

EMNLP-IJCNLP 2019

**Natural Language Processing for Internet Freedom:
Censorship, Disinformation, and Propaganda
NLP4IF 2019**

Proceedings of the Workshop

November 4, 2019
Hong Kong

©2019 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 978-1-950737-89-5

Preface

Welcome to the second edition of the Workshop on Natural Language Processing for Internet Freedom (NLP4IF 2019). This year, we focused on censorship, disinformation, and propaganda.

We further featured a shared task on the identification of propaganda in news articles. The task included two subtasks with different levels of complexity. Given a news article, the FLC subtask (fragment-level classification) asked for the identification of the propagandistic text fragments and also for the prediction of the specific propaganda technique used in this fragment (18-way classification task). The SLC subtask (sentence-level classification) is a binary classification task, which asked to detect the sentences that contain propaganda. A total of 39 teams submitted runs; 21 teams participated in the FLC subtask and 35 teams took part in the SLC subtask. Fourteen participants submitted a system description paper, which include models based on a wide range of learning models (e.g., neural networks, logistic regression) and representations (e.g., manually-engineered features, distributional representations).

We accepted a total of 24 papers: 10 for the regular track and 14 for the shared task. We are excited that the workshop includes a diverse set of topics: rumor and trolls detection, censorship and controversy, fake news vs. satire, uncovering propaganda and abusive language identification.

We are also thrilled to be able to bring an invited speaker, Elissa Redmiles from Princeton University and Microsoft Research, with a talk on measuring human perception to defend democracy, exploring a specific attack on the freedom of U.S. elections – the IRA Facebook advertisements, which successfully influenced people and avoided detection – and a defense against propaganda, which uses human perceptions to defend against the very propaganda that aims to influence those perceptions.

Last but not least, we would like to thank the program committee and the shared task participants for their help with reviewing the papers, and with advertising the workshop.

The NLP4IF 2019 Organizers:

Anna Feldman
Giovanni Da San Martino
Alberto Barrón-Cedeño
Chris Brew
Chris Leberknight
Preslav Nakov

Organizers

Anna Feldman, Montclair University (USA)
Giovanni Da San Martino, Qatar Computing Research Institute, HBKU (Qatar)
Alberto Barrón-Cedeño, Università di Bologna (Italy)
Chris Brew, Facebook (USA)
Chris Leberknight, Monclair University (USA)
Preslav Nakov, Qatar Computing Research Institute, HBKU (Qatar)

Program Committee

Banu Akdenizli, Northwestern University (Qatar)
Dyaa Albakour, Signal Media (UK)
Jisun An, Qatar Computing Research Institute, HBKU (Qatar)
Jed Crandall, University of New Mexico (USA)
Kareem Darwish, Qatar Computing Research Institute, HBKU (Qatar)
Anjalie Field, Carnegie Mellon University (USA)
Gianmarco De Francisci Morales, ISI Foundation (Italy)
Julio Gonzalo, UNED (Spain)
Heng Ji, Rensselaer Polytechnic Institute (USA)
Jeffrey Knockell, The Citizen Lab, University of Toronto (Canada)
Haewoon Kwak, Qatar Computing Research Institute, HBKU (Qatar)
Miguel Martinez, Signal Media, (UK)
Ivan Meza, National Autonomous University of Mexico (Mexico)
Rada Mihalcea, University Michigan, Ann Arbor (USA)
Prateek Mittal, Princeton University (USA)
Alessandro Moschitti, Amazon (USA)
Veronica Perez, University of Michigan (USA)
Hannah Rashkin, University of Washington (USA)
Paolo Rosso, Technical University of Valencia (Spain)
Anna Rumshisky, University of Massachusetts, Lowell (USA)
Mahmood Sharif, Carnegie Mellon University (USA)
Thamar Solorio, University of Houston (USA)
Benno Stein, Bauhaus University Weimar (Germany)
Denis Stukal, New York University (USA)
Yulia Tsvetkov, Carnegie Mellon University (USA)
Svetlana Volkova, Pacific Northwest National Laboratory (USA)
Henning Wachsmuth, University of Paderborn (Germany)
Brook Wu, New Jersey Institute of Technology (USA)

Shared Task Reviewers

Ali Fadel, Jordan University of Science and Technology (Jordan)
André Ferreira Cruz, University of Porto (Portugal)
George-Alexandru Vlad, University Politehnica of Bucharest (Romania)
Gil Rocha, University of Porto (Portugal)
Harish Tayyar Madabushi, University of Birmingham (England)
Henrique Lopes Cardoso, FEUP / LIACC (Portugal)
Ibraheem Tuffaha, Jordan University of Science and Technology (Jordan)

Jinfen Li, Syracuse University, New York (USA)
Kartik Aggarwal, NSIT, New Delhi (India)
Mahmoud Al-Ayyoub, Jordan University of Science and Technology (Jordan)
Malak Abdullah, Jordan University of Science and Technology (Jordan)
Mehdi Ghanimifard, University of Gothenburg (Sweden)
Norman Mapes, Louisiana Tech University, (USA)
Pankaj Gupta, University of Munich (LMU) and Siemens (Germany)
Shehel Yoosuf, Hamad Bin Khalifa University (Qatar)
Tariq Alhindi, Columbia University (USA)
Wenjun Hou, China Agricultural University, Beijing (China)
Yiqing Hua, Cornell University (USA)

Secondary Reviewers

Adam Ek, University of Gothenburg (Sweden)
Usama Yaseen, University of Munich and Siemens (Germany)

Invited Speaker

Elissa Redmiles, Princeton and Microsoft Research (USA)

Table of Contents

<i>Assessing Post Deletion in Sina Weibo: Multi-modal Classification of Hot Topics</i> Meisam Navaki Arefi, Rajkumar Pandi, Michael Carl Tschantz, Jedidiah R. Crandall, King-wa Fu, Dahlia Qiu Shi and Miao Sha	1
<i>Detecting context abusiveness using hierarchical deep learning</i> Ju-Hyoung Lee, Jun-U Park, Jeong-Won Cha and Yo-Sub Han	10
<i>How Many Users Are Enough? Exploring Semi-Supervision and Stylometric Features to Uncover a Russian Troll Farm</i> Nayeema Nasrin, Kim-Kwang Raymond Choo, Myung Ko and Anthony Rios	20
<i>Identifying Nuances in Fake News vs. Satire: Using Semantic and Linguistic Cues</i> Or Levi, Pedram Hosseini, Mona Diab and David Broniatowski	31
<i>Calls to Action on Social Media: Detection, Social Impact, and Censorship Potential</i> Anna Rogers, Olga Kovaleva and Anna Rumshisky	36
<i>Mapping (Dis-)Information Flow about the MH17 Plane Crash</i> Mareike Hartmann, Yevgeniy Golovchenko and Isabelle Augenstein	45
<i>Generating Sentential Arguments from Diverse Perspectives on Controversial Topic</i> ChaeHun Park, Wonsuk Yang and Jong Park	56
<i>Rumor Detection on Social Media: Datasets, Methods and Opportunities</i> Quanzhi Li, Qiong Zhang, Luo Si and Yingchi Liu	66
<i>Unraveling the Search Space of Abusive Language in Wikipedia with Dynamic Lexicon Acquisition</i> Wei-Fan Chen, Khalid Al Khatib, Matthias Hagen, Henning Wachsmuth and Benno Stein	76
<i>CAUnLP at NLP4IF 2019 Shared Task: Context-Dependent BERT for Sentence-Level Propaganda De- tection</i> Wenjun Hou and Ying Chen	83
<i>Fine-Grained Propaganda Detection with Fine-Tuned BERT</i> Shehel Yoosuf and Yin Yang	87
<i>Neural Architectures for Fine-Grained Propaganda Detection in News</i> Pankaj Gupta, Khushbu Saxena, Usama Yaseen, Thomas Runkler and Hinrich Schütze	92
<i>Fine-Tuned Neural Models for Propaganda Detection at the Sentence and Fragment levels</i> Tariq Alhindi, Jonas Pfeiffer and Smaranda Muresan	98
<i>Divisive Language and Propaganda Detection using Multi-head Attention Transformers with Deep Learn- ing BERT-based Language Models for Binary Classification</i> Norman Mapes, Anna White, Radhika Medury and Sumeet Dua	103
<i>On Sentence Representations for Propaganda Detection: From Handcrafted Features to Word Embed- dings</i> André Ferreira Cruz, Gil Rocha and Henrique Lopes Cardoso	107

<i>JUSTDeep at NLP4IF 2019 Task 1: Propaganda Detection using Ensemble Deep Learning Models</i>	
Hani Al-Omari, Malak Abdullah, Ola AlTiti and Samira Shaikh	113
<i>Detection of Propaganda Using Logistic Regression</i>	
Jinfen Li, Zhihao Ye and Lu Xiao	119
<i>Cost-Sensitive BERT for Generalisable Sentence Classification on Imbalanced Data</i>	
Harish Tayyar Madabushi, Elena Kochkina and Michael Castelle	125
<i>Understanding BERT performance in propaganda analysis</i>	
Yiqing Hua	135
<i>Pretrained Ensemble Learning for Fine-Grained Propaganda Detection</i>	
Ali Fadel, Ibrahim Tuffaha and Mahmoud Al-Ayyoub	139
<i>NSIT@NLP4IF-2019: Propaganda Detection from News Articles using Transfer Learning</i>	
Kartik aggarwal and Anubhav Sadana	143
<i>Sentence-Level Propaganda Detection in News Articles with Transfer Learning and BERT-BiLSTM-Capsule Model</i>	
George-Alexandru Vlad, Mircea-Adrian Tanase, Cristian Onose and Dumitru-Clementin Cercel	150
<i>Synthetic Propaganda Embeddings To Train A Linear Projection</i>	
Adam Ek and Mehdi Ghanimifard	157
<i>Findings of the NLP4IF-2019 Shared Task On Fine-grained Propaganda Detection</i>	
Giovanni Da San Martino, Alberto Barrón-Cedeño and Preslav Nakov	162

Workshop Program

- 9:00–9:10** Opening
- 9:10–9:30** *Assessing Post Deletion in Sina Weibo: Multi-modal Classification of Hot Topics*
Meisam Navaki Arefi, Rajkumar Pandi, Michael Carl Tschantz, Jedidiah R. Crandall, King-wa Fu, Dahlia Qiu Shi and Miao Sha
- 9:30–9:50** *Calls to Action on Social Media: Detection, Social Impact, and Censorship Potential*
Anna Rogers, Olga Kovaleva and Anna Rumshisky
- 9:50–10:10** *Identifying Nuances in Fake News vs. Satire: Using Semantic and Linguistic Cues*
Or Levi, Pedram Hosseini, Mona Diab and David Broniatowski
- 10:10–10:30** *Identifying Perspectives in Online News using Weakly Supervised Generative Models*
Srinivasan Iyer and Mike Lewis
- 10:30–11:00** Coffee Break
- 11:00–11:20** *Generating Sentential Arguments from Diverse Perspectives on Controversial Topic*
ChaeHun Park, Wonsuk Yang and Jong Park
- 11:20–11:40** *Unraveling the Search Space of Abusive Language in Wikipedia with Dynamic Lexicon Acquisition*
Wei-Fan Chen, Khalid Al Khatib, Matthias Hagen, Henning Wachsmuth and Benno Stein
- 11:40–12:00** *Findings of the NLP4IF-2019 Shared Task On Fine-grained Propaganda Detection*
Giovanni Da San Martino, Alberto Barrón-Cedeño and Preslav Nakov
- 12:00–12:20** *Divisive Language and Propaganda Detection using Multi-head Attention Transformers with Deep Learning BERT-based Language Models for Binary Classification*
Norman Mapes, Anna White, Radhika Medury and Sumeet Dua
- 12:20–12:40** *Fine-Grained Propaganda Detection with Fine-Tuned BERT*
Shehel Yoosuf and Yin Yang
- 12:40–14:00** Lunch Break
- 14:00–15:00** **Invited Talk: Elissa Redmiles (Princeton University/Microsoft), *Human Perception to Defend Democracy***
- 15:00–15:30** Coffee Break

Workshop Program (continued)

- 15:30-17:00** Poster Presentations:
- Detecting context abusiveness using hierarchical deep learning*
Ju-Hyoung Lee, Jun-U Park, Jeong-Won Cha and Yo-Sub Han
- How Many Users Are Enough? Exploring Semi-Supervision and Stylometric Features to Uncover a Russian Troll Farm*
Nayeema Nasrin, Kim-Kwang Raymond Choo, Myung Ko and Anthony Rios
- Mapping (Dis-)Information Flow about the MH17 Plane Crash*
Mareike Hartmann, Yevgeniy Golovchenko and Isabelle Augenstein
- Rumor Detection on Social Media: Datasets, Methods and Opportunities*
Quanzhi Li, Qiong Zhang, Luo Si and Yingchi Liu
- CAUnLP at NLP4IF 2019 Shared Task: Context-Dependent BERT for Sentence-Level Propaganda Detection*
Wenjun Hou and Ying Chen
- Neural Architectures for Fine-Grained Propaganda Detection in News*
Pankaj Gupta, Khushbu Saxena, Usama Yaseen, Thomas Runkler and Hinrich Schütze
- Fine-Tuned Neural Models for Propaganda Detection at the Sentence and Fragment levels*
Tariq Alhindi, Jonas Pfeiffer and Smaranda Muresan
- On Sentence Representations for Propaganda Detection: From Handcrafted Features to Word Embeddings*
André Ferreira Cruz, Gil Rocha and Henrique Lopes Cardoso
- JUSTDeep at NLP4IF 2019 Task 1: Propaganda Detection using Ensemble Deep Learning Models*
Hani Al-Omari, Malak Abdullah, Ola AlTiti and Samira Shaikh
- Detection of Propaganda Using Logistic Regression*
Jinfen Li, Zhihao Ye and Lu Xiao
- Cost-Sensitive BERT for Generalisable Sentence Classification on Imbalanced Data*
Harish Tayyar Madabushi, Elena Kochkina and Michael Castelle
- Understanding BERT performance in propaganda analysis*
Yiqing Hua
- Pretrained Ensemble Learning for Fine-Grained Propaganda Detection*
Ali Fadel, Ibrahim Tuffaha and Mahmoud Al-Ayyoub
- NSIT@NLP4IF-2019: Propaganda Detection from News Articles using Transfer Learning*
Kartik aggarwal and Anubhav Sadana
- Sentence-Level Propaganda Detection in News Articles with Transfer Learning and BERT-BiLSTM-Capsule Model*
George-Alexandru Vlad, Mircea-Adrian Tanase, Cristian Onose and Dumitru-Clementin Cercel
- Synthetic Propaganda Embeddings To Train A Linear Projection*
Adam Ek and Mehdi Ghanimifard