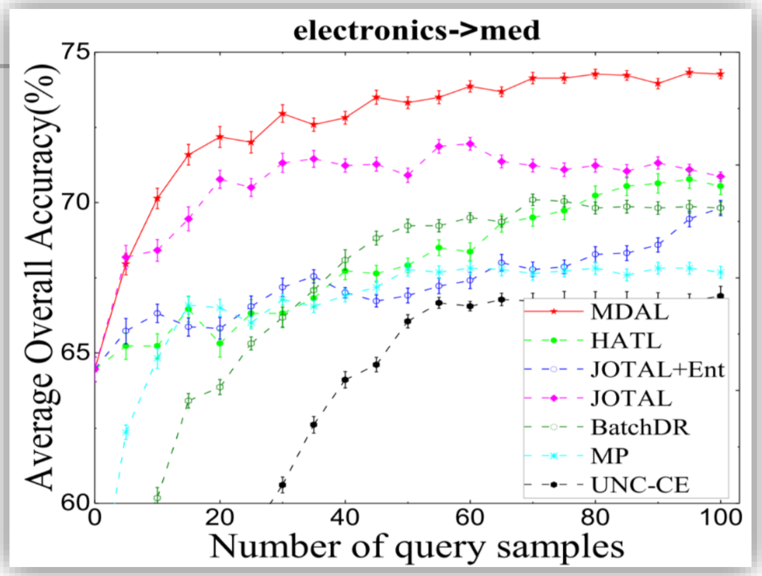


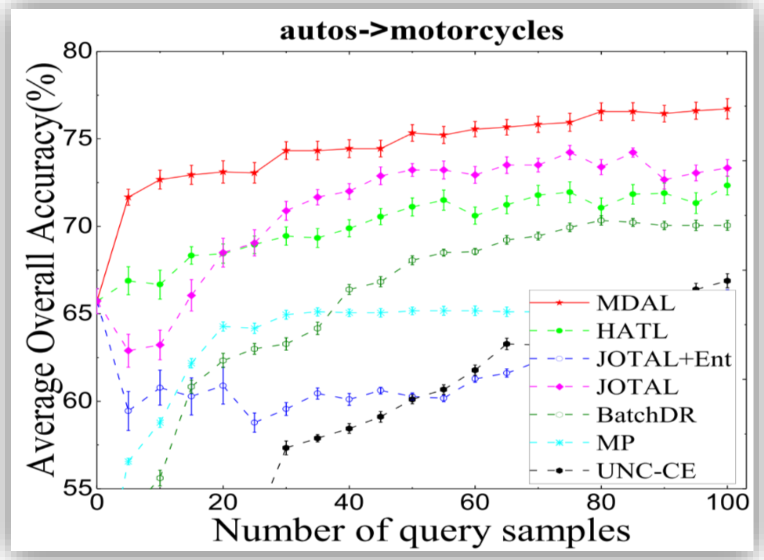
Lin et al. Active Self-Paced Learning for Cost-Effective and Progressive Face Identification. PAMI 2018.

Active learning with Weak Supervision

- Active query from source domains
  - Oracle is not available in the target domain
  - Insufficient labeled data in all domains







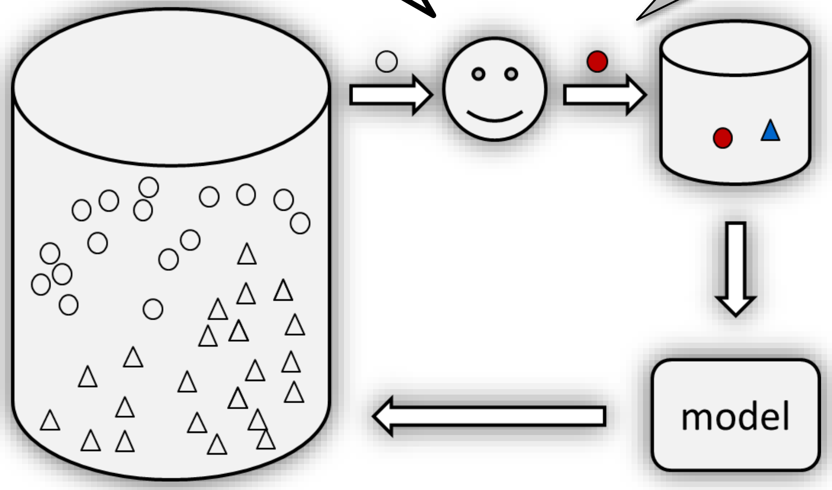
Wang et al. On Gleaning Knowledge from Multiple Domains for Active Learning. IJCAI 2017.

Recent Progress

Weak supervision
The oracle may be noisy or unavailable

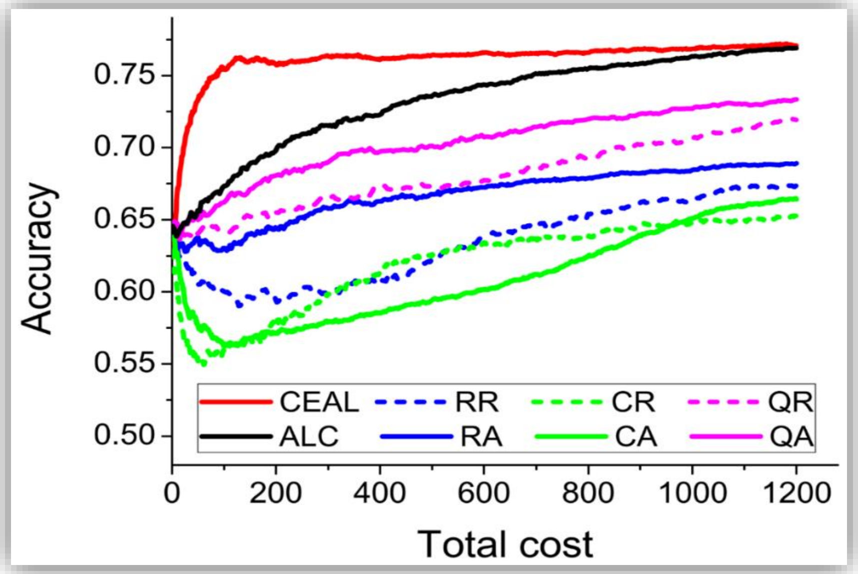
Cost sensitive
Care the cost rather than the number

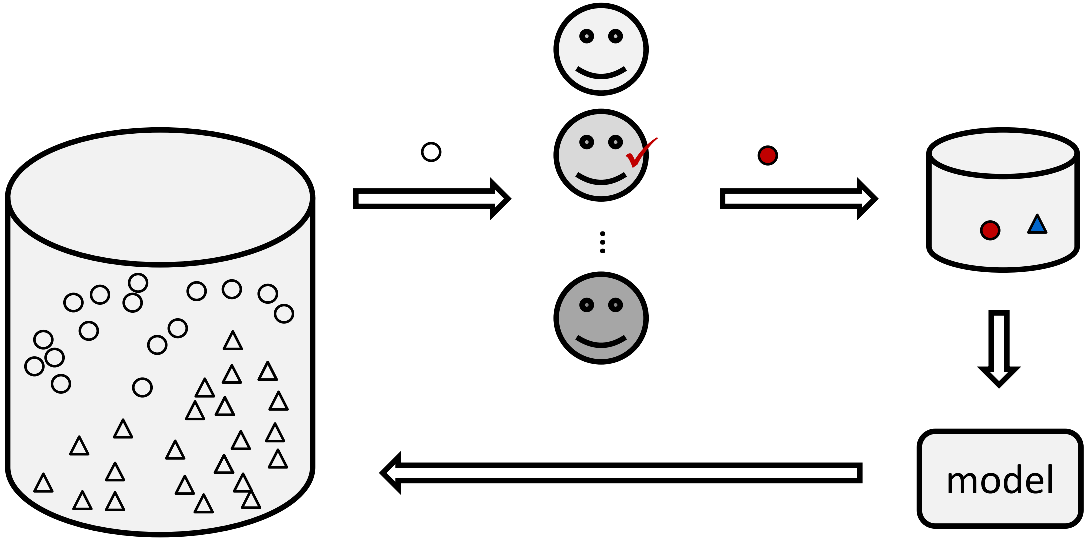
Model dependent
Different models may have diverse needs



Cost-Sensitive Active Learning

- Oracles are cost-sensitive
  - Different oracles have diverse prices
  - Selecting both instance and oracle
  - Accurate yet cheap annotations





Who is this ?

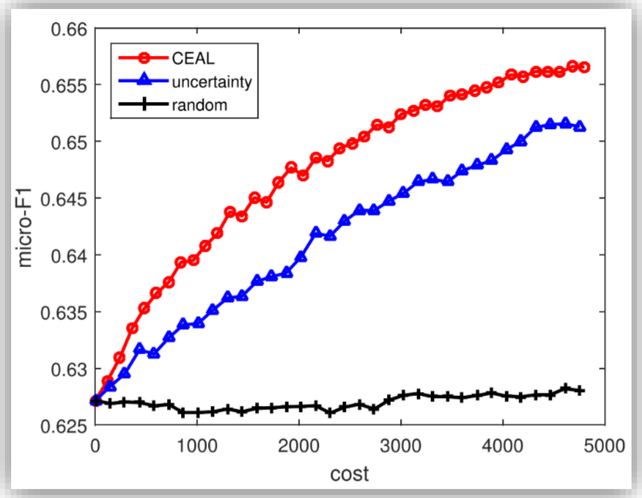
- Low overall quality
- Low price
- Expert for this query

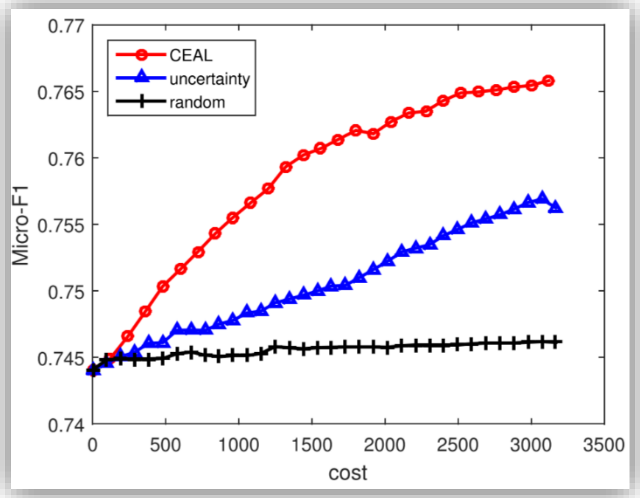
- High overall quality
- High price
- Less familiar with it

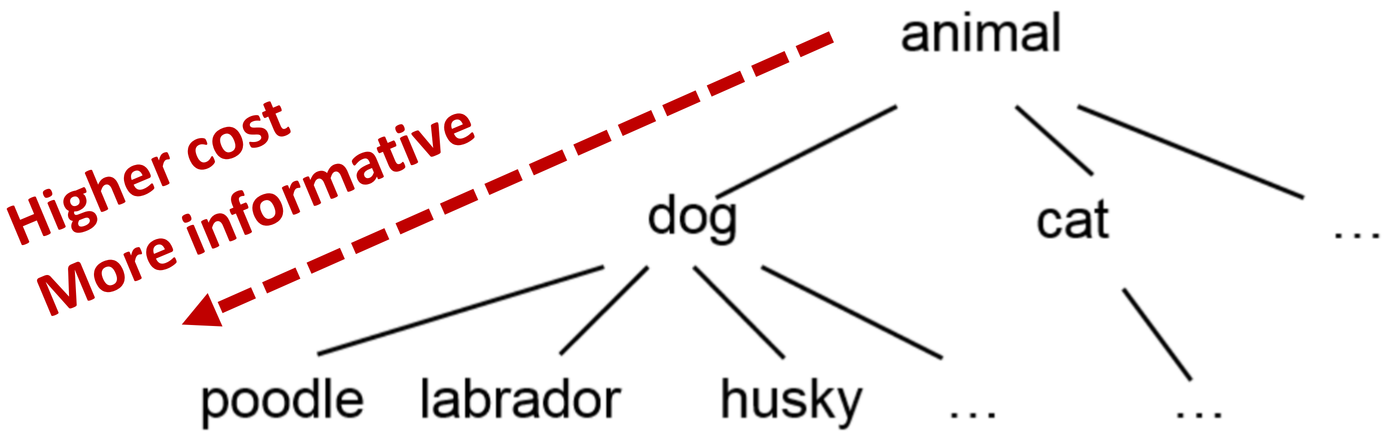
Huang et al. Cost-Effective Active Learning from Diverse Labelers. IJCAI 2017.

Cost-Sensitive Active Learning

- Labels are cost-sensitive
  - Labels have hierarchies
  - Bi-objective optimization to balance the cost and information







Yan et al. Cost-Effective Active Learning for Hierarchical Multi-Label Classification. IJCAI 2018.

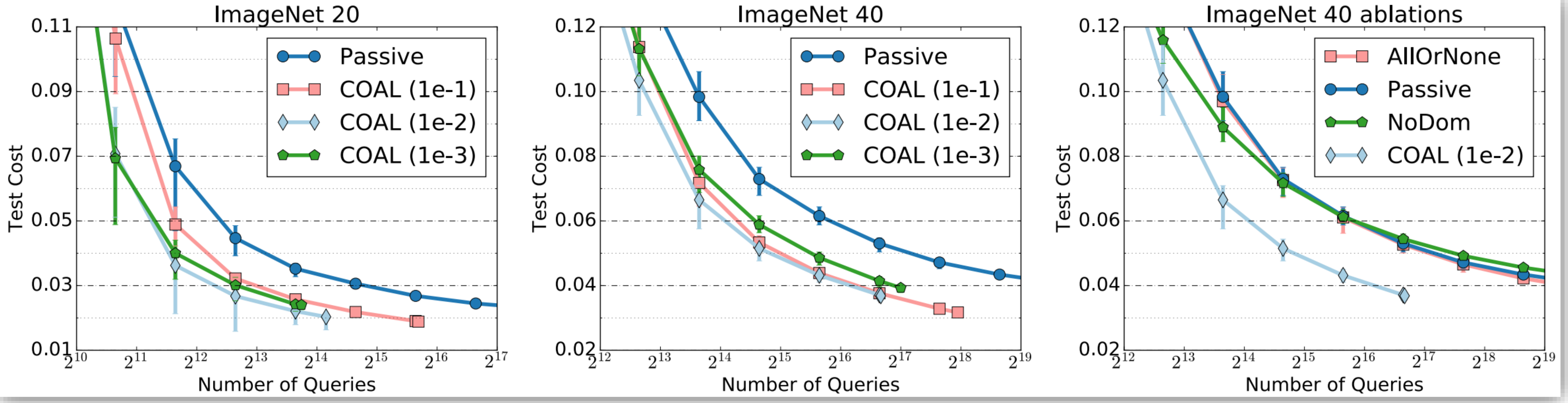
Cost-Sensitive Active Learning

- Learning task is cost-sensitive
  - Query the cost of predicting a specific label
  - Guarantee a polynomial improvement on label complexity for low noise case

Theorem 6. *Assume the Massart noise condition holds. With probability at least $1-2\delta$ the label complexity of the algorithm over n examples is at most,*

$$L_1 = \mathcal{O}\left(\frac{25^{1/\beta}}{\tau^2}\left(n^\beta K \log(n)\nu_n\theta_1 + \log(1/\delta)\right)\right)$$

$$L_2 = \mathcal{O}\left(\frac{25^{1/\beta}K}{\tau^2}\left(n^\beta \log(n)\nu_n\left[K\theta_1 + \theta_2\right] + \log(1/\delta)\right)\right)$$

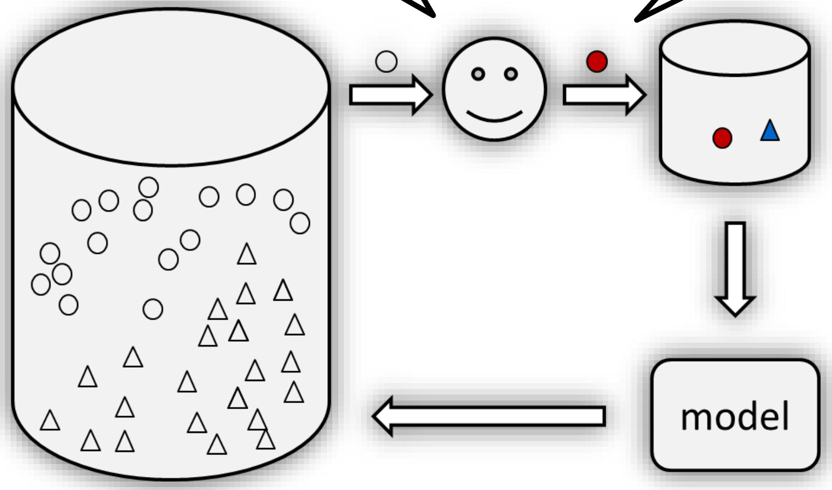


Krishnamurthy et al. Active Learning for Cost-Sensitive Classification. ICML 2017.

Recent Progress

Weak supervision
The oracle may be noisy or unavailable

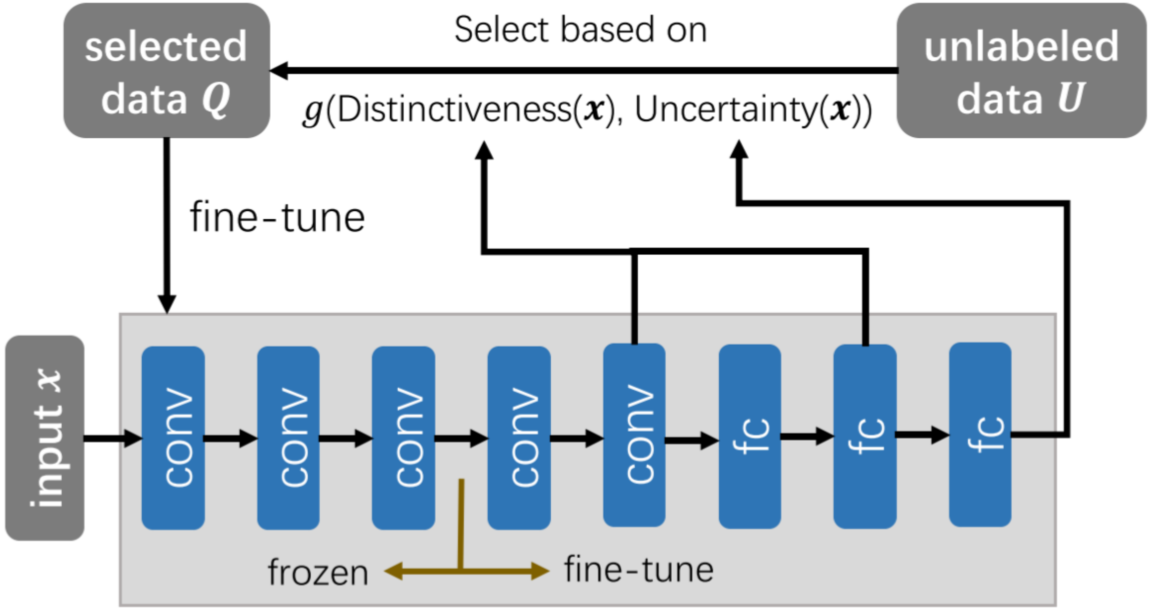
Cost sensitive
Care the cost rather than the number

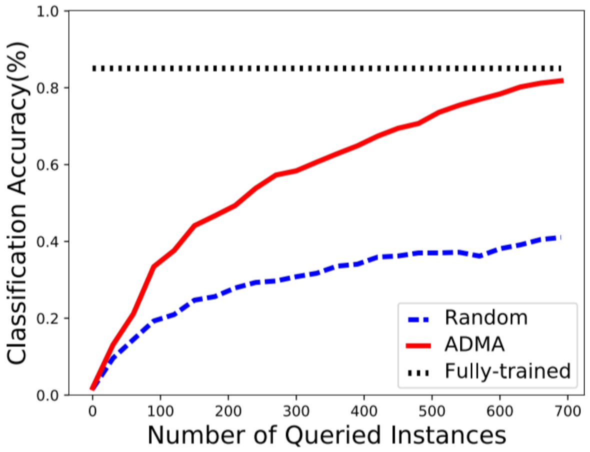


Model dependent
Different models may have diverse needs

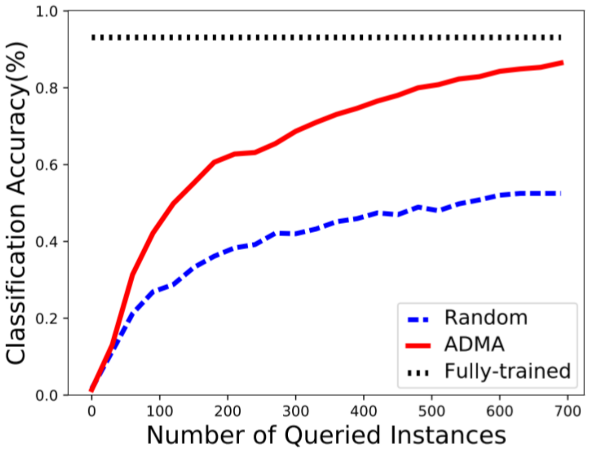
Active Learning with Deep Models

- Active model adaptation
  - A novel criterion "distinctiveness"
  - Reuse of pre-trained models
  - Less training data

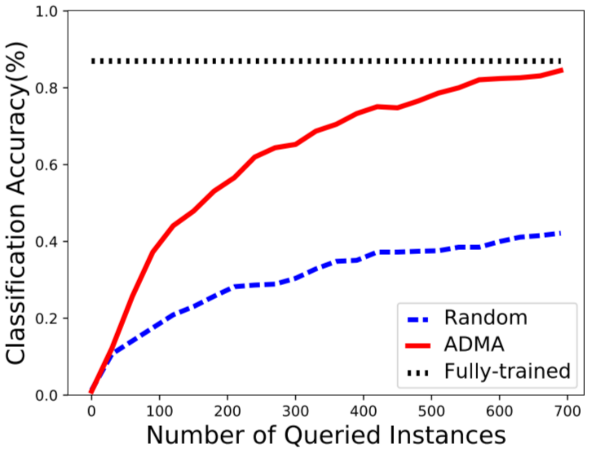




(d) AlexNet + Indoor



(e) VGG-16 + Indoor

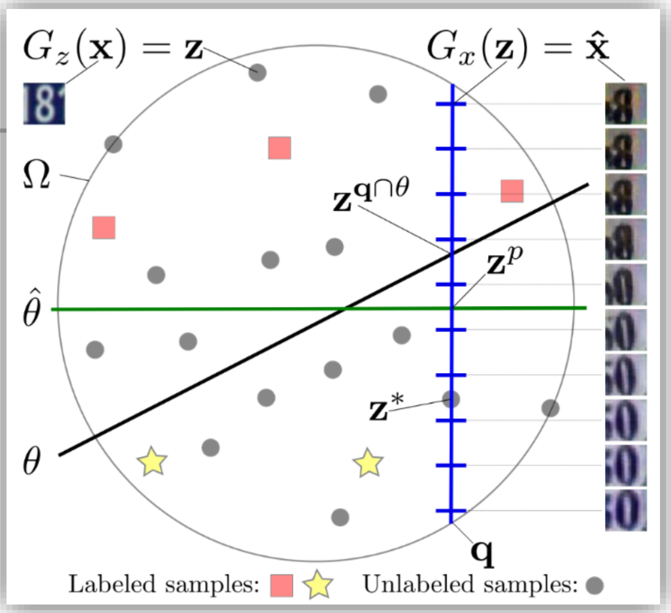


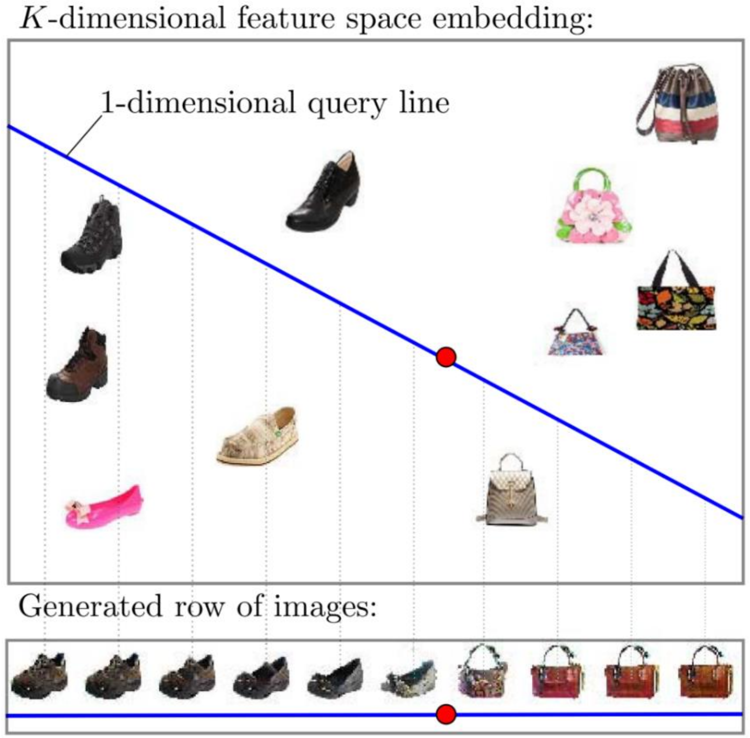
(f) ResNet-18 + Indoor

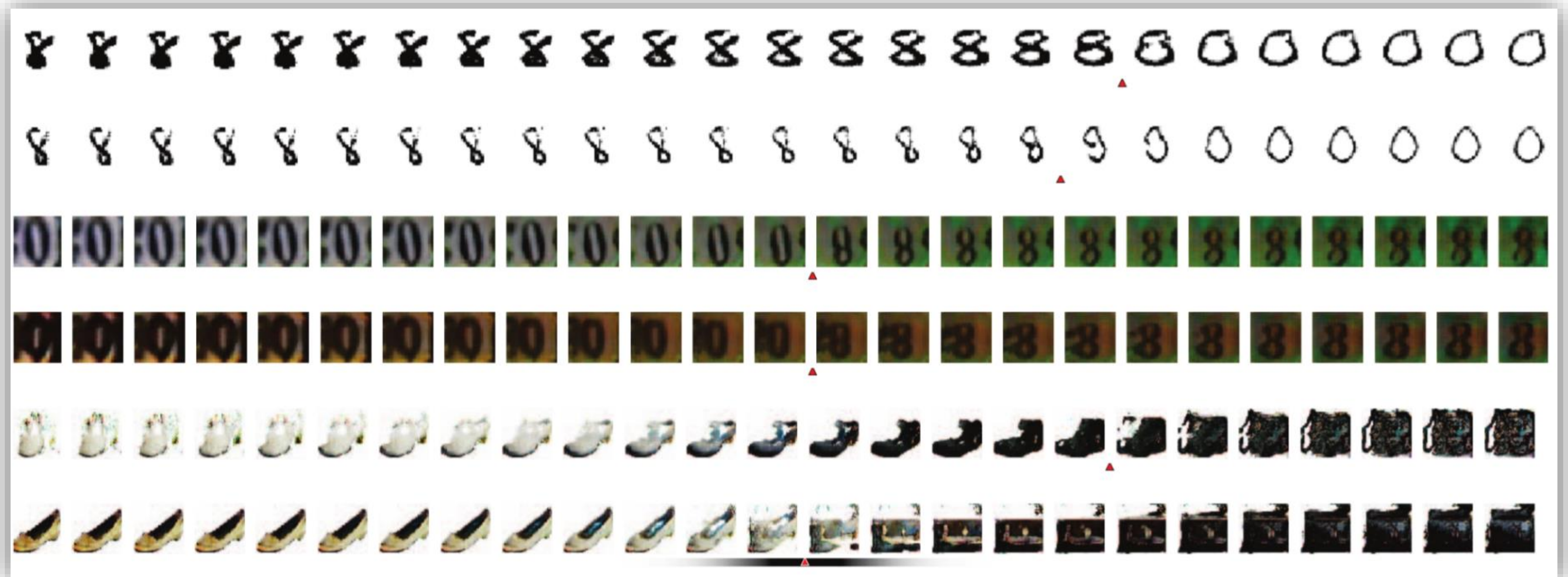
Huang et al. Cost-Effective Training of Deep CNNs with Active Model Adaptatio. arXiv 2018.

Active Learning with Deep Models

- Active annotation with deep generative models
  - Deep generative model to create novel instances
  - Oracle directly annotates the decision boundary







Huijser et al. Active Decision Boundary Annotation with Deep Generative Models. ICCV 2017.

Active Learning for Various Applications

- Human Pose Estimation [Liu & Ferrari ICCV'17]
- Face Identification [Lin et al. PAMI'18]
- Semantic Role Labeling [Wang et al. IJCAI'17]
- Biomedical Image Analysis [Zhou et al. CVPR'17]
- Quadcopter Control [Andersson et al. AAAI'17]
- Sentiment Analysis [Wu et al. ACL'17]
- Recommendation [Zhao et al. AIJ'17]
- Surveying [Lewenberg et al. AAAI'17]
- ……

Thank You !

huangsj@nuaa.edu.cn
http://parnec.nuaa.edu.cn/huangsj