Interactive Medical Word Sense Disambiguation with Instance and *Feature* Labeling

S16: Sub-language and Multi-lingual NLP

Yue Wang

PhD candidate

Computer Science and Engineering

University of Michigan

Disclosure

I and my spouse/partner have no relevant relationships with commercial interests to disclose.

Funding support

- NLM 2R01LM010681-05
- NSF IIS-1054199

Word sense disambiguation (WSD)

Words have multiple meanings (senses)

- cold: (1) low temperature; (2) common cold; (3) chronic obstructive lung disease (COLD)

Given ambiguous word in **context**, automatically assign a **sense** from a given set

context

"The child caught a ***cold*** and kept sneezing."

computer algorithm

sense

- low temperature
- common cold
- chronic obstructive lung disease (COLD)

Critical step for many medical NLP applications

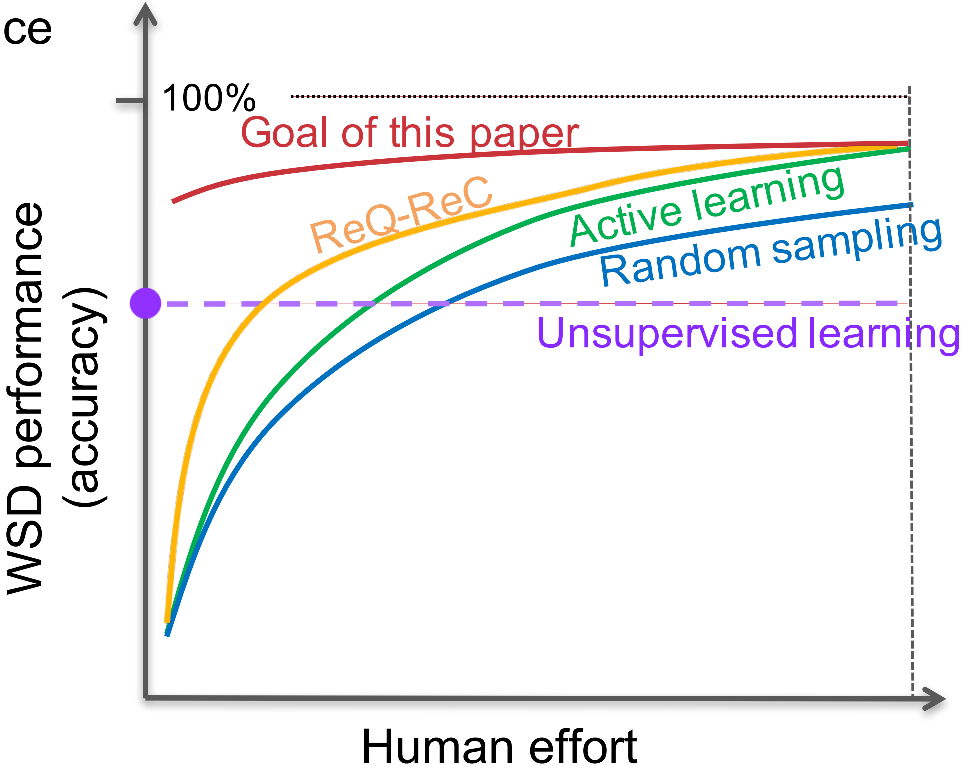
- Document indexing & classification; named entity extraction; computer-assisted review

WSD solutions

Human effort vs. algorithm performance
- Learning curve

Machine learning algorithms
- Unsupervised learning
- Supervised learning
  - labeling instances

Can we make better use of expert's domain knowledge?



Beyond instance labels

"cold"
- Chronic Obstructive Lung Disease
- Common Cold
- Low Temperature

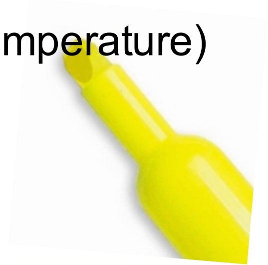
Indicative words:
- Chronic Obstructive Lung Disease: "COLD", "chronic", "obstructive", "lung", …
- Common Cold: "common", "cough", "sneeze", …

Unified
Medical
Language
System

"I know it when I see it":
- Rationale behind a sense label

"The tissue was exposed to a ***cold*** environment (5 degrees C)."

(Low Temperature)



Capture WSD domain knowledge

INFORMATICS PROFESSIONALS. LEADING THE WAY.

(1) Type in indicative words

Chronic Obstructive Lung Disease

COLD, chronic, obstructive, lung

Common Cold

common, cough, sneeze

(2) Highlight contextual cues when labeling

The tissue was exposed to a cold environment (5 degrees C).

- Chronic Obstructive Lung Disease
- Common Cold
- Low Temperature

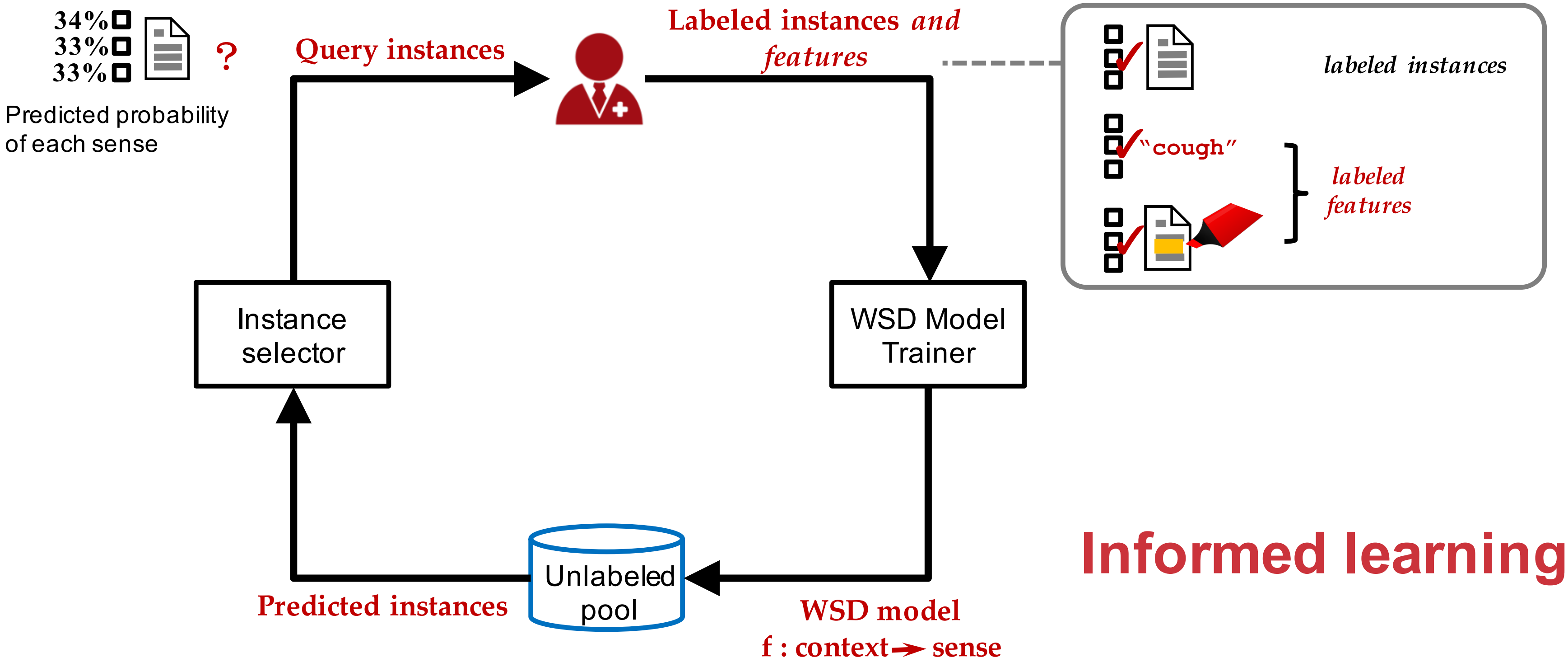
labeled instance
(without highlighting)

[textual pattern, sense] pairs

labeled features

["COLD" (all-cap),	Chronic Obstructive Lung Disease]
["chronic",	Chronic Obstructive Lung Disease]
["obstructive",	Chronic Obstructive Lung Disease]
["common",	Common Cold]
["cough",	Common Cold]
["sneeze",	Common Cold]
["<digit> degrees C",	Low Temperature]

Workflow



Informed learning

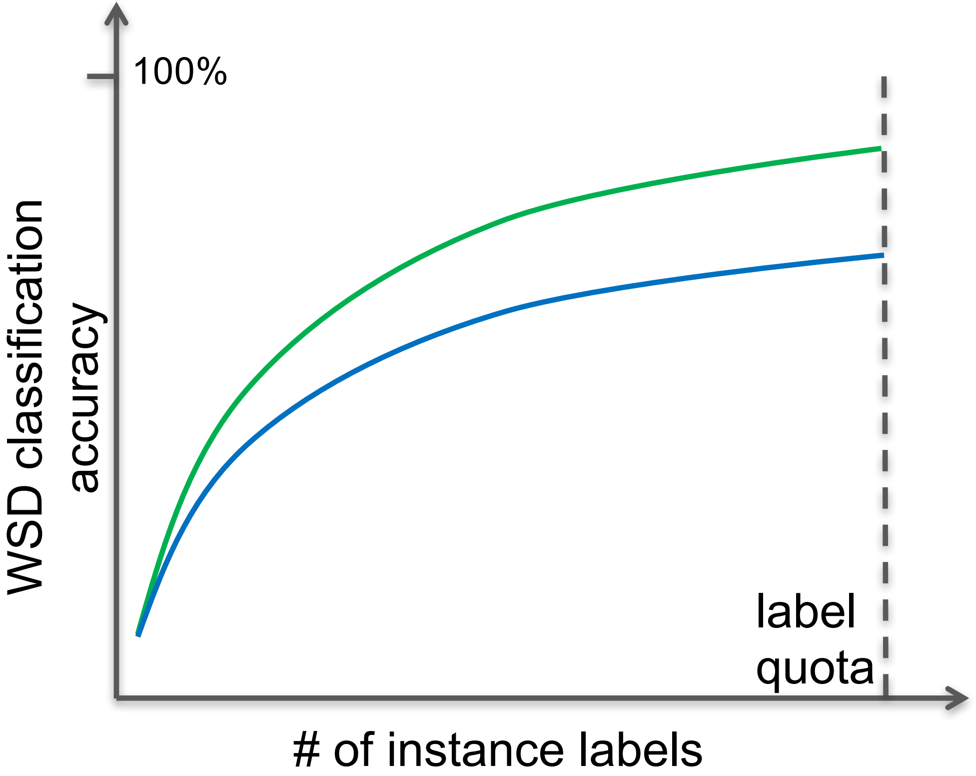
Experimental setup

Evaluation metric

- Learning curves
  - Labeled features only used at the first iteration
- Area under learning curve (ALC)

Compared methods

- Random sampling
- Active learning
- ReQ-ReC
- **Informed learning**



Evaluation Corpora

MSH: ambiguous words and abbreviations in MEDLINE abstracts

UMN: ambiguous clinical abbreviations, Univ. of Minnesota Fairview Hospital

VUH: ambiguous clinical abbreviations, Vanderbilt Univ. Hospital

	#ambiguous words	#sense/ word	#contexts/ word	#tokens/ context	majority guess accuracy
MSH	198	2.1	190	203	54.0%
UMN	75	5.4	500	61	73.8%
VUH	24	4.3	194	19	78.3%

Simulated human expert inputs

Provide labeled instances in active learning

- Sense labels in the corpora

Provide labeled features per sense (only at the first iteration)

- Words with top information gain

Chronic Obstructive Lung Disease	**Common Cold**	**Low Temperature**
obstructive	common	degrees
chronic	cough	C
lung	colds	temperatures
COLD	nasal	temperature
pulmonary	over-the-counter	exposure
COPD	OTC	gene

Simulated human expert inputs

Provide labeled instances in active learning

- Sense labels in the corpora

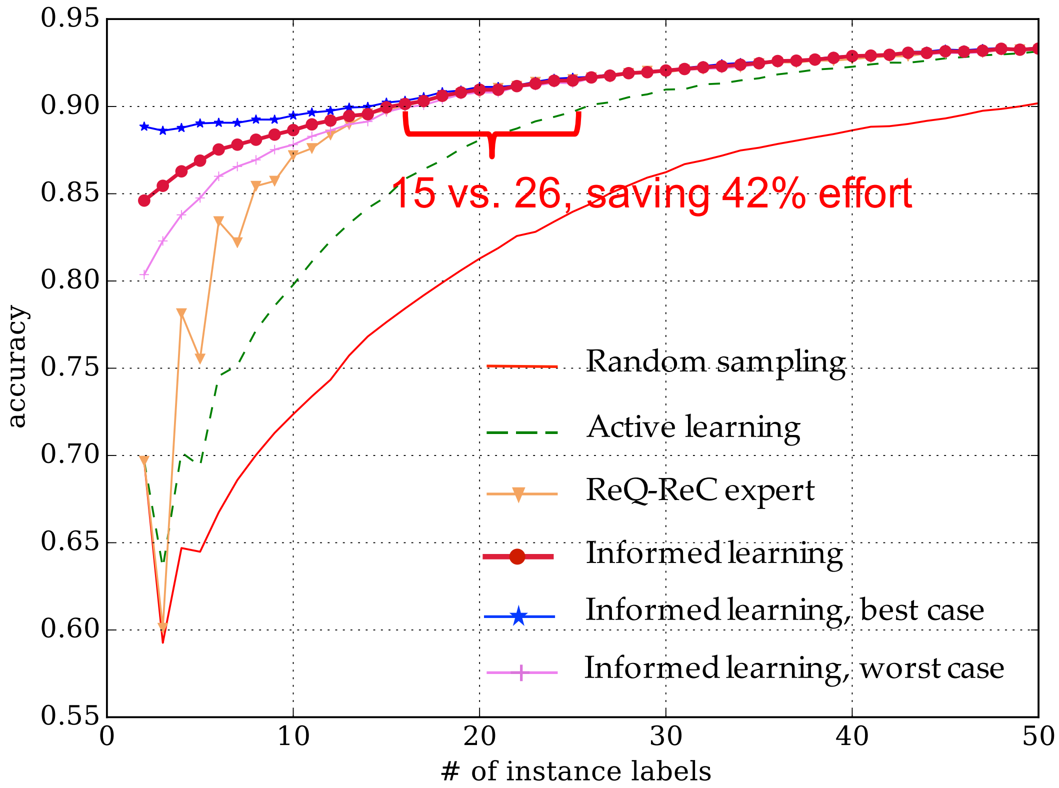
Provide labeled features per sense (only at the first iteration)

- Words with top information gain
- Making it more realistic:
  - Provide the best, 5th best, 10th best contextual words *upfront*
  - Highlight the best, 2nd best, 3rd best contextual words *in a given instance*

"Best-case scenario"

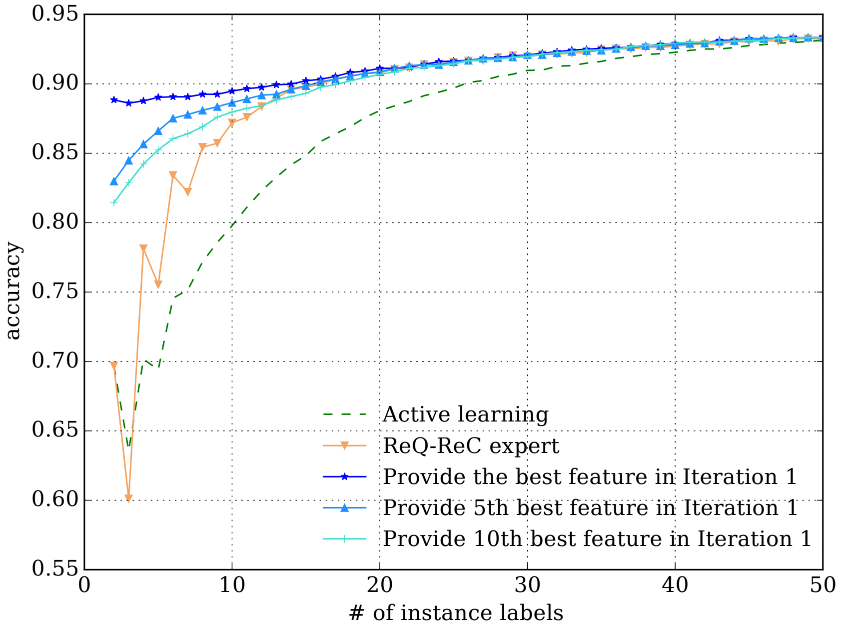
"Worst-case scenario"

MSH: 198 ambiguous words in MEDLINE

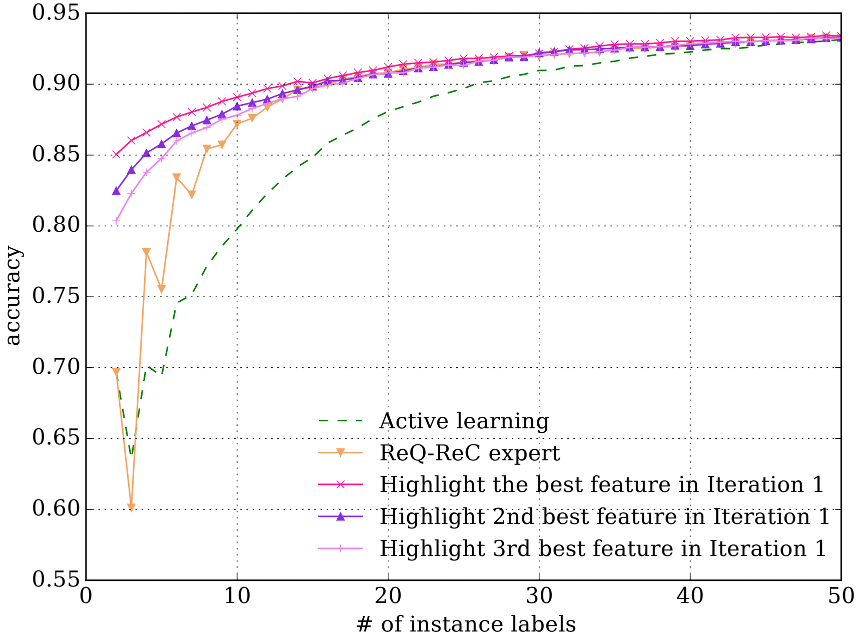


MSH drill-down analysis

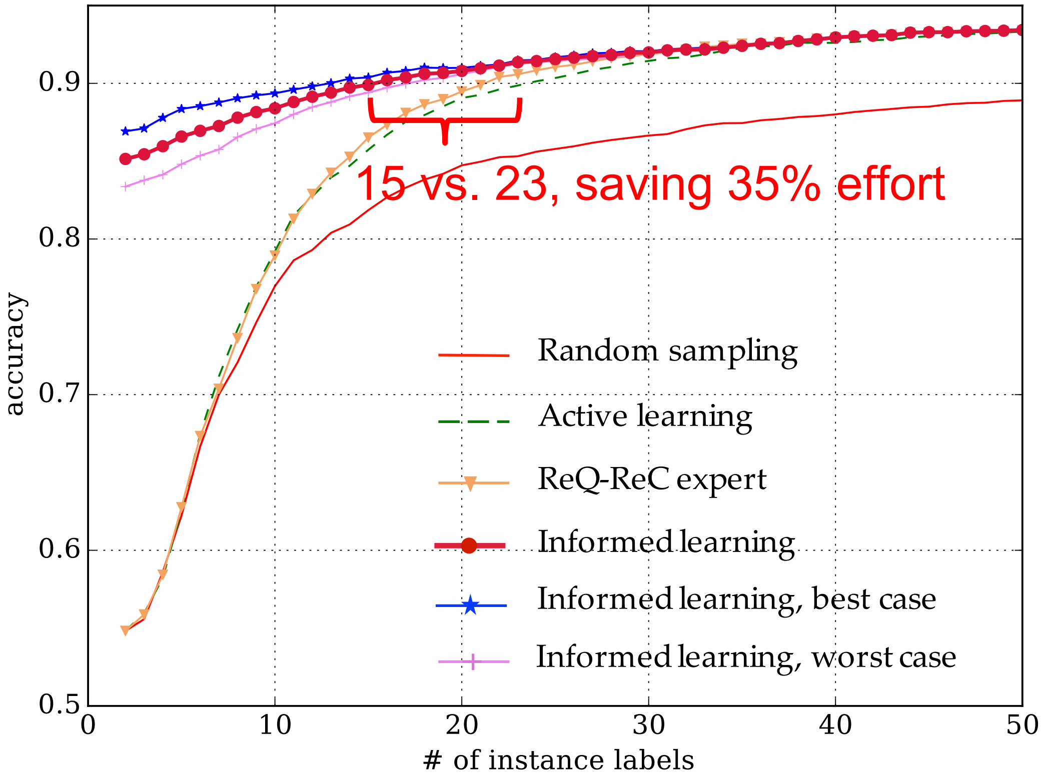
provide features upfront



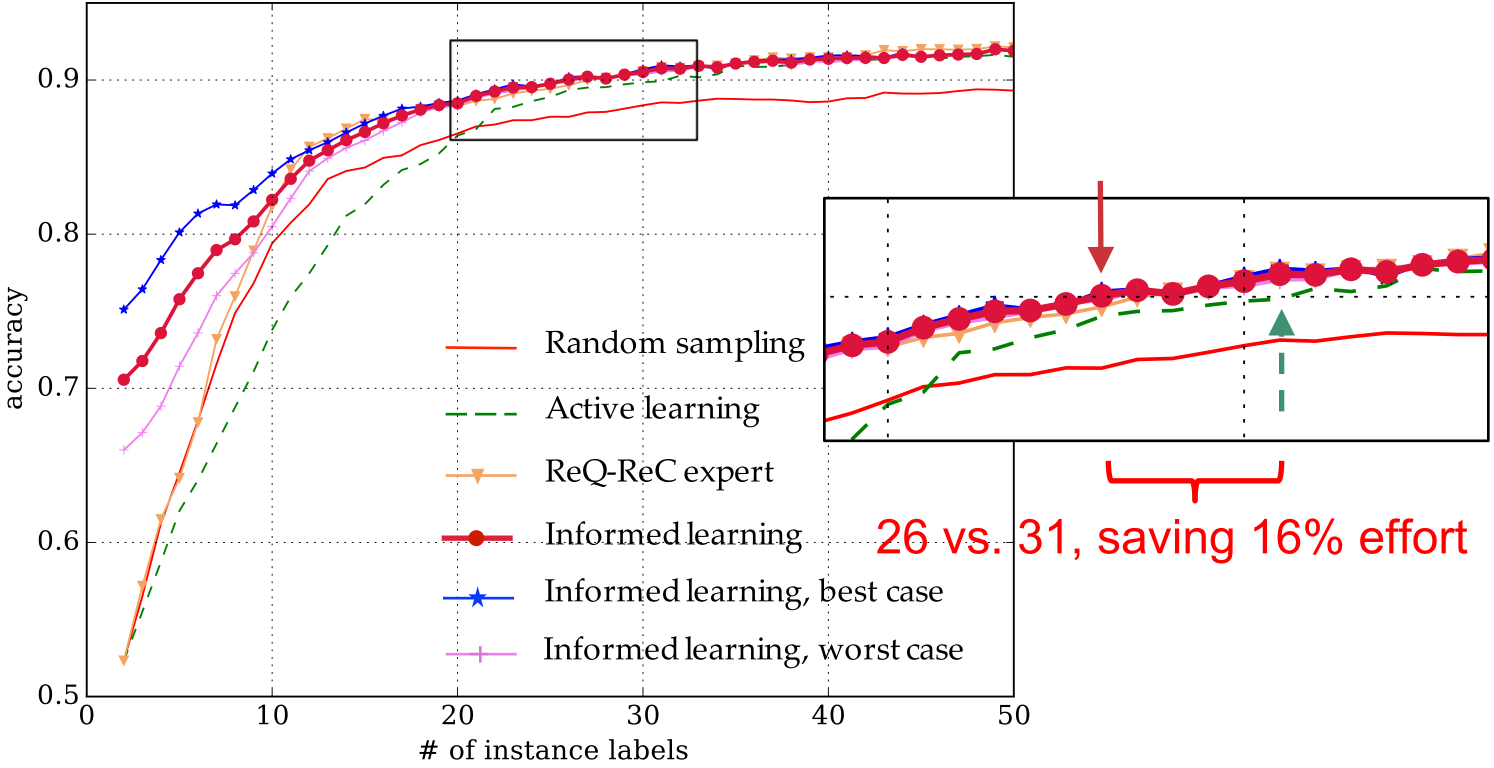
highlight features in examples



UMN: 75 ambiguous clinical abbreviations



VUH: 24 ambiguous clinical abbreviations



Area under learning curve (ALC)

	MSH	UMN	VUH
Random	0.816	0.817	0.831
Active learning	0.868	0.854	0.831
ReQ-ReC expert	0.893	0.857	0.852
Informed learning	**0.909***†	**0.909***†	**0.871***

Wilcoxon signed rank test
'$*$' : significant w.r.t. "Active learning" at $\alpha = 0.01$.
'$\dagger$' : significant w.r.t. "ReC-ReQ expert" at $\alpha = 0.01$.

Conclusion

Learning from labeled features is more efficient use of expert knowledge and time

- Labeling instances → providing/highlighting features
- Less human effort to reach desirable accuracy
- Even noisy feature labels are useful

Future work

- Inviting human experts in real-time evaluation

Acknowledgement

The
Co-authors

Kai Zheng

PhD, University of California, Irvine

Hua Xu

PhD, University of Texas Health Science Center

Qiaozhu Mei

PhD, University of Michigan

And **Anonymous Reviewers**

THANKS

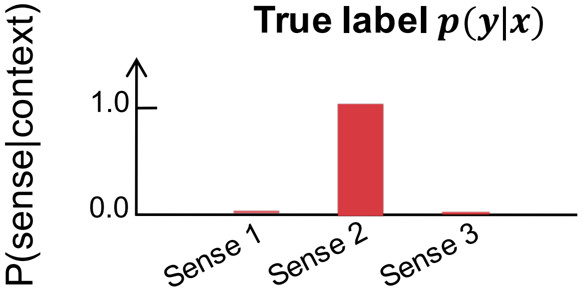
raywang@umich.edu

Training WSD model with labeled instances & features

INFORMATICS PROFESSIONALS. LEADING THE WAY.

Labeled instance: [context x, sense y]

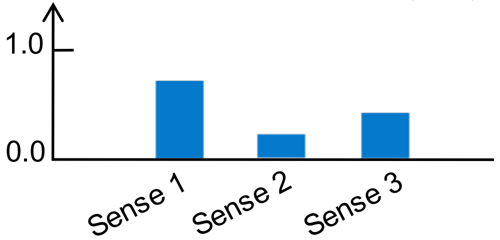
True label $p(y|x)$



cross entropy loss term

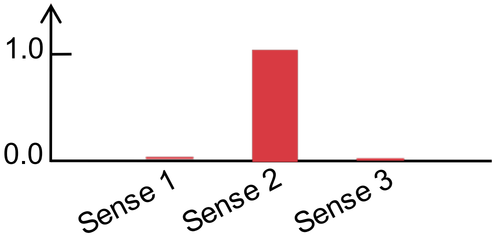
$$\sum_y p(y|x) \log \hat{p}(y|x)$$

Predicted label $\hat{p}(y|x)$



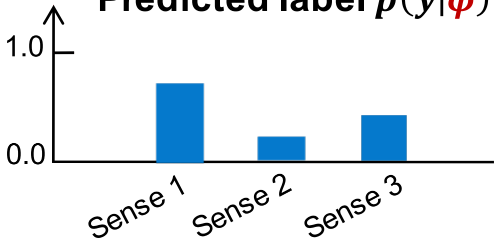
Labeled feature: [feature ϕ, sense y]

True label $p(y|\phi)$

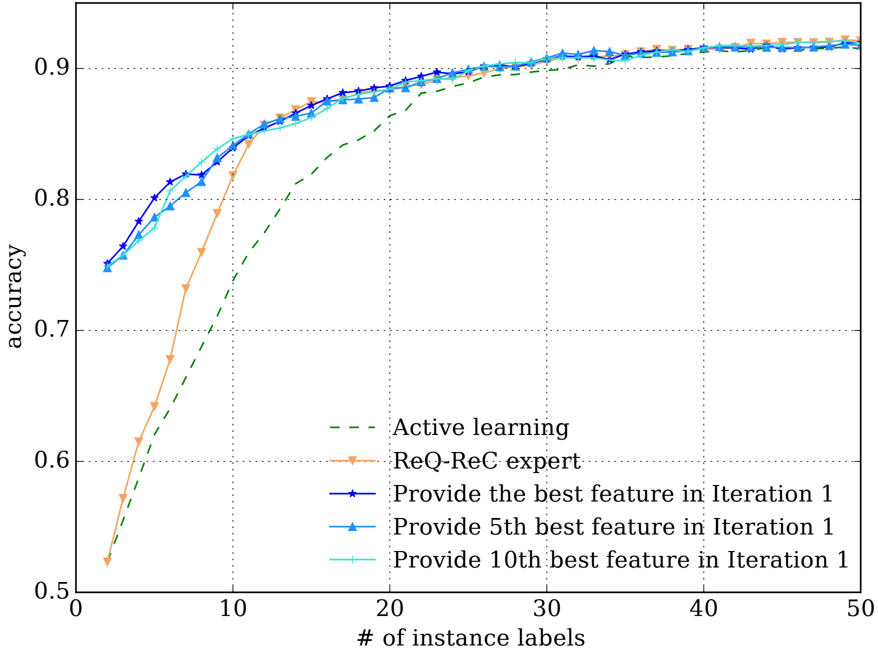


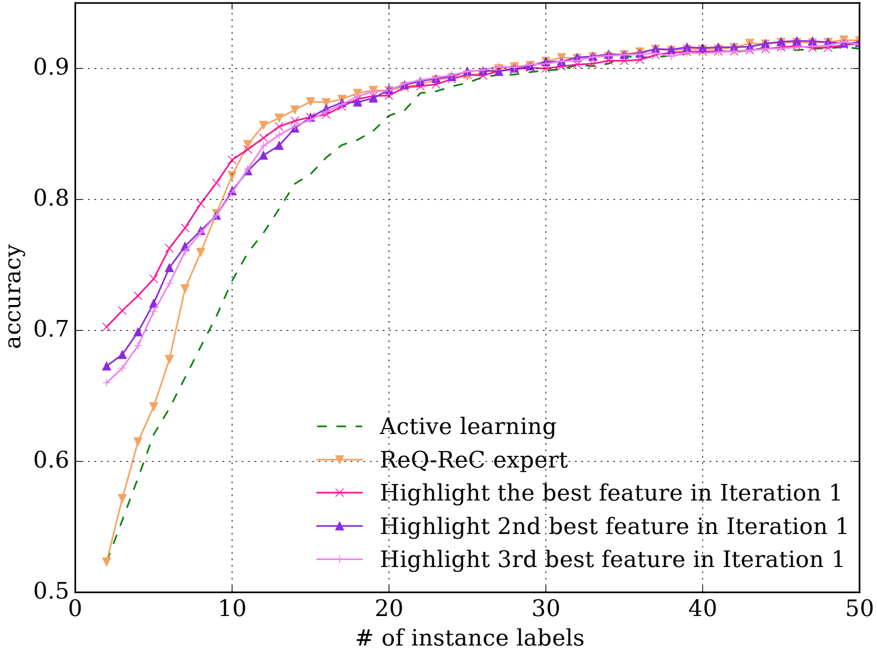
a new cross entropy loss term

Predicted label $\hat{p}(y|\phi)$ $= \dfrac{\sum_{x:\phi\in x} \hat{p}(y|x)}{|\{x:\phi\in x\}|}$



VUH drill-down analysis





UMN drill-down analysis