Active Multitask Learning Using Both Supervised and Latent Shared Topics

Ayan Acharya, Raymond J. Mooney, Joydeep Ghosh

UT Austin, Dept. of ECE & CS

April 24, 2014

Outline

- Background
- Act-DSLDA and Act-NPDSLDA
- Datasets & Empirical Results
- References

Motivation

- Multitask Learning: data from multiple tasks are collected and models are learnt simultaneously
- Active Learning: only the most informative examples are queried from the unlabeled pool
- Unify both of these approaches

Problem Setting

- In training corpus each document/image belongs to a known class and has a set of attributes (supervised topics).
- Classes from aYahoo data: carriage, centaur, bag, building, donkey, goat, jetski, monkey, mug, statue, wolf, and zebra
- Attributes: "has head", "has wheel", "has torso" and 61 others
- Train models using words, supervised topics and class labels
- An active MTL framework that can use and query over both attributes and class labels
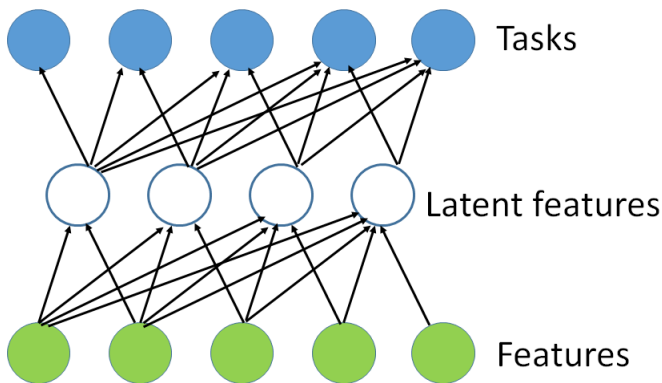

Class: Carriage

Attributes:
"has wheel?" Yes.
"has wood?" Yes.
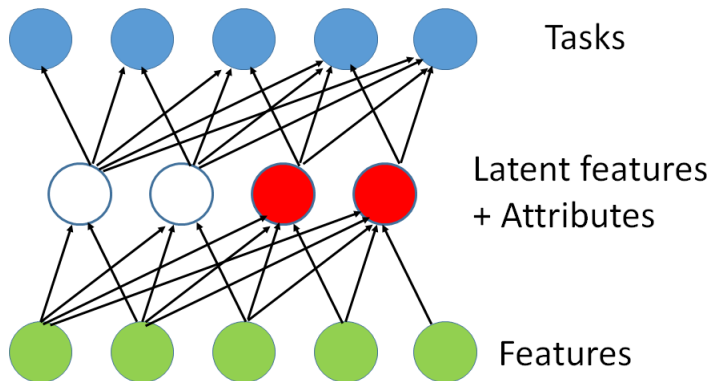
Transfer with Shared Supervised Attributes
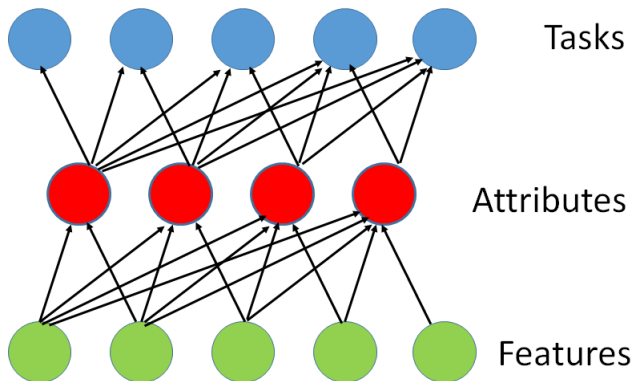
- Train to infer attributes from visual features
- Train to infer categories from attributes [Lampert et al., 2009]

Multitask Learning with Shared Latent Features

Tasks

Latent features

Features

Reference: [Caruana, 1997]

Background
○○○○●○○○○

Act-DSLDA & Act-NPDSLDA
○○○○

Datasets & Empirical Results
○○○○○○○○○○○○

References
○○○

Transfer with Shared Supervised and Latent Attributes

Tasks

Latent features
+ Attributes

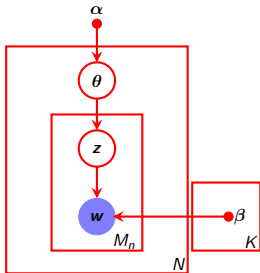Features

Topic Models: LDA
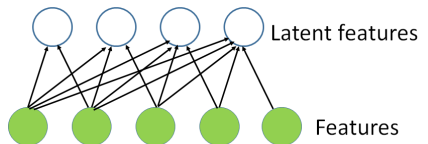

Figure : LDA

Figure : Visual Representation

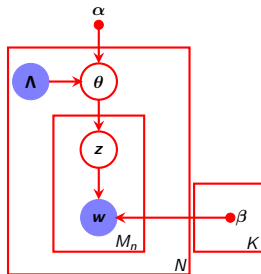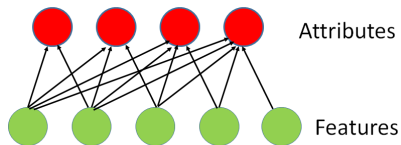## Topic Models: LLDA

Figure : LLDA
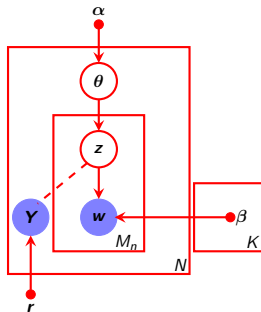

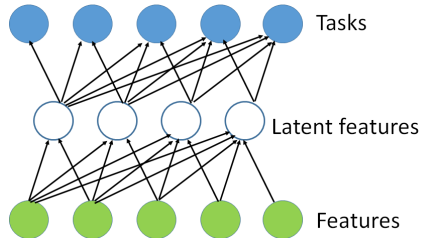Figure : Visual Representation

Topic Models: MedLDA

Figure : MedLDA

Figure : Visual Representation

Background
○○○○○○○○○●

Act-DSLDA & Act-NPDSLDA
○○○○

Datasets & Empirical Results
○○○○○○○○○○○○○

References
○○○

Topic Models: DSLDA

- Doubly Supervised LDA [Acharya et al., 2013]
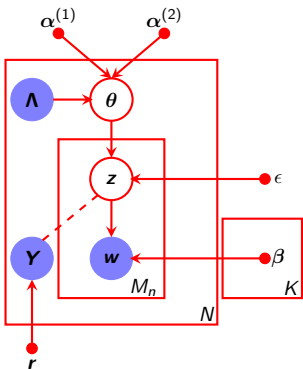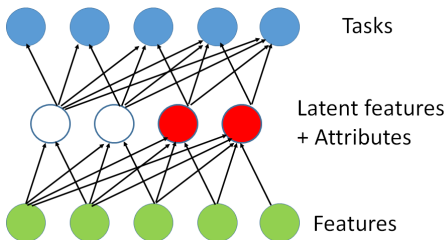- $\alpha^{(1)}, \alpha^{(2)}$: priors over supervised and latent topics

Figure : DSLDA

Figure : Visual Representation

Active DSLDA (Act-DSLDA)

- r_1 : weights for multiclass SVM
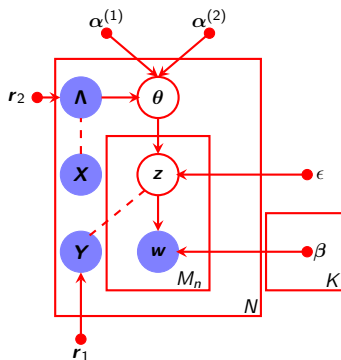- r_2 : weights for binary SVMs

Figure : Act-DSLDA
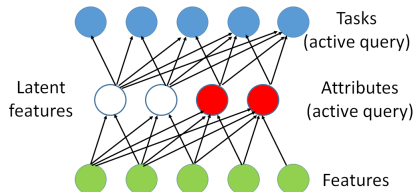

Figure : Visual Representation

Active NPDSLDA (Act-NPDSLDA)

- Non-parametric Doubly Supervised LDA [Acharya et al., 2013]
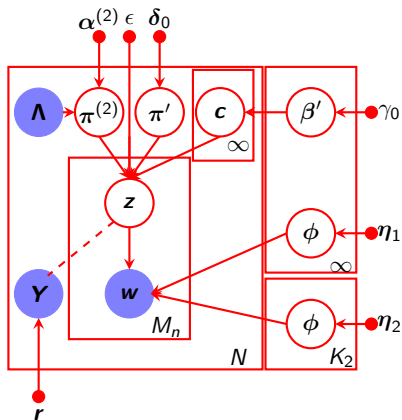

Figure : NPDSLDA

Active NPDSLDA (Act-NPDSLDA)

- Non-parametric Doubly Supervised LDA [Acharya et al., 2013]

Figure : NPDSLDA

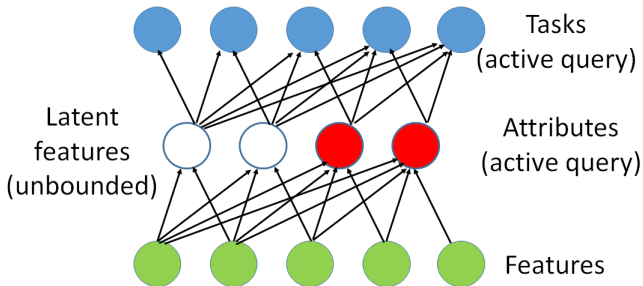Figure : Act-NPDSLDA

Visual Representation of Act-NPDSLDA

Figure : Visual Representation of Act-NPDSLDA

Inference and Learning

- Active learning measure: expected error reduction [Nigam et al., 1998]
- Batch mode: variational EM with completely factorized approximation to posterior, online SVM [Bordes et al., 2007]
- Active selection mode: incremental EM [Neal and Hinton, 1999], online SVM

Description of Dataset: ACM Conference

- **Classes:** Conference names: WWW, SIGIR, KDD, ICML, ISPD, DAC; abstracts of papers are treated as documents
- **Supervised topics:** keywords provided by the authors

Experimental Methodology

- Multitask training that evaluates benefits of sharing information among classes on the predictive accuracy of all classes
- Start with a completely labeled dataset $\mathcal{L}$ consisting of 300 documents
- In every active iteration, 50 labels (class labels or supervised topics) are queried for.

Compared Models

Model	Supervised Topics	Latent Topics	Class Labels
Act-DSLDA	present & queried	shared	queried
Act-NPDSLDA	present & queried	shared	queried
R-MedLDA-MTL	absent	shared	random selection
R-DSLDA	present & random selection	shared & random selection	random selection
Act-MedLDA-OVA	absent	not shared	queried
Act-MedLDA-MTL	absent	shared	queried
Act-DSLDA-OSST	present & queried	absent	queried
Act-DSLDA-NSLT	present & queried	not shared	queried

1. **R**andom MedLDA-MTL (R-MedLDA-MTL)
2. **R**andom DSLDA (R-DSLDA)
3. **A**ctive Learning in MedLDA with **o**ne-**v**s-**a**ll classification (Act-MedLDA-OVA)
4. **A**ctive Learning in MedLDA with **m**ulti**t**ask **l**earning (Act-MedLDA-MTL)
5. Act-DSLDA with **o**nly **s**hared **s**upervised **t**opics (Act-DSLDA-OSST)
6. Act-DSLDA with **n**o **s**hared **l**atent **t**opics (Act-DSLDA-NSLT)
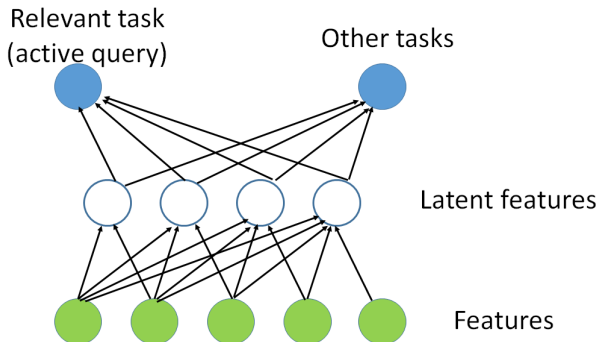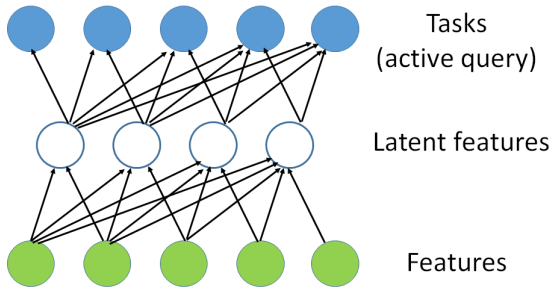
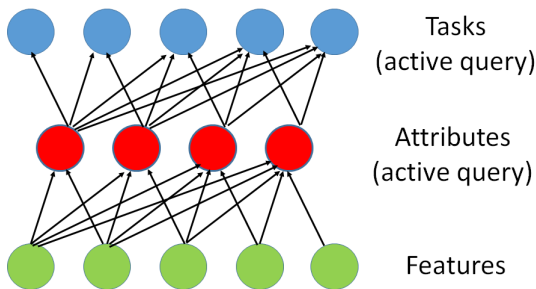# Random MedLDA-MTL (R-MedLDA-MTL)
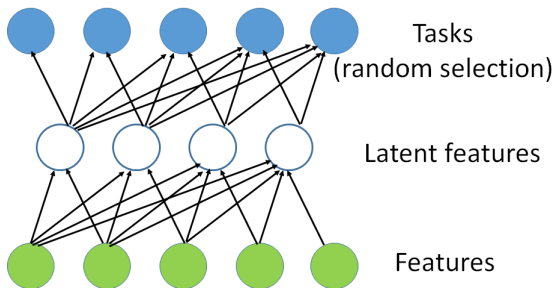
Random DSLDA (R-DSLDA)
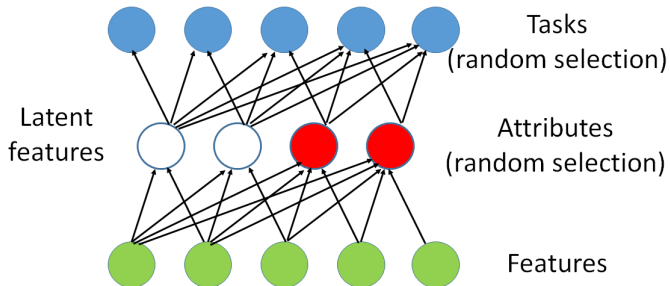
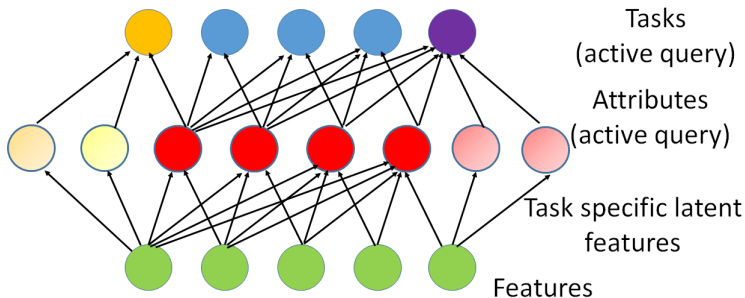# Active Learning in MedLDA with one-vs-all classification (Act-MedLDA-OVA)

Active Learning in MedLDA with Multitask Learning (Act-MedLDA-MTL)

Background
○○○○○○○○○○

Act-DSLDA & Act-NPDSLDA
○○○○

Datasets & Empirical Results
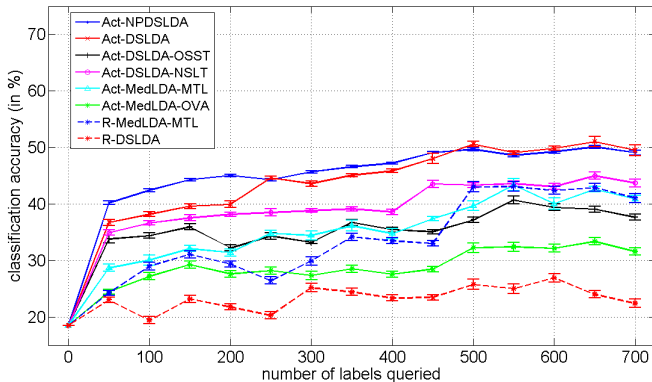○○○○○○○●○○○○○

References
○○○

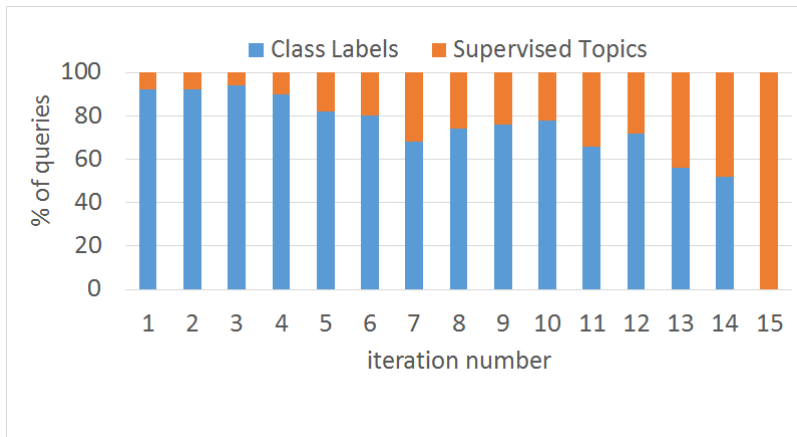# Act-DSLDA with Only Shared Supervised Topics (Act-DSLDA-OSST)

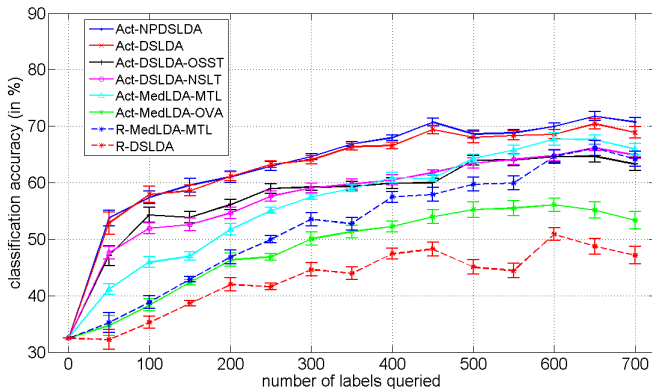# Act-DSLDA with No Shared Latent Topics (Act-DSLDA-NSLT)

aYahoo Learning Curves

aYahoo Query Distribution
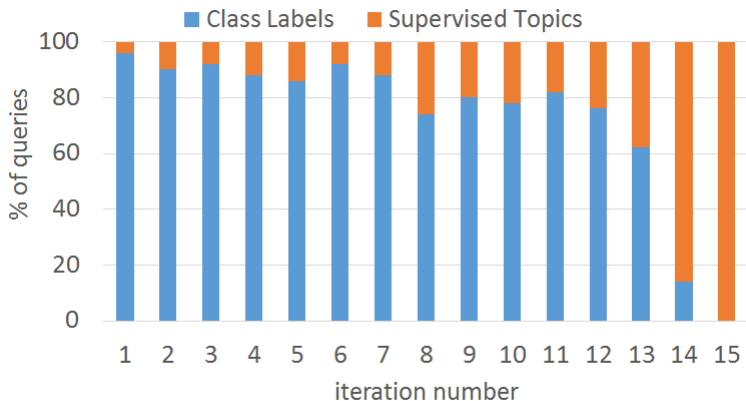
ACM Conference Learning Curves

ACM Conference Query Distribution

Conclusion and Future Work

- Experimental results demonstrate the utility of integrating active and multitask learning in one framework that also unifies latent and supervised shared topics.
- Better approximation techniques for active selection with large scale learning
- Active query with annotators' rationales

References

Acharya, A., Rawal, A., Mooney, R. J., and Hruschka, E. R. (2013).
Using both supervised and latent shared topics for multitask learning.
In *ECML PKDD, Part II, LNAI 8189*, pages 369–384.

Bordes, A., Bottou, L., Gallinari, P., and Weston, J. (2007).
Solving multiclass support vector machines with larank.
In *Proc. of ICML*, pages 89–96.

Caruana, R. (1997).
Multitask learning.
Machine Learning, 28:41–75.

Lampert, C. H., Nickisch, H., and Harmeling, S. (2009).
Learning to detect unseen object classes by betweenclass attribute transfer.
In *Proc. of CVPR*, pages 951–958.

Neal, R. M. and Hinton, G. E. (1999).
A view of the EM algorithm that justifies incremental, sparse, and other variants.

Nigam, K., McCallum, A., Thrun, S., and Mitchell, T. (1998).
Learning to classify text from labeled and unlabeled documents.
In *Proceedings of the Fifteenth National Conference on Artificial Intelligence*,
pages 792–799. AAAI Press.

Background
○○○○○○○○○

Act-DSLDA & Act-NPDSLDA
○○○○

Datasets & Empirical Results
○○○○○○○○○○○○

References
○○●

Questions?