Building Bridges across Social Platforms:

Mossaab Bagdouri Douglas W. Oard Computer Science iSchool & UMIACS

University of Maryland College Park, USA

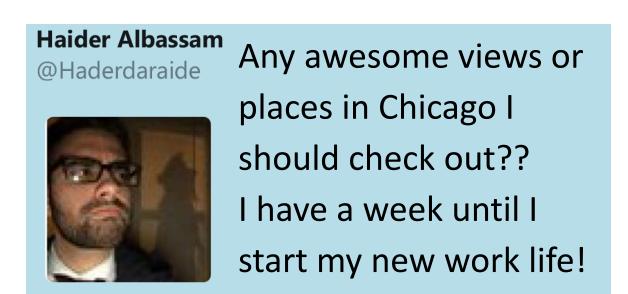
mossaab@umd.edu oard@umd.edu



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

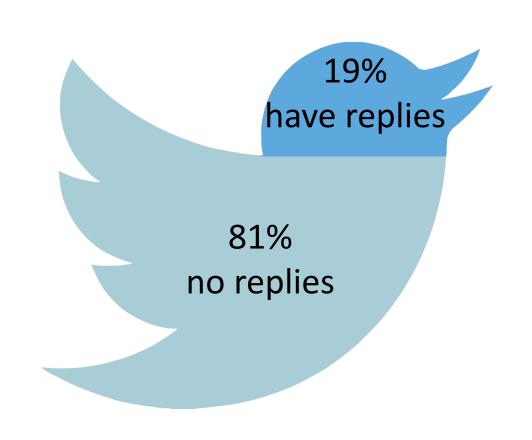


Title I'm goir for eati

Answer

I'm going to Chicago in the summer this is my first time going what places should i go check out? for eating or shopping or sightseeing....ECT?

**Best Answer: Museum of Science and Industry, Navy Pier (attractions and restaurants), Water Tower Place (shopping), Lincoln Park Zoo, Wrigley Field (take in a Cub's ball game), Shedd Aquarium, Adler Planetarium, Sears Tower (spectacular view of the city), Field Museum, Art Institute, Hard Rock Café (food and atmosphere), eat delicious Chicago pizza (Home Run Inn, Gino's, Pizzeria Uno, Giordano's), Oak Street Beach (sunbathing and people watching).







Experiments

Language Model: 1.6B tweets

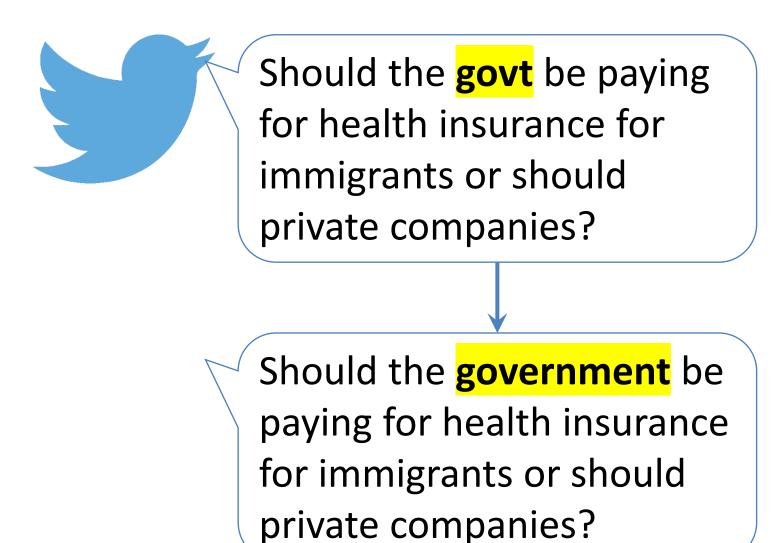
- Vocabulary
- N-gram language model
- Inverse document frequency
- Word embeddings

Hashtag Segmentation

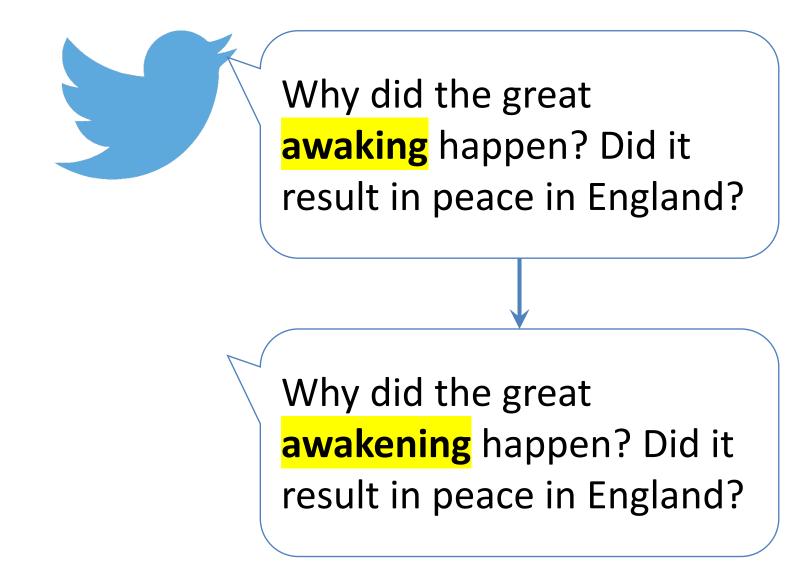
Should the govt be paying for #healthinsurance for #immigrants or should private companies? Should the govt be paying for health insurance for immigrants or should

private companies?

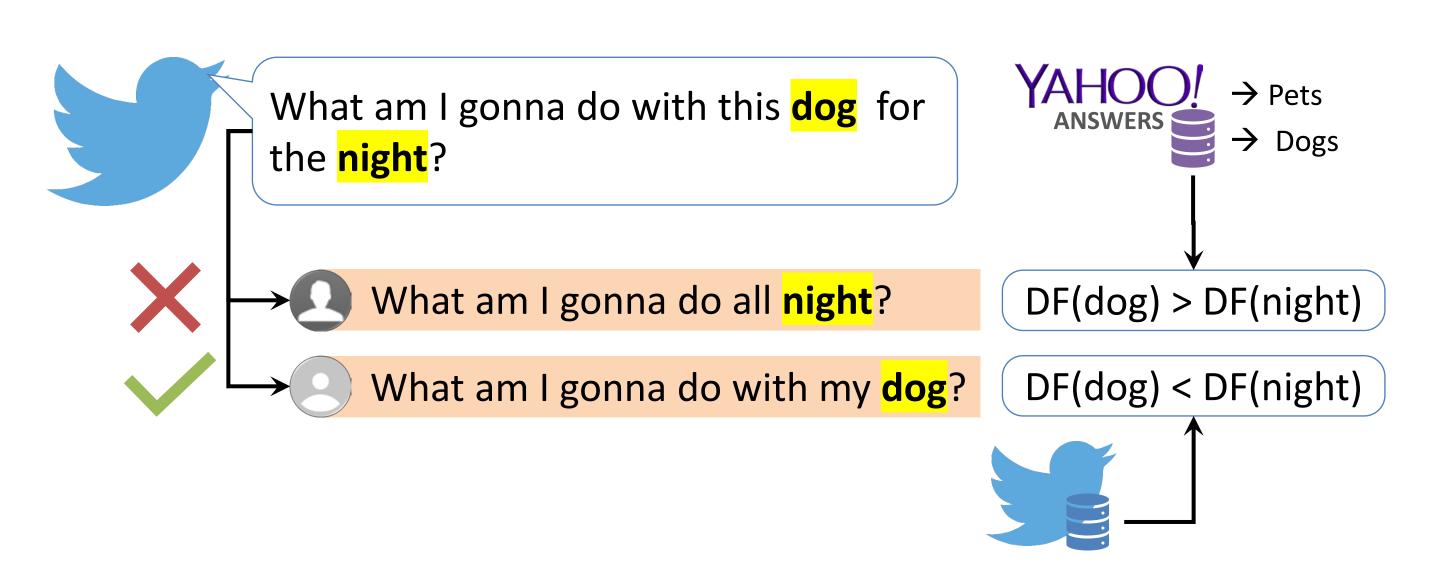
Synonyms



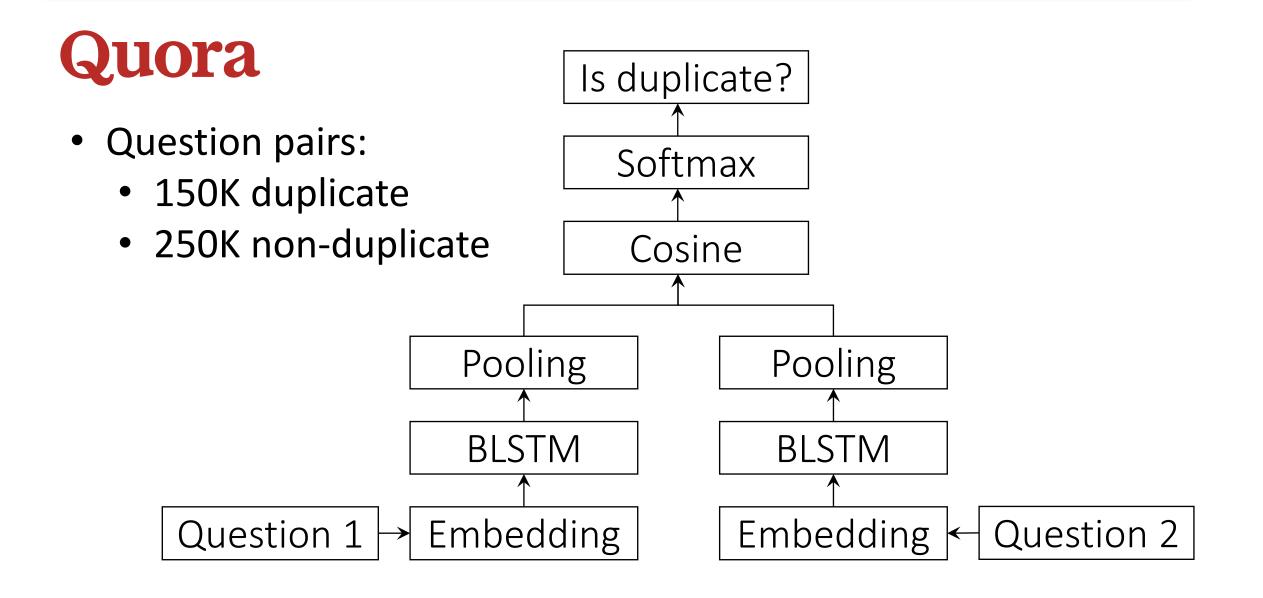
Misspelling Correction



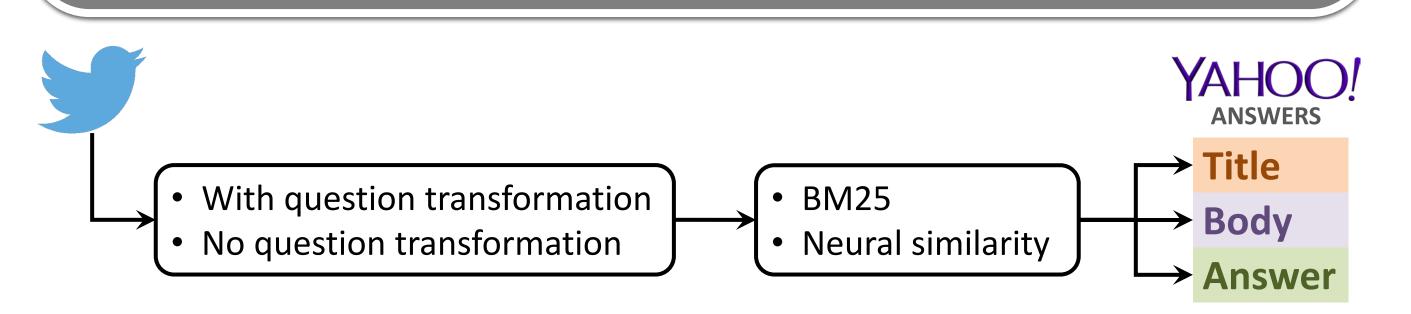
Inverse Document Frequency



Duplicate Questions



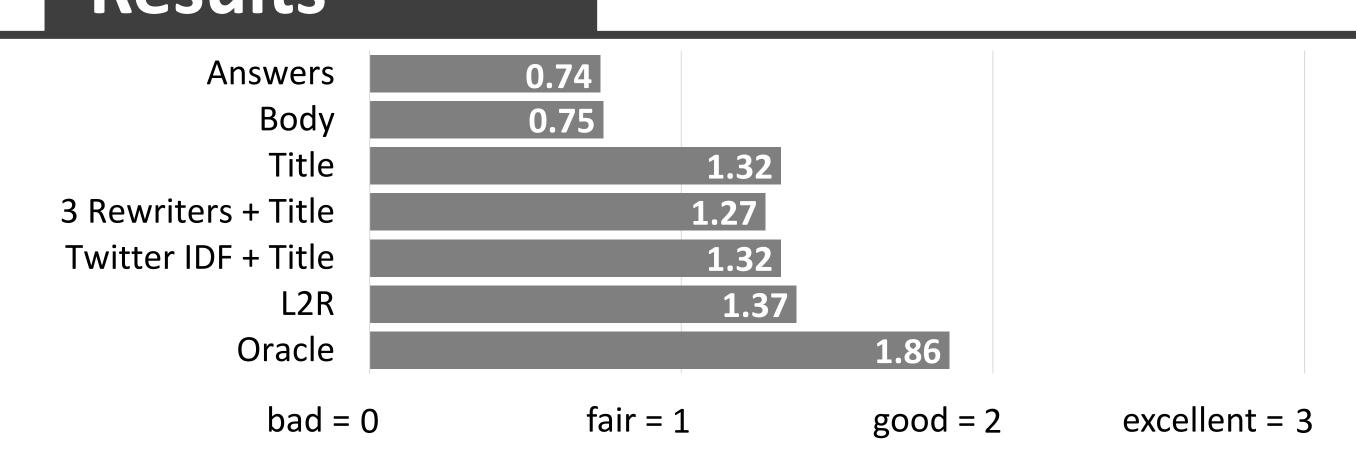
Learning to Rank



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



- 2/3 of aqweets have an excellent answer in Yahoo! Answers
- Title is the single best field
- No improvement observed from question rewriters
- Small improvement observed from L2R
- Size of corpus limits some comparisons