Improving Ad Relevance in Sponsored Search

Dustin Hillard, Stefan Schroedl, Eren Manavoglu, Hema Raghavan and Chris Leggetter
Motivation

• Improve ad relevance for search users
• Develop a model to predict relevance
• Leverage user interactions in learning
• Use predicted relevance to improve system
  – As a filter to remove bad ads
  – As a feature to improve ad ranking
  – As a score for improving ad page placement
Outline

- Motivation
- Ad Relevance Models
  - Baseline model
  - Learning from user clicks
- Sponsored Search Applications
- Conclusion
Outline

• Motivation
• Ad Relevance Models
  – Baseline model
  – Learning from user clicks
• Sponsored Search Applications
• Conclusion
An Ad Relevance Model

• Develop a model to predict how relevant an ad is for a particular query
• Incorporate typical IR features such as word and character overlap, word novelty
• Train a machine learned model based on human generated editorial judgments
### Relevance Modeling Data

- Retrieve about 20 ads per query with a typical information retrieval system
- Stratified query sample from web logs
- Binary ‘good’ vs. ‘bad’ editorial judgments

<table>
<thead>
<tr>
<th>Data</th>
<th>Queries</th>
<th>Query-Ad Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>4.8k</td>
<td>95k</td>
</tr>
<tr>
<td>Test</td>
<td>2.3k</td>
<td>47k</td>
</tr>
</tbody>
</table>
Baseline Model: Results

- Baseline features:
  - Character, word and bigram overlap
  - Ordered bigram overlap
  - Cosine match (TF/IDF)
  - Query length

<table>
<thead>
<tr>
<th>Model</th>
<th>Precision</th>
<th>Recall</th>
<th>F-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>maxent</td>
<td>0.658</td>
<td>0.458</td>
<td>0.540</td>
</tr>
<tr>
<td>adaBoost</td>
<td>0.670</td>
<td>0.543</td>
<td>0.600</td>
</tr>
<tr>
<td>GBDT</td>
<td>0.671</td>
<td>0.551</td>
<td>0.605</td>
</tr>
</tbody>
</table>
Baseline Precision/Recall

The graph shows the precision-recall curves for different models. The models compared are adaBoost, gbdt, maxent, and tfidf. The curves indicate the trade-off between precision and recall for various thresholds.
Outline

• Motivation

• Ad Relevance Models
  – Baseline model
  – Learning from user clicks

• Sponsored Search Applications

• Conclusion
Learning From User Clicks

• Incorporate information from user click behavior to improve relevance modeling
• Include historical click information:
  – Directly for specific observed click rates
  – Broadly with a query->ad click translation model
Observed Click History

- Previous click history is the best predictor of future click behavior
- Collect aggregate click rate statistics from our logs at multiple levels of granularity
  - Query-Ad, Query-Advertiser levels
  - Ad, Advertiser levels
  - Query level
- Broader aggregates are less precise but have higher coverage
Insufficient Click History

• No history is available for previously unseen ads, or infrequent query-ad pairs
• Develop a model that predicts click propensity based only on query-ad text
• Learn a relationship between a query and an ad title that can be applied to unseen query-ad pairs
A Click Translation Model

• Learn a query->title translation model

\[ p(D|Q) = p(Q|D)p(D)/p(Q) \]

• IBM Model I, with web logs as corpus

\[ p(Q|D) = \prod_{j=0}^{m} \sum_{i=0}^{n} \text{trans}(q_j|d_i) \]
\[ \text{trans}(q_j|d_i) = \frac{\sum_{\text{logs}} \text{count}(q_j|d_i)}{\sum_q \sum_{\text{logs}} \text{count}(q|d_i)} \]

• Compare 2 models: click-based, view-based

\[ \text{clickLikelihood} = \frac{p_{\text{click}}(Q|D)}{p_{\text{ec}}(Q|D)} \]
Using Clicks: Results

• Start with baseline GBDT model
  – Add observed click history features only
  – Add click translation scores only
  – Add both together

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline (GBDT)</td>
<td>0.671</td>
<td>0.551</td>
<td>0.605</td>
</tr>
<tr>
<td>+click history</td>
<td>0.699</td>
<td>0.557</td>
<td>0.620</td>
</tr>
<tr>
<td>+translations</td>
<td>0.658</td>
<td>0.590</td>
<td>0.622</td>
</tr>
<tr>
<td>+click +trans</td>
<td>0.673</td>
<td>0.584</td>
<td>0.625</td>
</tr>
</tbody>
</table>
Precision/Recall With Clicks
Outline

• Motivation
• Ad Relevance Models
  – Baseline model
  – Learning from user clicks
• Sponsored Search Applications
• Conclusion
Example of Sponsored Search
(typical)

Nikon Digital Cameras
Professional Quality Photos w/ Nikon's Easy to Use Digital Cameras.
www.NikonUSA.com

Save on Digital Camera
Lowest Prices on Name Brand Cameras in Stock. Buy Now.
www.TigerDirect.com

Kodak Sweetheart Deals
Save up to $70 on select Kodak Digital Cameras. Only at Kodak.
www.Kodak.com

Deals on Digital Cameras
Canon, Sony, Kodak Digital Cameras @ Low Prices. Save w/ Free Shipping.
www.circuitcity.com

Digital Cameras - Shopping Results

Brand
Sony (448)
Canon (824)
Nikon (513)
Olympus (520)

Popular Products
Panasonic Lumix DMC-TZ3S... ★★★★★ (3)
Panasonic Lumix DMC-TZ3-K... ★★★★★ (13)
Canon EOS 30D (Body Only)... ★★★★★ (15)
Nikon D200 (Body Only)... ★★★★★ (33)

Digital Camera - Wikipedia, the free encyclopedia
Types of... | Conversion of... | History | Image...
A digital camera is a camera that takes video or still photographs,
Example of Sponsored Search
(could be better)
Ad Filtering

- Problem: Remove low quality ads
- Approach: Filter low relevance score ads
- Impact:
  - Filtered 50% of Bad ads, less than 10% of Good
  - Bucket metrics:

<table>
<thead>
<tr>
<th>Metric</th>
<th>Relative Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>coverage</td>
<td>-8.7%</td>
</tr>
<tr>
<td>ad depth</td>
<td>-11.9%</td>
</tr>
<tr>
<td>ad CTR</td>
<td>+10.1%</td>
</tr>
<tr>
<td>total ad clicks</td>
<td>+0.5%</td>
</tr>
</tbody>
</table>
Ad Ranking

• Problem: Rank ads by bid and $p(\text{click})$
• Approach: Provide relevance as feature
• Impact:
  – Improves click model when history is sparse
Optimization

• Problem: Place ads on the search page
• Approach: Consider ad and web relevance
• Impact:
  – Reduced low quality ads above search results
  – Bucket metrics:

<table>
<thead>
<tr>
<th></th>
<th>Relative Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Ad Impact</td>
<td>-4.5%</td>
</tr>
<tr>
<td>North ad CTR</td>
<td>+1.5%</td>
</tr>
<tr>
<td>total ad clicks</td>
<td>+0.8%</td>
</tr>
</tbody>
</table>
Outline

• Motivation
• Ad Relevance Models
  – Baseline model
  – Learning from user clicks
• Sponsored Search Applications
• Conclusion
Summary

- Developed a useful ad relevance model
- Improved performance with user click data
- Extend to new ads with click trans. model
- Incorporated in sponsored search system:
  - Removed low quality ads
  - Improved ad ranking
  - Improved ad placement
THANKS!

QUESTIONS?