Coached Active Learning for Interactive Video Search

Xiao-Yong Wei, Zhen-Qun Yang

Machine Intelligence Laboratory
College of Computer Science
Sichuan University, China
Coached Active Learning for Interactive Video Search

Xiao-Yong Wei, Zhen-Qun Yang

Machine Intelligence Laboratory
College of Computer Science
Sichuan University, China
Coached Active Learning for Interactive Video Search

Xiao-Yong Wei, Zhen-Qun Yang

Machine Intelligence Laboratory
College of Computer Science
Sichuan University, China
Search “car on road” with Google
Search “car on road” with Google
Interactive Search

Relevance Feedback
Interactive Search

Relevance Feedback

Query Modeling, i.e., to figure out what the searcher wants?
Feature Space with Unlabeled Instances
Space Exploration for Query Modeling

How to explore effectively?

Querying Strategy: which instances to query next (i.e., being presented to the searcher for labeling)?
Active Learning with Uncertainty Sampling

The 1st Round – Training Classifier
Active Learning with Uncertainty Sampling

The 1st Round – Query the Searcher
Active Learning with Uncertainty Sampling

The 2nd Round – Training Classifier
Active Learning with Uncertainty Sampling

The 2nd Round – Query the Searcher
Active Learning with Uncertainty Sampling

The 3rd Round
Active Learning with Uncertainty Sampling

The 4th Round
Active Learning with Uncertainty Sampling

The 5th Round
Active Learning with Uncertainty Sampling
Active Learning with Uncertainty Sampling may not be a “kindred soul” with Video Search
Active Learning with Uncertainty Sampling may not be a “kindred soul” with Video Search

The spirit of Active Learning
(to reduce labeling efforts)

The goal of Query Modeling
(to harvest relevant instances)

Also found by A.G. Hauptmann and et al. [MM’06]
The Dilemma of Exploration and Exploitation

- Exploration: to explore more unknown area and get more about the Query Distribution and to boost future gains
- Exploitation: to harvest more relevant instances and to obtain immediate rewards
The Idea of Coached Active Learning

• **Estimate** the *Query Distribution* after each round to avoid the risk of learning on a completely unknown distribution
The Idea of Coached Active Learning

• Estimate the Query Distribution

*pdf* of the underlying Query Distribution
The Idea of Coached Active Learning

- Estimate the *Query Distribution*
- Query users with instances picked from both dense and uncertain areas of the distribution
- Balance the proportion of the two types of instances
The Idea of Coached Active Learning

Modeling the Query Distribution

Balancing the Exploration and Exploitation
Modeling the Query Distribution
Modeling the Query Distribution

$L^+$: An incomplete sampling of the Query Distribution
Modeling the Query Distribution

• A practical implementation
  – Training: organize training examples into *nonexclusive* semantic groups
  – Testing: check $L^+$ with each group, see whether they are *statistically* from the same distribution using two-sample Hotelling T-Square Test
  – Testing: estimate Query Distribution with **GMM**
Modeling the Query Distribution
Modeling the Query Distribution

Distribution of the semantic group “road”
Modeling the Query Distribution

Distribution of the semantic group “road”

Distribution of the semantic group “car”

Distribution of the query “car on road”
Balancing the Exploration and Exploitation
Balancing the Exploration and Exploitation

- The priority of selecting an instance $x$ to query next (i.e., present to the searcher for labeling)

Balancing Factor

$$\rho(x) = \lambda \text{Harvest}(x) + (1 - \lambda) \text{Explore}(x)$$

Exploitative Priority

Explorative Priority

Balancing Factor
Harvest($x$): Exploitative Priority

- How likely the exploitation will be boosted when $x$ is selected to query next.

$$Harvest(x) = P(x|Q)P(Q|x)$$
How likely the exploitation will be boosted when $x$ is selected to query next

$Harvest(x) = P(x|Q)P(Q|x)$
Harvest($x$): Exploitative Priority

- How likely the exploitation will be boosted when $x$ is selected to query next

\[ \text{Harvest}(x) = P(x|Q)P(Q|x) \]
Harvest($x$): Exploitative Priority

- How likely the exploitation will be boosted when $x$ is selected to query next

\[ \text{Harvest}(x) = P(x|Q)P(Q|x) \]
Explore($x$): Explorative Priority

- How likely the exploration will be improved when $x$ is selected to query next

$$Explore(x) = \left\{ -P(x|Q) \log(P(x|Q)) \right\} \times \left\{ -P(Q|x) \log(P(Q|x)) \right\}$$

Entropy of the prior distribution

Entropy of the posterior distribution
Explore($x$): Explorative Priority

- How likely the exploration will be improved when $x$ is selected to query next

$$Explore(x) = \left\{ -P(x|Q) \log(P(x|Q)) \right\} \times \left\{ -P(Q|x) \log(P(Q|x)) \right\}$$
\( \lambda: \text{Balancing Factor} \)

- Updated by investigating the Harvest History, i.e., number of relevant instances found in previous rounds

\[
\lambda_{i+1} = \lambda_i + \frac{\left| \mathcal{L}_{i+1}^+ \right| - \bar{R}}{R}
\]

The expected harvest
Experiments

- Learning $pdf(s)$ of the semantic groups
  - Large Scale Concept Ontology for Multimedia (LSCOM), 449 concept labeled on TRECVID 2005 development dataset
  - Using *concept combination* to create groups
  - Results: 23,064 groups and their $pdfs$
    (http://www.cs.cityu.edu.hk/~xiaoyong/CAL/)
Experiments

• Comparison with Uncertainty Sampling (US)
  – 12 volunteers (age 19~26, 3 girls and 9 boys)
  – TREVID 2005 test dataset (45,765 video shots in English/Chinese/Arabic)
  – 24 TRECVID queries and ground truth
  – Mean Average Precision (MAP)
  – Concept-based search for the initial round
  – Interface
Experiments
Experiments

- Comparison with Uncertainty Sampling (US)
Experiments

- Comparison with Uncertainty Sampling (US)
  - Purely explorative US causes early give-up

“difficult” queries
Experiments

- Study of user patience on labeling

<table>
<thead>
<tr>
<th>Method</th>
<th>#Examined Shots</th>
<th>#Irrelevant Shots</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Max</td>
<td>Avg.</td>
</tr>
<tr>
<td>CAL</td>
<td>5462</td>
<td>2445</td>
</tr>
<tr>
<td>US</td>
<td>4806</td>
<td>2037</td>
</tr>
</tbody>
</table>
Experiments

• Can GMM stimulate the query distribution well?

Query “the road with one more cars”
Experiments

• Comparison with the state-of-the-art
  – CAL (with virtual searchers) vs. The best results reported in TRECVID 06-09

<table>
<thead>
<tr>
<th></th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>0.303</td>
<td>0.362</td>
<td>0.194</td>
<td>0.246</td>
</tr>
<tr>
<td>CAL</td>
<td>0.422</td>
<td>0.493</td>
<td>0.331</td>
<td>0.380</td>
</tr>
</tbody>
</table>
Conclusions

• Advantages of CAL
  – Predictable Query Distribution
  – Fast Convergence
  – Balance between Exploitation and Exploration in a principled way

• Toys and demos available at http://www.cs.cityu.edu.hk/~xiaoyong/CAL/
Conclusions

• Future work
  – Replace the *Harvest()* and *Explore()* with more advanced ones
  – Try it on larger datasets or web videos
Thanks!
Thanks!