NAACL HLT 2018

The 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies

Proceedings of the Conference
Volume 3 (Industry Papers)

June 1-June 6, 2018
New Orleans, Louisiana

©2018 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 978-1-948087-30-8

Introduction

It is our pleasure to welcome you to the inaugural Industry Track in the *ACL family of conferences at NAACL 2018.

The idea of organizing an industry track stemmed from the challenging issues encountered while attempting to apply state-of-the-art techniques to real-world language problems. As those who have attempted these problems know, practical applications are rarely as well defined as in laboratory settings and the data never as clean. In addition, there may be practical constraints such as computational requirements, processing speed, memory footprint, latency requirements, ease of updating a deployed solution that need to be balanced judiciously, and capability to be embedded as part of a larger system. The NAACL 2018 Industry Track was born out of the desire to provide a forum for researchers, engineers and application developers who are experienced in these issues to exchange ideas, share results and discuss use-cases of successful development and deployments of language technologies in real-world settings.

Although we thought that the time is ripe in the NLP field for such a forum, and hoped that the community will embrace the opportunity to share their experience with others, it was nonetheless a guessing game as to the amount of interest the track would actually generate. As submissions drew to a close in late February, we were happy to report that we received 91 submissions, far exceeding our expectations (which led to last-minute scrambling to recruit more reviewers, but we're not complaining!). Six of the papers were desk rejects due to non-conformance with submission requirements, and the remaining 85 papers were reviewed by 65 reviewers. We accepted 28 papers – an acceptance rate of 32.9% (one paper was subsequently withdrawn after acceptance) of which 19 papers will be presented in oral sessions that run as a parallel track during the main conference, and 8 papers will be presented during poster sessions. Of course, none of this would have been possible without the participation of authors and reviewers, and we would like to convey our heartfelt "thank you" to all the authors for submitting papers and the reviewers for their efforts in the paper selection process.

We analyzed our submissions along a couple of dimensions and would like to share some interesting statistics. First we looked at the submissions with respect to the distribution of author affiliations. As one would expect, the industry track focuses on problems that manifest themselves more readily in industry than in academia. Indeed, of the 85 papers reviewed, 55 papers are authored by researchers/engineers in industry laboratories. The particularly encouraging statistic, however, is that 25 papers are the results of collaboration between those in industry and academia. it would be interesting to track these statistics in future years to see if the collaboration increases as the field continues to mature. The second dimension we analyzed is the geographic distribution of authors by contact author. This being a NAACL conference, it is no surprise that 62% of the papers came from North America. We are pleased with the participation of authors from other regions, including 22% from Europe, 14% from Asia, and 1% from Africa.

In addition to paper presentations, we will have two plenary keynote speeches. For the keynote speeches, we aimed to feature researchers who also have first hand experience applying research results to practical applications. To that end, we are honored to have two illustrious members of NLP community – Daniel Marcu, who co-founded Language Weaver more than 15 years ago and is now the director of MT/NLP at Amazon, and Mari Ostendorf, professor at the University of Washington, who led a team of students to build a social bot that won the 2017 Alexa Prize competition. We are confident that their experiences would be of immense interest to the larger NLP community.

Another highlight of the industry track includes two panel discussions on topics of increasing importance in the community. The first panel, "Careers in Industry", moderated by Philip Resnik, professor at University of Maryland, is primarily geared toward students and recent graduates who are exploring careers in industry versus academia. The panel will feature experienced professionals who have worked

in both environments to share their experience and offer advice, based on questions gathered from the *ACL community earlier this year . The second panel, "Ethics in NLP", will be moderated by Dirk Hovy, professor at Università Bocconi, and will focus on raising awareness of the emerging issues of biases present in NLP/AI solutions, the social implications of such biases, and what we, as NLP practitioners, can do to reduce them.

With the overwhelming response to the call for papers, the language community has unambiguously endorsed the relevance of the Industry track in the milieu of annual conferences. As organizers, we have attempted to amplify this endorsement by bringing to the participants a invigorating technical program. We hope through your engaging discussions and active participation during the sessions, you will unanimously support and nurture the concept of an Industry track in NLP conferences over the years to come.

Srinivas Bangalore (Interactions Labs)
Jennifer Chu-Carroll (Elemental Cognition)
Yunyao Li (IBM Research - Almaden)

NAACL 2018 Industry Track Co-Chairs

Organizers:

Srinivas Bangalore, Interactions Labs
Jennifer Chu-Carroll, Elemental Cognition
Yunyao Li, IBM Research - Almaden

Program Committee:

Apoorv Agarwal,Columbia University (United States)
Alan Akbik,Zalando Research (Germany)
Miguel Ballesteros,IBM Research (United States)
Loïc Barrault,LIUM, Le Mans University (France)
Akash Bharadwaj,Carnegie Mellon University (United States)
Ciprian Chelba,Google (United States)
John Chen,Interactions LLC (United States)
Minhua Chen,Interactions LLC (United States)
Laura Chiticariu,IBM Research - Almaden (United States)
Deborah Dahl,Conversational Technologies (United States)
Lingjia Deng,Bloomberg (United States)
Giuseppe Di Fabbrizio,Rakuten Institute of Technology (United States)
Christine Doran,Self (United States)
Kavita Ganesan,uiuc (United States)
Anna Lisa Gentile,IBM Research Almaden (United States)
Yufan Guo,University of Cambridge (United Kingdom)
Bo Han,RedMarker/Kaplan Inc. (Australia)
Peter Heeman,OHSU / CSLU (United States)
John Henderson,MITRE (United States)
Lynette Hirschman,MITRE (United States)
Dirk Hovy,Bocconi University (Italy)
Badrinath Jayakumar,Interactions LLC (United States)
Michael Johnston,Interactions Corporation (United States)
Adi Kalyanpur,Elemental Cognition (United States)
Doo Soon Kim,Adobe Research (United States)
Kevin Knight,USC/ISI (United States)
Gourab Kundu,IBM T.J. Watson Research Center (United States)
Gakuto Kurata,IBM Research (Japan)
Yocheved Levitan,Interactions LLC (United States)
Bing Liu,University of Illinois at Chicago (United States)
Alexander Loeser,Beuth-University of Applied Sciences Berlin (Germany)
Stephanie Lukin,US Army Research Laboratory (United States)
Xiaoqiang Luo,Google (United States)
Yuji Matsumoto,Nara Institute of Science and Technology (Japan)
Irina Matveeva,NexLP, Inc (United States)
David McClosky,Google (United States)
Mahnoosh Mehrabani,Interactions LLC (United States)
Marie Meteer,Brandeis University (United States)
Margaret Mitchell,Google Research and Machine Intelligence (United States)
Helena Moniz,INESC-ID, FLUL/Unbabel (Portugal)
Stefano Pacifico,Elemental Cognition (United States)

Shimei Pan,UMBC (United States)
Youngja Park,IBM T. J. Watson Research Center (United States)
Siddharth Patwardhan,Apple (United States)
Alexandros Potamianos,National Technical University of Athens (Greece)
John Prager,IBM Research (United States)
Rashmi Prasad,Interactions Corporation (United States)
Jerry Quinn,IBM Research (United States)
Preethi Raghavan,IBM Research TJ Watson (United States)
Owen Rambow,Columbia University (United States)
Ehud Reiter,University of Aberdeen (United Kingdom)
Giuseppe Riccardi,University of Trento (Italy)
Stefan Riezler,Heidelberg University (Germany)
Brian Roark,Google Inc. (United States)
Nicholas Ruiz,Interactions, LLC (United States)
Frank Seide,Microsoft Research (United States)
Ethan Selfridge,Interactions Corp (United States)
Chaitanya Shivade,IBM Research (United States)
Svetlana Stoyanchev,Interactions Corporation (United States)
Fabian Suchanek,Telecom ParisTech University (France)
Aniruddha Tammewar,Interactions Corporation (India)
Isabel Trancoso,INESC-ID / IST (Portugal)
Ling Tsou,SDL (United States)
Gokhan Tur,Uber (United States)
Jason D Williams,Microsoft Research (United States)
Fei Xia,University of Washington (United States)
Luke Zettlemoyer,University of Washington (United States)

Keynote Speaker:

Daniel Marcu, Amazon
Mari Ostendorf, University of Washington

Panel Moderator:

Philip Resnik, University of Maryland - College Park
Dirk Hovy, Università Bocconi

Table of Contents

Scalable Wide and Deep Learning for Computer Assisted Coding
Marilisa Amoia, Frank Diehl, Jesus Gimenez, Joel Pinto, Raphael Schumann, Fabian Stemmer, Paul Vozila and Yi Zhang 1

Neural Network based Extreme Classification and Similarity Models for Product Matching
Kashif Shah, Selcuk Kopru and Jean David Ruvini 8

A Scalable Neural Shortlisting-Reranking Approach for Large-Scale Domain Classification in Natural Language Understanding
Young-Bum Kim, Dongchan Kim, Joo-Kyung Kim and Ruhi Sarikaya 16

What we need to learn if we want to do and not just talk
Rashmi Gangadharaiah, Balakrishnan Narayanaswamy and Charles Elkan 25

Data Collection for Dialogue System: A Startup Perspective
Yiping Kang, Yunqi Zhang, Jonathan K. Kummerfeld, Lingjia Tang and Jason Mars 33

Bootstrapping a Neural Conversational Agent with Dialogue Self-Play, Crowdsourcing and On-Line Reinforcement Learning
Pararth Shah, Dilek Hakkani-Tur, Bing Liu and Gokhan Tur 41

Quality Estimation for Automatically Generated Titles of eCommerce Browse Pages
Nicola Ueffing, José G. C. de Souza and Gregor Leusch 52

Atypical Inputs in Educational Applications
Su-Youn Yoon, Aoife Cahill, Anastassia Loukina, Klaus Zechner, Brian Riordan and Nitin Madnani 60

Using Aspect Extraction Approaches to Generate Review Summaries and User Profiles
Christopher Mitcheltree, Veronica Wharton and Avneesh Saluja 68

SystemT: Declarative Text Understanding for Enterprise
laura chiticariu, Marina Danilevsky, Yunyao Li, Frederick Reiss and Huaiyu Zhu 76

Construction of the Literature Graph in Semantic Scholar
Waleed Ammar, Dirk Groeneveld, Chandra Bhagavatula, Iz Beltagy, Miles Crawford, Doug Downey, Jason Dunkelberger, Ahmed Elgohary, Sergey Feldman, Vu Ha, Rodney Kinney, Sebastian Kohlmeier, Kyle Lo, Tyler Murray, Hsu-Han Ooi, Matthew Peters, Joanna Power, Sam Skjonsberg, Lucy Wang, Chris Willhelm, Zheng Yuan, Madeleine van Zuylen and oren etzioni 84

Can Neural Machine Translation be Improved with User Feedback?
Julia Kreutzer, Shahram Khadivi, Evgeny Matusov and Stefan Riezler 92

Accelerating NMT Batched Beam Decoding with LMBR Posteriors for Deployment
Gonzalo Iglesias, William Tambellini, Adrià de Gispert, Eva Hasler and Bill Byrne 106

Pieces of Eight: 8-bit Neural Machine Translation
Jerry Quinn and Miguel Ballesteros 114

From dictations to clinical reports using machine translation
Gregory Finley, Wael Salloum, Najmeh Sadoughi, Erik Edwards, Amanda Robinson, Nico Axtmann, Michael Brenndoerfer, Mark Miller and David Suendermann-Oeft 121

Benchmarks and models for entity-oriented polarity detection
Lidia Pivovarova, Arto Klami and Roman Yangarber 129

Selecting Machine-Translated Data for Quick Bootstrapping of a Natural Language Understanding System
Judith Gaspers, Penny Karanasou and Rajen Chatterjee 137

Fast and Scalable Expansion of Natural Language Understanding Functionality for Intelligent Agents
Anuj Kumar Goyal, Angeliki Metallinou and Spyros Matsoukas 145

Bag of Experts Architectures for Model Reuse in Conversational Language Understanding
Rahul Jha, Alex Marin, Suvamsh Shivaprasad and Imed Zitouni 153

Multi-lingual neural title generation for e-Commerce browse pages
Prashant Mathur, Nicola Ueffing and Gregor Leusch 162

A Novel Approach to Part Name Discovery in Noisy Text
Nobal Bikram Niraula, Daniel Whyatt and Anne Kao 170

The Alexa Meaning Representation Language
Thomas Kollar, Danielle Berry, Lauren Stuart, Karolina Owczarzak, Tagyoung Chung, Lambert Mathias, Michael Kayser, Bradford Snow and Spyros Matsoukas 177

Practical Application of Domain Dependent Confidence Measurement for Spoken Language Understanding Systems
Mahnoosh Mehrabani, David Thomson and Benjamin Stern 185

Prediction for the Newsroom: Which Articles Will Get the Most Comments?
Carl Ambroselli, Julian Risch, Ralf Krestel and Andreas Loos 193

Demand-Weighted Completeness Prediction for a Knowledge Base
Andrew Hopkinson, Amit Gurdasani, Dave Palfrey and Arpit Mittal 200

Personalized neural language models for real-world query auto completion
Nicolas Fiorini and Zhiyong Lu 208

Document-based Recommender System for Job Postings using Dense Representations
Ahmed Elsafty, Martin Riedl and Chris Biemann 216

Conference Program

Saturday, June 2, 2018

10:30–11:30 Machine Learning – Classification

10:30–10:47 *Scalable Wide and Deep Learning for Computer Assisted Coding*
Marilisa Amoia, Frank Diehl, Jesus Gimenez, Joel Pinto, Raphael Schumann, Fabian Stemmer, Paul Vozila and Yi Zhang

10:48–11:05 *Neural Network based Extreme Classification and Similarity Models for Product Matching*
Kashif Shah, Selcuk Kopru and Jean David Ruvini

11:06–11:23 *A Scalable Neural Shortlisting-Reranking Approach for Large-Scale Domain Classification in Natural Language Understanding*
Young-Bum Kim, Dongchan Kim, Joo-Kyung Kim and Ruhi Sarikaya

11:30–12:30 Dialog

11:30–11:47 *What we need to learn if we want to do and not just talk*
Rashmi Gangadharaiah, Balakrishnan Narayanaswamy and Charles Elkan

11:48–12:05 *Data Collection for Dialogue System: A Startup Perspective*
Yiping Kang, Yunqi Zhang, Jonathan K. Kummerfeld, Lingjia Tang and Jason Mars

12:05–12:23 *Bootstrapping a Neural Conversational Agent with Dialogue Self-Play, Crowdsourcing and On-Line Reinforcement Learning*
Pararth Shah, Dilek Hakkani-Tur, Bing Liu and Gokhan Tur

Saturday, June 2, 2018 (continued)

10:30am–12:00pm Session Posters and Demos: Generation

Quality Estimation for Automatically Generated Titles of eCommerce Browse Pages
Nicola Ueffing, José G. C. de Souza and Gregor Leusch

10:30–12:00 Session Posters and Demos: NLP Applications -

Atypical Inputs in Educational Applications
Su-Youn Yoon, Aoife Cahill, Anastassia Loukina, Klaus Zechner, Brian Riordan and Nitin Madnani

Using Aspect Extraction Approaches to Generate Review Summaries and User Profiles
Christopher Mitcheltree, Veronica Wharton and Avneesh Saluja

SystemT: Declarative Text Understanding for Enterprise
laura chiticariu, Marina Danilevsky, Yunyao Li, Frederick Reiss and Huaiyu Zhu

Construction of the Literature Graph in Semantic Scholar
Waleed Ammar, Dirk Groeneveld, Chandra Bhagavatula, Iz Beltagy, Miles Crawford, Doug Downey, Jason Dunkelberger, Ahmed Elgohary, Sergey Feldman, Vu Ha, Rodney Kinney, Sebastian Kohlmeier, Kyle Lo, Tyler Murray, Hsu-Han Ooi, Matthew Peters, Joanna Power, Sam Skjonsberg, Lucy Wang, Chris Willhelm, Zheng Yuan, Madeleine van Zuylen and oren etzioni

14:00–15:00 Industry Track Keynote 1

14:00–15:00 *Building a Socialbot: Lessons Learned from 10M Conversations*
Mari Ostendorf

Saturday, June 2, 2018 (continued)

15:30–17:00 Careers in the Industry Panel

17:00–18:30 Machine Translation

17:00–17:17 *Can Neural Machine Translation be Improved with User Feedback?*
Julia Kreutzer, Shahram Khadivi, Evgeny Matusov and Stefan Riezler

17:18–17:35 *Accelerating NMT Batched Beam Decoding with LMBR Posteriors for Deployment*
Gonzalo Iglesias, William Tambellini, Adrià de Gispert, Eva Hasler and Bill Byrne

17:36–17:53 *Pieces of Eight: 8-bit Neural Machine Translation*
Jerry Quinn and Miguel Ballesteros

17:54–18:12 *From dictations to clinical reports using machine translation*
Gregory Finley, Wael Salloum, Najmeh Sadoughi, Erik Edwards, Amanda Robinson, Nico Axtmann, Michael Brenndoerfer, Mark Miller and David Suendermann-Oeft

15:30–17:00 Session Posters and Demos: Sentiment Analysis

Benchmarks and models for entity-oriented polarity detection
Lidia Pivovarova, Arto Klami and Roman Yangarber

Sunday, June 3, 2018

Sunday, June 3, 2018 (continued)

10:30–11:30 Machine Learning – Domain Adaptation

10:30–10:47 *Selecting Machine-Translated Data for Quick Bootstrapping of a Natural Language Understanding System*
Judith Gaspers, Penny Karanasou and Rajen Chatterjee

10:48–11:05 *Fast and Scalable Expansion of Natural Language Understanding Functionality for Intelligent Agents*
Anuj Kumar Goyal, Angeliki Metallinou and Spyros Matsoukas

11:05–11:22 *Bag of Experts Architectures for Model Reuse in Conversational Language Understanding*
Rahul Jha, Alex Marin, Suvamsh Shivaprasad and Imed Zitouni

11:30–12:30 Named Entity and Language Generation

11:30–11:47 *Multi-lingual neural title generation for e-Commerce browse pages*
Prashant Mathur, Nicola Ueffing and Gregor Leusch

11:48–12:05 *A Novel Approach to Part Name Discovery in Noisy Text*
Nobal Bikram Niraula, Daniel Whyatt and Anne Kao

12:06–12:23 *The Alexa Meaning Representation Language*
Thomas Kollar, Danielle Berry, Lauren Stuart, Karolina Owczarzak, Tagyoung Chung, Lambert Mathias, Michael Kayser, Bradford Snow and Spyros Matsoukas

Sunday, June 3, 2018 (continued)

14:00–15:00 Industry Track Keynote 2

14:00–15:00 *Building innovative startups, products, and services – personal insights*
Daniel Marcu

15:30–17:00 Ethics in NLP Panel

15:30–17:00 Session Posters and Demos: Machine Learning

Practical Application of Domain Dependent Confidence Measurement for Spoken Language Understanding Systems
Mahnoosh Mehrabani, David Thomson and Benjamin Stern

Monday, June 4, 2018

11:30–12:30 Machine Learning - Prediction

11:30–11:47 *Prediction for the Newsroom: Which Articles Will Get the Most Comments?*
Carl Ambroselli, Julian Risch, Ralf Krestel and Andreas Loos

11:48–12:05 *Demand-Weighted Completeness Prediction for a Knowledge Base*
Andrew Hopkinson, Amit Gurdasani, Dave Palfrey and Arpit Mittal

12:05–12:22 *Personalized neural language models for real-world query auto completion*
Nicolas Fiorini and Zhiyong Lu

10:30–12:00 Session Posters and Demos: Social Media and Computational Social Science

Document-based Recommender System for Job Postings using Dense Representations
Ahmed Elsafty, Martin Riedl and Chris Biemann