

54th Annual Meeting of the Association for Computational Linguistics (ACL 2016)

Student Research Workshop

Berlin, Germany
7 – 12 August 2016

ISBN: 978-1-5108-2760-8

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© (2016) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2016)

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: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>Controlled and Balanced Dataset for Japanese Lexical Simplification</i> Tomonori Kodaira, Tomoyuki Kajiwara and Mamoru Komachi	1
<i>Dependency Forest based Word Alignment</i> Hitoshi Otsuki, Chenhui Chu, Toshiaki Nakazawa and Sadao Kurohashi	8
<i>Identifying Potential Adverse Drug Events in Tweets Using Bootstrapped Lexicons</i> Eric Benzschawel	15
<i>Generating Natural Language Descriptions for Semantic Representations of Human Brain Activity</i> Eri Matsuo, Ichiro Kobayashi, Shinji Nishimoto, Satoshi Nishida and Hideki Asoh	22
<i>Improving Twitter Community Detection through Contextual Sentiment Analysis</i> Alron Jan Lam	30
<i>Significance of an Accurate Sandhi-Splitter in Shallow Parsing of Dravidian Languages</i> Devadath V V and Dipti Misra Sharma	37
<i>Improving Topic Model Clustering of Newspaper Comments for Summarisation</i> Clare Llewellyn, Claire Grover and Jon Oberlander	43
<i>Arabizi Identification in Twitter Data</i> Taha Tobaili	51
<i>Robust Co-occurrence Quantification for Lexical Distributional Semantics</i> Dmitrijs Milajevs, Mehrnoosh Sadrzadeh and Matthew Purver	58
<i>Singleton Detection using Word Embeddings and Neural Networks</i> Hessel Haagsma	65
<i>A Dataset for Joint Noun-Noun Compound Bracketing and Interpretation</i> Murhaf Fares	72
<i>An Investigation on The Effectiveness of Employing Topic Modeling Techniques to Provide Topic Awareness For Conversational Agents</i> Omid Moradiannasab	80
<i>Improving Dependency Parsing Using Sentence Clause Charts</i> Vincent Križ and Barbora Hladka	86
<i>Graph- and surface-level sentence chunking</i> Ewa Muszyńska	93
<i>From Extractive to Abstractive Summarization: A Journey</i> Parth Mehta	100
<i>Putting Sarcasm Detection into Context: The Effects of Class Imbalance and Manual Labelling on Supervised Machine Classification of Twitter Conversations</i> Gavin Abercrombie and Dirk Hovy	107
<i>Unsupervised Authorial Clustering Based on Syntactic Structure</i> Alon Daks and Aidan Clark	114

<i>Suggestion Mining from Opinionated Text</i>	
Sapna Negi	119
<i>An Efficient Cross-lingual Model for Sentence Classification Using Convolutional Neural Network</i>	
Yandi Xia, Zhongyu Wei and Yang Liu	126
<i>QA-It: Classifying Non-Referential It for Question Answer Pairs</i>	
Timothy Lee, Alex Lutz and Jinho D. Choi	132
<i>Building a Corpus for Japanese Wikification with Fine-Grained Entity Classes</i>	
Davaajav Jargalsaikhan, Naoaki Okazaki, Koji Matsuda and Kentaro Inui	138
<i>A Personalized Markov Clustering and Deep Learning Approach for Arabic Text Categorization</i>	
Vasu Jindal	145