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

Reference: [Caruana, 1997]
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
Topic Models: DSLDA

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

![Diagram of DSLDA](image.png)

**Figure**: DSLDA

**Figure**: Visual Representation
Active DSLDA (Act-DSLDA)

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

![Act-DSLDA Diagram](image)

**Figure**: Act-DSLDA
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
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

<table>
<thead>
<tr>
<th>Model</th>
<th>Supervised Topics</th>
<th>Latent Topics</th>
<th>Class Labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Act-DSLDA</td>
<td>present &amp; queried</td>
<td>shared</td>
<td>queried</td>
</tr>
<tr>
<td>Act-NPDSLDA</td>
<td>present &amp; queried</td>
<td>shared</td>
<td>queried</td>
</tr>
<tr>
<td>R-MedLDA-MTL</td>
<td>absent</td>
<td>shared &amp; random selection</td>
<td>random selection</td>
</tr>
<tr>
<td>R-DSLDA</td>
<td>present &amp; random selection</td>
<td>shared &amp; random selection</td>
<td>random selection</td>
</tr>
<tr>
<td>Act-MedLDA-OVA</td>
<td>absent</td>
<td>not shared</td>
<td>queried</td>
</tr>
<tr>
<td>Act-MedLDA-MTL</td>
<td>absent</td>
<td>shared</td>
<td>queried</td>
</tr>
<tr>
<td>Act-DSLDA-OSST</td>
<td>present &amp; queried</td>
<td>absent</td>
<td>queried</td>
</tr>
<tr>
<td>Act-DSLDA-NSLT</td>
<td>present &amp; queried</td>
<td>not shared</td>
<td>queried</td>
</tr>
</tbody>
</table>

1. **Random** MedLDA-MTL (R-MedLDA-MTL)
2. **Random** DSLDA (R-DSLDA)
3. Active Learning in MedLDA with one-vs-all classification (Act-MedLDA-OVA)
4. Active Learning in MedLDA with multitask learning (Act-MedLDA-MTL)
5. Act-DSLDA with only shared supervised topics (Act-DSLDA-OSST)
6. Act-DSLDA with no shared latent topics (Act-DSLDA-NSLT)
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)
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

![Chart showing query distribution across iterations](image-url)
ACM Conference Learning Curves

![Learning Curves Graph]

- Act-NPDSLDA
- Act-DLSDA
- Act-DLSDA-OSST
- Act-DLSDA-NSLT
- Act-MedLDA-MTL
- Act-MedLDA-OVA
- R-MedLDA-MTL
- R-DLSDA

The graph shows the classification accuracy (in %) over the number of labels queried. Each dataset is represented by a different line color and style. The accuracy increases as more labels are queried, with some variations observed across different datasets.
ACM Conference Query Distribution

![Graph showing query distribution with class labels and supervised topics.](image)
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


Questions?