Australasian Language Technology Association Workshop (ALTA 2017)

Brisbane, Australia
6 – 8 December 2017

Editors:

Jojo Sze-Meng!Wong
Gholamreza Haffari

ISBN: 978-1-5108-5939-5

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© (2017) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2018)

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

Contents

Invited talks **1**

Tutorials **3**

Long papers **5**

Stock Market Prediction with Deep Learning: A Character-based Neural Language Model for Event-based Trading
Leonardo Dos Santos Pinheiro and Mark Dras **6**

Improving End-to-End Memory Networks with Unified Weight Tying
Fei Liu, Trevor Cohn and Timothy Baldwin **16**

Joint Sentence-Document Model for Manifesto Text Analysis
Shivashankar Subramanian, Trevor Cohn, Timothy Baldwin and Julian Brooke **25**

Leveraging linguistic resources for improving neural text classification
Ming Liu, Gholamreza Haffari, Wray Buntine and Michelle Ananda-Rajah **34**

A Hybrid Model for Quality Assessment of Wikipedia Articles
Aili Shen, Jianzhong Qi and Timothy Baldwin **43**

Phonemic Transcription of Low-Resource Tonal Languages
Oliver Adams, Trevor Cohn, Graham Neubig and Alexis Michaud **53**

A Comparative Study of Two Statistical Modelling Approaches for Estimating Multivariate Likelihood Ratios in Forensic Voice Comparison
Shunichi Ishihara **61**

Automatic Negation and Speculation Detection in Veterinary Clinical Text
Katharine Cheng, Timothy Baldwin and Karin Verspoor **70**

Medication and Adverse Event Extraction from Noisy Text
Xiang Dai, Sarvnaz Karimi and Cecile Paris **79**

Incremental Knowledge Acquisition Approach for Information Extraction on both Semi-structured and Unstructured Text from the Open Domain Web
Maria Myunghee Kim **88**

Short papers **97**

On Extending Neural Networks with Loss Ensembles for Text Classification
Hamideh Hajiabadi, Diego Mollá-Alliod and Reza Monsefi **98**

Towards the Use of Deep Reinforcement Learning with Global Policy For Query-based Extractive Summarisation
Diego Mollá-Alliod **103**

From Word Segmentation to POS Tagging for Vietnamese
Dat Quoc Nguyen, Thanh Vu, Dai Quoc Nguyen, Mark Dras and Mark Johnson **108**

ALTA Shared Task papers **114**

Overview of the 2017 ALTA Shared Task: Correcting OCR Errors
Diego Mollá-Alliod and Steve Cassidy **115**

OCR Post-Processing Text Correction using Simulated Annealing (OPTeCA)
Gitansh Khirbat **119**

SuperOCR for ALTA 2017 Shared Task
Yufei Wang **124**