2nd Workshop on Language in Social Media 2012

Held at the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2012)

Montreal, Quebec, Canada
3-8 June 2012

ISBN: 978-1-62276-549-2

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2012) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2013)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

Analyzing Urdu Social Media for Sentiments using Transfer Learning with Controlled Translations
Smruthi Mukund and Rohini Srihari 1

Detecting Distressed and Non-distressed Affect States in Short Forum Texts
Michael Thaul Lehrman, Cecilia Ovesdotter Alm and Ruben A. Proano 9

Detecting Hate Speech on the World Wide Web
William Warner and Julia Hirschberg 19

A Demographic Analysis of Online Sentiment during Hurricane Irene
Benjamin Mandel, Aron Culotta, John Boulahanis, Danielle Stark, Bonnie Lewis and Jeremy Rodrigue 27

Detecting Influencers in Written Online Conversations
Or Biran, Sara Rosenthal, Jacob Andreas, Kathleen McKeown and Owen Rambow 37

Re-tweeting from a linguistic perspective
Aobo Wang, Tao Chen and Min-Yen Kan 46

Robust kaomoji detection in Twitter
Steven Bedrick, Russell Beckley, Brian Roark and Richard Sproat 56

Language Identification for Creating Language-Specific Twitter Collections
Shane Bergsma, Paul McNamee, Mossaab Bagdouri, Clayton Fink and Theresa Wilson 65

Processing Informal, Romanized Pakistani Text Messages
Ann Irvine, Jonathan Weese and Chris Callison-Burch 75