ACL 2014

Joint Workshop on Social Dynamics and Personal Attributes in Social Media

Proceedings of the Workshop

June 27, 2014
Baltimore, Maryland, USA

Introduction

These proceedings contain papers presented at the Joint Workshop on Social Dynamics and Personal Attributes in Social Media. The workshop was held in Baltimore, Maryland, USA and hosted in conjunction with the 52nd Annual Meeting of the Association for Computational Linguistics.

This workshop was intended to serve as a forum for sharing research in:

I NLP and Social Dynamics: Language is a set of publicly agreed conventions that serves the purpose of inter-personal communication. Speakers (or writers) try to convey a message, instill an idea or make an impression on the listeners. Listeners (or readers), in turn, are affected by the message and may respond to it. Language, in that sense, is an important vehicle that shapes (and is shaped by) social dynamics.

Traditional NLP research, however, focuses on "documents" (either of full length or on the sentence level), rather than on the communication process as reflected by language use. Common examples of traditional NLP research are parsing, document classification, machine translation, and sentiment analysis at the sentence and document level without considering the social dynamics of the people who are writing and reading those texts. We propose to move beyond analyzing the informational aspect of documents and discuss ways in which NLP can contribute to gaining insights about the interplay between language use and various levels of social dynamics.

II Personal Attributes in Social Media: There are many important social science questions and commercial applications that are impacted by the large amounts of diverse personalized data emerging from social media. These data can reveal user interests, preferences and opinions, as well as trends and activity patterns for companies and their products.

The automatic prediction of latent attributes from discourse in social media includes topics such as: inferring user/customer demographic profiles (gender, age, religion, social status, race, ethnicity, origin); predicting user interests (sports, movies) and preferences (political favorites or product likes); classifying sentiment, personality, emotional states (onset of depression), and opinions held by an author; and analyzing general trends and influence for companies and products.

We invited original and unpublished research papers on all topics related to NLP and social dynamics and text-driven attribute prediction in social media, including but not limited to the topics listed below.

NLP and Social Dynamics:

- Emergence and diffusion of slang, neologisms and metaphors

- Emotion dynamics in social media conversation threads

- Evolution of word formation and word meaning

- Language coordination and lexical entrainment

- Language evolution through history and language variation across communities

- Linguistic and social factors in acceptance of new words and phrases

- Persuasive language and (online) campaigns

- Pragmatics of language

- Social dynamics in (blog/news story) comment threads

Personal Attributes in Social Media:

- Dynamic and streaming nature of social media data

- Data collection, sharing and labeling biases for personal analytics in social media

- Joint latent attribute prediction (e.g., age together with political preference)

- Emotional states, distress, mental condition classification in social networks

- Mood, sentiment, emotion and opinion analysis of authors in social media

- Multi-relational aspect of social media (e.g., networks of friends, followers, user mentions etc.)

- Scalability to other understudied languages and dialects in social media

- Security, identity and privacy issues for personal analytics in social media.

The Workshop Committee received 24 submissions. Three reviewers reviewed each submission. For the final workshop program, 15 regular papers, 5 to 11 pages each, were selected, 3 papers were rejected and 3 withdrawn. 3 papers have chosen to be non-archived. Authors affiliations include Computer and Information Science, Linguistics, Political Science, Nutrition and Health Promotion, History, English and Business and Management.

We would like to thank all Program Committee members and external reviewers for their effort in providing high-quality reviews. We thank all the authors who submitted their papers. Many thanks to our invited speakers and panel participants.

Alice Oh,
Benjamin Van Durme,
David Yarowsky,
Oren Tsur,
Svitlana Volkova

Organizers:

Alice Oh, Korea Advanced Institute of Science and Technology
Benjamin Van Durme, Johns Hopkins University
David Yarowsky, Johns Hopkins University
Oren Tsur, Harvard University, Northeastern University
Svitlana Volkova, Johns Hopkins University

Program Committee:

Abigail Jacobs, University of Colorado at Boulder, USA
Alan Ritter, Carnegie Mellon University, USA
Alejandro Jaimes, Yahoo Research, Barcelona
Alessandro Moschitti, Qatar Computing Research Institute, Qatar
Ancsa Hannak, Northeastern University, USA
Ari Rappaport, The Hebrew University, Israel
Aron Culotta, Illinois Institute of Technology, USA
Brendan O'Connor, Carnegie Mellon University, USA
Brian Keegan, Northeastern University, USA
Carlo Strapparava, FBK, Italy
Chin-Yew Lin, Microsoft Research Asia
Chris Dyer, Carnegie Mellon University, USA
Cristian Danescu-Niculescu-Mizil, Max Planck Institute SWS
Dan Jurafsky, Stanford University, USA
Daniel Romero, University of Michigan, USA
David Smith, Northeastern University, USA
Delip Rao, Amazon, USA
Derek Ruths, McGill University, Canada
Dong Nguyen, University of Twente, Netherlands
Eugene Kharitonov, Yandex, Russia
Francisco Iacobelli, Northeastern Illinois University, USA
Gideon Dror, Yahoo! Research, USA
Glen Coppersmith, Johns Hopkins University, USA
Haewoon Kwak, Qatar Computing Research Institute, Qatar
Idan Szpektor, Yahoo! Research, USA
Ilia Chetviorkin, Lomonosov Moscow State University, Russia
Ingmar Weber, QCRI, Qatar
Jacob Eisenstein, Georgia Institute of Technology, USA
James Caverlee, Texas University, USA
John Henderson, MITRE, USA
Margaret Mitchell, Microsoft Research, USA
Mark Dredze, Johns Hopkins University, USA
Michael Gamon, Microsoft Research, USA
Michael Paul, Johns Hopkins University, USA
Omri Abend, University of Edinburgh, UK
Patrick Pantel, Microsoft Research, USA

Paul Cook, University of Melbourne, Australia
Pavel Braslavski, KonturLabs, Russia
Pavel Sergyukov, Yandex, Russia
Philip Resnik, University of Maryland, USA
Rebecca Knowles, Johns Hopkins University, USA
Reut Tzarfati, Weizman Institute of Science, Israel
Rivka Levitan, Columbia University, USA
Roi Reichart, Israel Institute of Technology, Israel
Roy Schwartz, The Hebrew University, Israel
Saif Mohammad, National Research Council, Canada
Silviu-Petru Cucerzan, Microsoft Research, USA
Souneil Park, University of Michigan, USA
Vasileios Lampos, University College London, UK
Vivi Nastase, FBK, Italy
Yael Netzer, Ben Gurion University, Israel
Yoav Goldberg, Bar Ilan University, Israel
Yu-Ru Lin, University of Pittsburgh, USA
Yuval Pinter, Yahoo! Research, Israel

Invited Speakers:

Derek Ruths, McGill University
Henry Kautz, University of Rochester

Panelists:

Hanna Wallach, UMass Amherst
Jacob Eisenstein, Georgia Tech
Cristian Danescu-Niculescu-Mizil, Max Planck Institute SWS
Jimmy Lin, University of Maryland

Table of Contents

Conference Program

Friday, June 27, 2014

(8:40–8:50) Welcome Notes

8:50–9:50   Invited talk by Derek Ruths

Oral Session 1

10:05–10:25   *Detecting Sociostructural Beliefs about Group Status Differences in Online Discussions*
Brian Riordan, Heather Wade and Afzal Upal

10:25–10:45   *Using County Demographics to Infer Attributes of Twitter Users*
Ehsan Mohammady and Aron Culotta

Oral Session 2

10:55–11:10   *The Enrollment Effect: A Study of Amazon's Vine Program*
Dinesh Puranam and Claire Cardie

11:10–11:30   *Discourse Analysis of User Forums in an Online Weight Loss Application*
Lydia Manikonda, Heather Pon-Barry, Subbarao Kambhampati, Eric Hekler and David W. McDonald

11:30–12:30   Panel by Hanna Wallach, Jacob Eisenstein, Cristian Danescu-Niculescu-Mizil and Jimmy Lin

Oral Session 3

1:30–1:50   *A Unified Topic-Style Model for Online Discussions*
Ying Ding, Jing Jiang and Qiming Diao

1:50–2:10   *Self-disclosure topic model for Twitter conversations*
JinYeong Bak, Chin-Yew Lin and Alice Oh

2:10–2:25   *Detecting and Evaluating Local Text Reuse in Social Networks*
Shaobin Xu, David Smith, Abigail Mullen and Ryan Cordell

2:35–3:35   Invited talk

Oral Session 4

3:45–4:00   *Generating Subjective Responses to Opinionated Articles in Social Media: An Agenda-Driven Architecture and a Turing-Like Test*
Tomer Cagan, Stefan L. Frank and Reut Tsarfaty

4:00–4:15   *A Semi-Automated Method of Network Text Analysis Applied to 150 Original Screenplays*
Starling Hunter

Poster Session

4:30–5:30   *Power of Confidence: How Poll Scores Impact Topic Dynamics in Political Debates*
Vinodkumar Prabhakaran, Ashima Arora and Owen Rambow

4:30–5:30   *As Long as You Name My Name Right: Social Circles and Social Sentiment in the Hollywood Hearings*
Oren Tsur, Dan Calacci and David Lazer

4:30–5:30   *Towards Tracking Political Sentiment through Microblog Data*
Yu Wang, Tom Clark, Jeffrey Staton and Eugene Agichtein

4:30–5:30   *Innovation of Verbs in Hebrew*
Ornan Uzzi

4:30–5:30   *User Type Classification of Tweets with Implications for Event Recognition*
Lalindra De Silva and Ellen Riloff

4:30–5:30   *Collective Stance Classification of Posts in Online Debate Forums*
Dhanya Sridhar, Lise Getoor and Marilyn Walker

Friday, June 27, 2014 (continued)

(5:30–5:40) Closing Remarks