Building Bridges across Social Platforms:

Mossaab Bagdouri  
Computer Science  
iSchool & UMIACS  
University of Maryland  
College Park, USA  
mossaab@umd.edu

Douglas W. Oard  
Computer Science  
iSchool & UMIACS  
University of Maryland  
College Park, USA  
oard@umd.edu

Answering Twitter Questions with Yahoo! Answers

Introduction

Sponsored by Qatar National Research Fund (NPRP 6-1377-1-257), an IBM Ph.D. Fellowship, and SIGIR Student Travel Grant

19% have replies
81% no replies

Experiments

Language Model: 1.6B tweets

- Vocabulary
- N-gram language model
- Inverse document frequency
- Word embeddings

Hashtag Segmentation

Should the govt be paying for #healthinsurance for immigrants or should private companies?

Synonyms

Should the #government be paying for health insurance for immigrants or should private companies?

Misspelling Correction

Why did the great awaking happen? Did it result in peace in England?

Duplicate Questions

Quora

- Question pairs:  
  - 150K duplicate
  - 250K non-duplicate

Annotations

- 362 aqweets (answerable question tweets)
- CrowdFlower workers
- 4-level relevance scale: [0 = bad, 1 = fair, 2 = good, 3 = excellent]
- Score = weighted average label

Results

<table>
<thead>
<tr>
<th>Answers</th>
<th>Body</th>
<th>Title</th>
<th>3 Rewriters + Title</th>
<th>Twitter IDF + Title</th>
<th>L2R</th>
<th>Oracle</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.74</td>
<td>0.75</td>
<td>1.32</td>
<td>1.27</td>
<td>1.32</td>
<td>1.37</td>
<td>1.86</td>
</tr>
</tbody>
</table>

bad = 0  
fair = 1  
good = 2  
excellent = 3

1.4 B QA Pairs
English: 673M ANSWERS

Mossaab Bagdouri  
Douglas W. Oard

Computer Science  
iSchool & UMIACS  
University of Maryland  
College Park, USA

mossaab@umd.edu  
oard@umd.edu