ACL 2014

5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis
WASSA 2014

Proceedings of the Workshop

June 27, 2014
Baltimore, Maryland, USA

Endorsed by SIGSEM - ACL's Special Interest Group on Computational Semantics
Endorsed by SIGNLL - ACL's Special Interest Group in Natural Language Learning

©2014 The Association for Computational Linguistics

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-941643-11-2

Introduction

Research in automatic Subjectivity and Sentiment Analysis (SSA), as subtasks of Affective Computing and Natural Language Processing (NLP), has flourished in the past years. The growth in interest in these tasks was motivated by the birth and rapid expansion of the Social Web that made it possible for people all over the world to share, comment or consult content on any given topic. In this context, opinions, sentiments and emotions expressed in Social Media texts have been shown to have a high influence on the social and economic behavior worldwide. SSA systems are highly relevant to many real-world applications (e.g. marketing, eGovernance, business intelligence, social analysis) and also to many tasks in Natural Language Processing (NLP) - information extraction, question answering, textual entailment, to name just a few. The importance of this field has been proven by the high number of approaches proposed in research in the past decade, as well as by the interest that it raised from other disciplines (Economics, Sociology, Psychology) and the applications that were created using its technology.

Despite the large interest shown by the research community and the development of a set of benchmarking resources and methods to tackle sentiment analysis, SSA remains far from being a solved issue. While systems working for English on customer reviews obtain good results in sentiment classification, systems working for other languages or on Social Media texts are still struggling to surpass the baseline. As such, it is necessary to continue the sentiment analysis community's efforts to develop new resources and methods, as well as to bring knowledge and experience from other disciplines that have been dealing with affect phenomena (e.g. Psychology, Sociology, etc.).

The aim of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA 2014) was to continue the line of the previous editions, bringing together researchers in Computational Linguistics working on Subjectivity and Sentiment Analysis and researchers working on interdisciplinary aspects of affect computation from text. Starting with 2013, WASSA has extended its scope and focus to Social Media phenomena and the impact of affect-related phenomena in this context.

WASSA 2014 was organized in conjunction to the 52nd Annual Meeting of the Association for Computational Linguistics (ACL 2014), on June 27, 2014, in Baltimore, Maryland, United States of America.

For this year's edition of WASSA, we received a total of 40 submissions, from universities and research centers all over the world, out of which 8 were accepted as long and another 14 as short papers. Each paper has been thoroughly reviewed by at least 2 members of the Program Committee. The accepted papers were all highly assessed by the reviewers, the best paper receiving an average punctuation (computed as an average of all criteria used to assess the papers) of 4.5 out of 5.

The main topics of the accepted papers are related to computational and cognitive modelling of affect, especially in Social Media - the creation and evaluation of resources for subjectivity, sentiment and emotion resources for Twitter mining, the development of semantic analysis-based methods for sentiment detection, argumentation and inference analysis, cross-lingual and multilingual resource creation and use, the detection of irony and sarcasm.

The invited talks reflected the multimodal nature of affect expressions and the strong connection between human affect-sensing mechanisms. At the same time, the talks drew our attention on the possible misuses of social media platforms that can bias opinion analysis, both to humans, as well as automatic systems. Finally, the talk by the organizers described the difficulties involved in porting research to the real-life application scenario.

This year's edition has again shown that the topics put forward to discussion by WASSA are of high interest to the research community and that the papers chosen to be debated in this forum bring an

important development to the SSA research area.

We would like to thank the ACL 2014 Organizers for the help and support at the different stages of the workshop organization process. We are also especially grateful to the Program Committee members and the external reviewers for the time and effort spent assessing the papers. We would like to extend our thanks to our invited speakers – Dr. Saif Mohammad and Dr. Myle Ott - for accepting to deliver the keynote talks.

Secondly, we would like to express our gratitude for the official endorsement we received from SIGNLL, the ACL Special Interest Group on Natural Language Learning, and SIGSEM, ACL's Special Interest Group on Computational Semantics.

We would like to express our gratitude to Yaniv Steiner, who created the WASSA logo and to the entire Europe Media Monitor team at the European Commission Joint Research Centre, for the technical support they provided.

Alexandra Balahur, Erik van der Goot, Ralf Steinberger and Andrés Montoyo
WASSA 2014 Chairs

Organizers:

Alexandra Balahur
European Commission Joint Research Centre
Institute for the Protection and Security of the Citizen

Erik van der Goot
European Commission Joint Research Centre
Institute for the Protection and Security of the Citizen

Ralf Steinberger
European Commission Joint Research Centre
Institute for the Protection and Security of the Citizen

Andrés Montoyo
University of Alicante
Department of Software and Computing Systems

Program Committee:

Nicoletta Calzolari, CNR Pisa (Italy)
Erik Cambria, University of Stirling (U.K.)
Fermin Cruz Mata, University of Seville (Spain)
Montse Cuadros - Vicomtech (Spain)
Leon Derczynski - University of Sheffield (U.K.)
Michael Gamon, Microsoft (U.S.A.)
Veronique Hoste, University of Ghent (Belgium)
Ruben Izquierdo Bevia - Vrije Universiteit Amsterdam (The Netherlands)
Isa Maks - Vrije Universiteit Amsterdam (The Netherlands)
Diana Maynard - University of Sheffield (U.K.)
Saif Mohammad, National Research Council (Canada)
Karo Moilanen, University of Oxford (U.K.)
Günter Neumann, DFKI (Germany)
Constantin Orasan, University of Wolverhampton (U.K.)
Viktor Pekar, University of Wolverhampton (U.K.)
Jose-Manuel Perea-Ortega - European Commission Joint Research Centre (Italy)
Paolo Rosso, Technical University of Valencia (Spain)
Josef Steinberger, Charles University Prague (The Czech Republic)
Mike Thelwall, University of Wolverhampton (U.K.)
Dan Tufis, RACAI (Romania)
Alfonso Ureña, University of Jaén (Spain)
Janyce Wiebe - University of Pittsburgh (U.S.A.)
Michael Wiegand, Saarland University (Germany)

Taras Zagibalov, Brantwatch (U.K.)

Invited Speakers:

Saif M. Mohammad, National Research Council Canada, Canada
Myle Ott, Facebook, U.S.A.

Table of Contents

Words: Evaluative, Emotional, Colourful, Musical!
Saif Mohammad 1

Robust Cross-Domain Sentiment Analysis for Low-Resource Languages
Jakob Elming, Barbara Plank and Dirk Hovy 2

An Investigation for Implicatures in Chinese : Implicatures in Chinese and in English are similar !
Lingjia Deng and Janyce Wiebe 8

Inducing Domain-specific Noun Polarity Guided by Domain-independent Polarity Preferences of Adjectives
Manfred Klenner, Michael Amsler and Nora Hollenstein 18

Aspect-Level Sentiment Analysis in Czech
Josef Steinberger, Tomáš Brychcín and Michal Konkol 24

Linguistic Models of Deceptive Opinion Spam
Myle Ott 31

Semantic Role Labeling of Emotions in Tweets
Saif Mohammad, Xiaodan Zhu and Joel Martin 32

An Impact Analysis of Features in a Classification Approach to Irony Detection in Product Reviews
Konstantin Buschmeier, Philipp Cimiano and Roman Klinger 42

Modelling Sarcasm in Twitter, a Novel Approach
Francesco Barbieri, Horacio Saggion and Francesco Ronzano 50

Emotive or Non-emotive: That is The Question
Michal Ptaszynski, Fumito Masui, Rafal Rzepka and Kenji Araki 59

Challenges in Creating a Multilingual Sentiment Analysis Application for Social Media Mining
Alexandra Balahur, Hristo Tanev and Erik van der Goot 66

Two-Step Model for Sentiment Lexicon Extraction from Twitter Streams
Ilia Chetviorkin and Natalia Loukachevitch 67

Linguistically Informed Tweet Categorization for Online Reputation Management
Gerard Lynch and Pádraig Cunningham 73

Credibility Adjusted Term Frequency: A Supervised Term Weighting Scheme for Sentiment Analysis and Text Classification
Yoon Kim and Owen Zhang 79

Opinion Mining and Topic Categorization with Novel Term Weighting
Tatiana Gasanova, Roman Sergienko, Shakhnaz Akhmedova, Eugene Semenkin and Wolfgang Minker 84

Sentiment classification of online political discussions: a comparison of a word-based and dependency-based method
Hugo Lewi Hammer, Per Erik Solberg and Lilja Øvrelid 90

Improving Agreement and Disagreement Identification in Online Discussions with A Socially-Tuned Sentiment Lexicon
Lu Wang and Claire Cardie 97

Lexical Acquisition for Opinion Inference: A Sense-Level Lexicon of Benefactive and Malefactive Events
Yoonjung Choi, Lingjia Deng and Janyce Wiebe 107

Dive deeper: Deep Semantics for Sentiment Analysis
Nikhilkumar Jadhav and Pushpak Bhattacharyya 113

Evaluating Sentiment Analysis Evaluation: A Case Study in Securities Trading
Siavash Kazemian, Shunan Zhao and Gerald Penn 119

Sentiment Classification on Polarity Reviews: An Empirical Study Using Rating-based Features
Dai Quoc Nguyen, Dat Quoc Nguyen, Thanh Vu and Son Bao Pham 128

Effect of Using Regression on Class Confidence Scores in Sentiment Analysis of Twitter Data
Itir Onal, Ali Mert Ertugrul and Ruken Cakici 136

A cognitive study of subjectivity extraction in sentiment annotation
Abhijit Mishra, Aditya Joshi and Pushpak Bhattacharyya 142

The Use of Text Similarity and Sentiment Analysis to Examine Rationales in the Large-Scale Online Deliberations
Wanting Mao, Lu Xiao and Robert Mercer 147

A Conceptual Framework for Inferring Implicatures
Janyce Wiebe and Lingjia Deng 154

Conference Program

Friday June 27, 2014

(8:30) Opening Remarks

(8:35) Invited talk: Dr. Saif Mohammad

Words: Evaluative, Emotional, Colourful, Musical!
Saif Mohammad

(9:10) Session 1: Cross-domain and Multilingual Sentiment Analysis

9:10 *Robust Cross-Domain Sentiment Analysis for Low-Resource Languages*
Jakob Elming, Barbara Plank and Dirk Hovy

9:35 *An Investigation for Implicatures in Chinese : Implicatures in Chinese and in English are similar !*
Lingjia Deng and Janyce Wiebe

10:00 *Inducing Domain-specific Noun Polarity Guided by Domain-independent Polarity Preferences of Adjectives*
Manfred Klenner, Michael Amsler and Nora Hollenstein

10:15 *Aspect-Level Sentiment Analysis in Czech*
Josef Steinberger, Tomáš Brychcín and Michal Konkol

(10:30) Break

(10:50) Invited talk: Dr. Myle Ott

10:50 *Linguistic Models of Deceptive Opinion Spam*
Myle Ott

Friday June 27, 2014 (continued)

(11:25) Session 2: Emotion, Irony and Sarcasm Classification

11:25 *Semantic Role Labeling of Emotions in Tweets*
Saif Mohammad, Xiaodan Zhu and Joel Martin

11:50 *An Impact Analysis of Features in a Classification Approach to Irony Detection in Product Reviews*
Konstantin Buschmeier, Philipp Cimiano and Roman Klinger

12:15 *Modelling Sarcasm in Twitter, a Novel Approach*
Francesco Barbieri, Horacio Saggion and Francesco Ronzano

12:30 *Emotive or Non-emotive: That is The Question*
Michal Ptaszynski, Fumito Masui, Rafal Rzepka and Kenji Araki

(12:45) Lunch Break

(14:00) Demo talk: Dr. Alexandra Balahur

14:00 *Challenges in Creating a Multilingual Sentiment Analysis Application for Social Media Mining*
Alexandra Balahur, Hristo Tanev and Erik van der Goot

(14:30) Session 3: Lexical Acquisition and Feature Weighting for Sentiment Analysis

14:30 *Two-Step Model for Sentiment Lexicon Extraction from Twitter Streams*
Ilia Chetviorkin and Natalia Loukachevitch

14:45 *Linguistically Informed Tweet Categorization for Online Reputation Management*
Gerard Lynch and Pádraig Cunningham

15:00 *Credibility Adjusted Term Frequency: A Supervised Term Weighting Scheme for Sentiment Analysis and Text Classification*
Yoon Kim and Owen Zhang

15:15 *Opinion Mining and Topic Categorization with Novel Term Weighting*
Tatiana Gasanova, Roman Sergienko, Shakhnaz Akhmedova, Eugene Semenkin and Wolfgang Minker

Friday June 27, 2014 (continued)

(15:30) Break

(16:00) Session 4: Sentiment Analysis from Discourse and Dialogues

16:00 *Sentiment classification of online political discussions: a comparison of a word-based and dependency-based method*
Hugo Lewi Hammer, Per Erik Solberg and Lilja Øvrelid

16:25 *Improving Agreement and Disagreement Identification in Online Discussions with A Socially-Tuned Sentiment Lexicon*
Lu Wang and Claire Cardie

16:50 *Lexical Acquisition for Opinion Inference: A Sense-Level Lexicon of Benefactive and Malefactive Events*
Yoonjung Choi, Lingjia Deng and Janyce Wiebe

17:05 *Dive deeper: Deep Semantics for Sentiment Analysis*
Nikhilkumar Jadhav and Pushpak Bhattacharyya

(17:20) Break

(17:30) Session 5: Sentiment Analysis Evaluation. Going Beyond Current Sentiment Analysis Approaches

17:30 *Evaluating Sentiment Analysis Evaluation: A Case Study in Securities Trading*
Siavash Kazemian, Shunan Zhao and Gerald Penn

17:55 *Sentiment Classification on Polarity Reviews: An Empirical Study Using Rating-based Features*
Dai Quoc Nguyen, Dat Quoc Nguyen, Thanh Vu and Son Bao Pham

18:20 *Effect of Using Regression on Class Confidence Scores in Sentiment Analysis of Twitter Data*
Itir Onal, Ali Mert Ertugrul and Ruken Cakici

18:35 *A cognitive study of subjectivity extraction in sentiment annotation*
Abhijit Mishra, Aditya Joshi and Pushpak Bhattacharyya

18:50 *The Use of Text Similarity and Sentiment Analysis to Examine Rationales in the Large-Scale Online Deliberations*
Wanting Mao, Lu Xiao and Robert Mercer

19:05 *A Conceptual Framework for Inferring Implicatures*
Janyce Wiebe and Lingjia Deng

Friday June 27, 2014 (continued)

(19:20) Closing remarks