Overview of the 2014 NLP Unshared Task in PoliInformatics

Noah A. Smith* **Claire Cardie†** **Anne L. Washington‡** **John D. Wilkerson§**
*School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA
†Department of Computer Science, Cornell University, Ithaca, NY 14853, USA
‡School of Public Policy, George Mason University, Arlington, VA 22201, USA
§Department of Political Science, University of Washington, Seattle, WA 98195, USA
*Corresponding author: nasmith@cs.cmu.edu

Abstract

We describe a research activity carried out during January–April 2014, seeking to increase engagement between the natural language processing research community and social science scholars. In this activity, participants were offered a corpus of text relevant to the 2007–8 financial crisis and an open-ended prompt. Their responses took the form of a short paper and an optional demonstration, to which a panel of judges will respond with the goal of identifying efforts with the greatest potential for future interdisciplinary collaboration.

1 Introduction

In recent years, numerous interdisciplinary research meetings have sought to bring together computer scientists with expertise in automated text data analysis and scholars with substantive interests that might make use of text data. The latter group has included political scientists, economists, and communications scholars. An NSF Research Coordination Network grant to encourage research using open government data was awarded to co-authors Washington and Wilkerson in 2013. The network for Political Informatics, or PoliInformatics, brought together a steering committee from diverse research backgrounds that convened in February 2013. At that meeting, a substantive focus on the 2007–8 nancial crisis was selected.

Drawing inspiration from the "shared task" model that has been successful in the natural language processing community, we designed a research competition for computer scientists. In a shared task, a gold-standard dataset is created in advance of the competition, inputs and outputs are defined by the organizers, typically creating a supervised learning setup with held-out data used for evaluation. Constraints on the resources that may be used are typically set in place as well, to focus the energies of participants on a core problem, and the official evaluation scores are published, usually as open-source software. Final systems (or system output) is submitted by a deadline and judged automatically against the gold-standard. Participants report on their systems in short papers, typically presented at a meeting associated with a conference or workshop.

With neither a clear definition of what the final outcome might be, nor the resources to create the necessary gold-standard data, we developed a more open-ended competition. A text corpus was collected and made available, and a prompt was offered. Participants were given freedom in how to respond; competition entries took the form of short research papers and optional demonstrations of the results of the projects. Rather than an objective score, a panel of judges organized by the PoliInformatics steering committee offered public reviews of the work, with an emphasis on potential for future interdisciplinary research efforts that might stem from these preliminary projects.

2 Setup

The prompts offered to participants were:

Who was the financial crisis? We seek to understand the participants in the lawmaking and regulatory processes that formed the government's response to the crisis: the individuals, industries, and professionals targeted by those policies; the agencies and organizations responsible for implementing them; and the lobbyists, witnesses, advocates, and politicians who were actively involved—and the connections among them.

What was the financial crisis? We seek to understand the cause(s) of the crisis, proposals for reform, advocates for those proposals, arguments

Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science, pages 5–7,
Baltimore, Maryland, USA, June 26, 2014.

for and against, policies ultimately adopted by the government, and the impact of those policies.

The set of datasets made available is listed in Table 1. Several additional datasets were suggested on the website,[1] but were not part of the official data.

3 Response

Forty teams initially registered to participate in the unshared task; ten submitted papers. The teams came from a variety of institutions spread across six countries. Half of the teams included links to online demonstrations or browsable system output. At this writing, the papers are under review by the panel of judges. We provide a very brief summary of the contributions of each team.

3.1 Who was the financial crisis?

Bordea et al. (2014) inferred importance and hierarchy of topics along with expertise mining to find which participants in the discourse might be experts (e.g., Paul Volcker and "proprietary trading") based on FOMC, FCIC, and Congressional hearing and report data.

Baerg et al. (2014) considered transcripts of the FOMC, developing a method for scaling the preferences of its members with respect to inflation (hawks to doves); the method incorporates automatic dimensionality reduction and expert topic interpretation.

Zirn et al. (2014) also focused on the transcripts, distinguishing between position-taking statements and shorter "discussion elements" that express agreement or disagreement rather than substance, and used this analysis to quantify similarity among FOMC members and take first steps toward extraction of sub-dialogues among them.

Bourreau and Poibeau (2014) focused on the FCIC report and the two Congressional reports, identifying named entities and then visualizing correlations among mentions both statically (as networks) and dynamically. Clark et al. (2014) considered Congressional hearings, applying a reasoning model that integrates analysis of social roles and relationships with analysis of individual beliefs in hope of detecting opinion shifts and signs of influence.

With an eye toward substantive hypotheses about dependencies among banks' access to bailout funds relating to underlying social connections, Morales et al. (2014) automatically extracted a social network from the corpus alongside structured data in Freebase.

3.2 What was the financial crisis?

Miller and McCoy (2014) considered FOMC transcripts, applying topic models for dimensionality reduction and viewing topic proportions as time series.

In a study of the TARP, Dodd-Frank, and the health reform bills, Li et al. (2014) explored the ideas expressed in those bills, applying models of text reuse from bills introduced in the 110th and 111th Congresses.

Wang et al. (2014) implemented a query-focused summarization system for FOMC and FCIC meeting transcripts and Congressional hearings, incorporating topic and expertise measures into the score, and queried the corpus with candidate causes for the crisis, derived from Wikipedia (e.g., "subprime lending" and "growth housing bubble").

Kleinnijenhuis et al. (2014) considered Congressional hearings alongside news text from the United States and the United Kingdom, carrying out keyword analysis to compare and measure directional effects between the two, on different dimensions.

4 Conclusion

The unshared task was successful in attracting the interest of forty participants working on ten teams. A highly diverse range of activities ensued, each of which is being reviewed at this writing by a panel of judges. Reviews and final outcomes will be posted at the `https://sites.google.com/site/unsharedtask2014` as soon as they are available, and a presentation summarizing the competition will be part of the ACL 2014 Workshop on Language Technologies and Computational Social Science.

Acknowledgments

We thank the participants and judges for their time and effort. This activity was supported in part by NSF grants 1243917 and 1054319.

[1] `https://sites.google.com/site/unsharedtask2014`

- Federal Open Market Committee (FOMC):
  - Meeting transcripts are only made available five years after each meeting date. (The 2008 transcripts came available around the time of the activity and were kindly made available by participant William Li.)
  - Meeting minutes are available for all meetings to date.
- Federal Crisis Inquiry Commission (FCIC; an independent commission created by Congress to investigate the causes of the crisis):
  - Report
  - Transcript of the first public hearing
- Congressional reports:
  - Senate Committee on Homeland Security and Governmental Affairs: "Wall Street and the financial crisis: anatomy of a financial collapse"
  - House Committee on Financial Services: "The stock market plunge: what happened and what is next?"
- Congressional bills:
  - Troubled Assets Relief Program, 2008 (TARP)
  - Dodd-Frank Wall Street Reform and Consumer Protection Act (2010)
  - American Recovery and Reinvestment Act of 2009 (Stimulus)
  - Housing and Economic Recovery Act of 2008
  - Public Company Accounting Reform and Investor Protection Act of 2002 (Sarbanes-Oxley)
  - Financial Services Modernization Act of 1999 (Gramm-Leach-Bliley)
  - In addition to the above financial reform bills, the text of all versions of all Congressional bills introduced in the 110th and 111th Congresses
- Congressional hearings, segmented into turns:
  - Monetary policy (26)
  - TARP (12)
  - Dodd-Frank (61)
  - Other selected committee hearings relating to financial reform (15)

Table 1: Text datasets made available to unshared task participants. These can be downloaded at `https://sites.google.com/site/unsharedtask2014`.

References

Nicole Rae Baerg, Will Lowe, Simone Ponzetto, Heiner Stuckenschmidt, and Cäcilia Zirn. 2014. Estimating central bank preferences.

Georgeta Bordea, Kartik Asooja, Paul Buitelaar, and Leona O'Brien. 2014. Gaining insights into the global financial crisis using Saffron.

Pierre Bourreau and Thierry Poibeau. 2014. Mapping the economic crisis: Some preliminary investigations.

Micah Clark, Adam Dalton, Tomas By, Yorick Wilks, Samira Shaikh, Ching-Sheng Lin, and Tomek Strzalkowski. 2014. Influence and belief in Congressional hearings.

Jan Kleinnijenhuis, Wouter van Atteveldt, and Antske Fokkens. 2014. Chicken or egg? the reciprocal influence of press and politics.

William P. Li, David Larochelle, and Andrew W. Lo. 2014. Estimating policy trajectories during the financial crisis.

John E. Miller and Kathleen F. McCoy. 2014. Changing focus of the FOMC through the financial crisis.

Michelle Morales, David Brizan, Hussein Ghaly, Thomas Hauner, Min Ma, and Andrew Rosenberg. 2014. Application of social network analysis in the estimation of bank financial strength during the financial crisis.

Lu Wang, Parvaz Mahdabi, Joonsuk Park, Dinesh Puranam, Bishan Yang, and Claire Cardie. 2014. Cornell expert aided query-focused summarization (CEAQS): A summarization framework to PoliInformatics.

Cäcilia Zirn, Michael Schäfer, Simone Paolo Ponzetto, and Michael Strube. 2014. Exploring structural features for position analysis in political discussions.