NAACL-HLT 2021

Teaching NLP

Proceedings of the Fifth Workshop

June 10 - 11, 2021
Mexico City (virtual)

©2021 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 978-1-954085-36-7

Introduction

Welcome to the Fifth Workshop on Teaching Natural Language Processing (NLP). This online workshop featured an exciting mix of papers, teaching material submissions, panels, talks, and participatory activities.

The field of NLP is growing rapidly, with new state-of-the-art methods emerging every year, if not sooner. As educators in NLP, we struggle to keep up. We need to make decisions about what to teach and how to teach it with every offering of a course, sometimes even as a course is being offered. The fast-paced nature of NLP brings unique challenges for curriculum design, and the immense growth of the field has lead to not just core NLP courses, but also to more specialized classes and seminars in subareas such as Natural Language Understanding, Computational Social Science, Machine Translation, and many more. We also have an increasing number of students interested in NLP, bringing with them a wide range of backgrounds and experiences.

We were happy to accept 13 long papers and 13 short papers on teaching materials. The latter were accompanied by exercises and assignments as Jupyter notebooks, software, slides, and teaching guidelines that will be made available via a repository created as result of the workshop. Both types of papers cover many topics: curriculum selection, teaching strategies, adapting to different student audiences, resources for assignments, and course or program design.

Our workshop also featured two panels, one on "What should we be teaching?" and another on "What does industry need?". The first panel featured Isabelle Augenstein (University of Copenhagen), Emily M. Bender (University of Washington), Yoav Goldberg (Bar Ilan University), and Dan Jurafsky (Stanford University). The second panel featured Lenny Bronner (The Washington Post), Delip Rao (Allen Institute for AI), Frank Rudzicz (University of Toronto), and Rachael Tatman (Rasa). We also had two amazing invited speakers, Ines Montani (Explosion) and Jason Eisner (Johns Hopkins University).

We thank the Program Committee, who thoughtfully reviewed these papers this year. We also appreciate the sponsorship funding we received from Google and Duolingo. Finally, we thank the workshop participants, whose interests in teaching allow us to establish and grow the next generation of NLP researchers and practitioners.

David Jurgens, Varada Kolhatkar, Lucy Li, Margot Mieskes, and Ted Pedersen (the co-organizers)

Organizing Committee

David Jurgens, University of Michigan
Varada Kolhatkar, University of British Columbia
Lucy Li, University of California, Berkeley
Margot Mieskes, Darmstadt University of Applied Sciences
Ted Pedersen, University of Minnesota, Duluth

Keynote Speakers

Jason Eisner, Johns Hopkins University
Ines Montani, Explosion

Panelists

Isabelle Augenstein, University of Copenhagen
Emily M. Bender, University of Washington
Lenny Bronner, The Washington Post
Yoav Goldberg, Bar-Ilan University
Dan Jurafsky, Stanford University
Delip Rao, Allen Institute for AI
Frank Rudzicz, University of Toronto
Rachael Tatman, Rasa

Program Committee

Benedikt Adelmann, University of Hamburg
Thomas Arnold, Technische Universität Darmstadt
Denilson Barbosa, University of Alberta
Austin Blodgett, Georgetown University
Su Lin Blodgett, Microsoft Research
Brendon Boldt, Carnegie Mellon University
Chris Brew, LivePerson
Julian Brooke, University of British Columbia
Paul Cook, University of New Brunswick
Carrie Demmans Epp, University of Alberta
Ivan Derzhanski, Bulgarian Academy of Sciences
Anna Feldman, Montclair State University
Alvin Grissom II, Haverford College
Bradley Hauer, University of Alberta

Marti A. Hearst, University of California, Berkeley
Katherine Keith, University of Massachusetts, Amherst
Grzegorz Kondrak, University of Alberta
Valia Kordoni, Humboldt University Berlin
Sandra Kübler, Indiana University
Vladislav Kubon, Charles University in Prague
James Kunz, Google / University of California, Berkeley
Lori Levin, Carnegie Mellon University
Deryle Lonsdale, Brigham Young University
Olga Lyashevskaya, HSE University
Christian M. Meyer, Technische Universitat Darmstadt
Bridget McInnes, Virginia Commonwealth University
Julie Medero, Harvey Mudd College
Gerald Penn, University of Toronto
Alexander Piperski, HSE University
Christopher Potts, Stanford University
Anoop Sarkar, Simon Fraser University
Nathan Schneider, Georgetown University
Alexandra Schofield, Harvey Mudd College
Melanie Siegel, Hochschule Darmstadt
Sowmya Vajjala, National Research Council
Gerhard Van Huyssteen, North-West University
Shira Wein, Georgetown University
Richard Wicentowski, Swarthmore College
Shuly Wintner, University of Haifa
Torsten Zesch, University of Duisburg-Essen
Heike Zinsmeister, University of Hamburg

Table of Contents

Pedagogical Principles in the Online Teaching of Text Mining: A Retrospection
Rajkumar Saini, György Kovács, Mohamadreza Faridghasemnia, Hamam Mokayed, Oluwatosin Adewumi, Pedro Alonso, Sumit Rakesh and Marcus Liwicki 1

Teaching a Massive Open Online Course on Natural Language Processing
Ekaterina Artemova, Murat Apishev, Denis Kirianov, Veronica Sarkisyan, Sergey Aksenov and Oleg Serikov 13

Natural Language Processing 4 All (NLP4All): A New Online Platform for Teaching and Learning NLP Concepts
Rebekah Baglini and Hermes Hjorth 28

A New Broad NLP Training from Speech to Knowledge
Maxime Amblard and Miguel Couceiro 34

Applied Language Technology: NLP for the Humanities
Tuomo Hiippala 46

A Crash Course on Ethics for Natural Language Processing
Annemarie Friedrich and Torsten Zesch 49

A dissemination workshop for introducing young Italian students to NLP
Lucio Messina, Lucia Busso, Claudia Roberta Combei, Alessio Miaschi, Ludovica Pannitto, Gabriele Sarti and Malvina Nissim 52

MiniVQA - A resource to build your tailored VQA competition
Jean-Benoit Delbrouck 55

From back to the roots into the gated woods: Deep learning for NLP
Barbara Plank 59

Learning PyTorch Through A Neural Dependency Parsing Exercise
David Jurgens 62

A Balanced and Broadly Targeted Computational Linguistics Curriculum
Emma Manning, Nathan Schneider and Amir Zeldes 65

Gaining Experience with Structured Data: Using the Resources of Dialog State Tracking Challenge 2
Ronnie Smith 70

The Flipped Classroom model for teaching Conditional Random Fields in an NLP course
Manex Agirrezabal 80

Flamingos and Hedgehogs in the Croquet-Ground: Teaching Evaluation of NLP Systems for Undergraduate Students
Brielen Madureira 87

An Immersive Computational Text Analysis Course for Non-Computer Science Students at Barnard College
Adam Poliak and Jalisha Jenifer 92

Introducing Information Retrieval for Biomedical Informatics Students
Sanya Taneja, Richard Boyce, William Reynolds and Denis Newman-Griffis 96

Contemporary NLP Modeling in Six Comprehensive Programming Assignments
Greg Durrett, Jifan Chen, Shrey Desai, Tanya Goyal, Lucas Kabela, Yasumasa Onoe and Jiacheng Xu 99

Interactive Assignments for Teaching Structured Neural NLP
David Gaddy, Daniel Fried, Nikita Kitaev, Mitchell Stern, Rodolfo Corona, John DeNero and Dan Klein 104

Learning about Word Vector Representations and Deep Learning through Implementing Word2vec
David Jurgens 108

Naive Bayes versus BERT: Jupyter notebook assignments for an introductory NLP course
Jennifer Foster and Joachim Wagner 112

Natural Language Processing for Computer Scientists and Data Scientists at a Large State University
Casey Kennington 115

On Writing a Textbook on Natural Language Processing
Jacob Eisenstein 125

Learning How To Learn NLP: Developing Introductory Concepts Through Scaffolded Discovery
Alexandra Schofield, Richard Wicentowski and Julie Medero 131

The Online Pivot: Lessons Learned from Teaching a Text and Data Mining Course in Lockdown, Enhancing online Teaching with Pair Programming and Digital Badges
Beatrice Alex, Clare Llewellyn, Pawel Orzechowski and Maria Boutchkova 138

Teaching NLP outside Linguistics and Computer Science classrooms: Some challenges and some opportunities
Sowmya Vajjala 149

Teaching NLP with Bracelets and Restaurant Menus: An Interactive Workshop for Italian Students
Ludovica Pannitto, Lucia Busso, Claudia Roberta Combei, Lucio Messina, Alessio Miaschi, Gabriele Sarti and Malvina Nissim 160

Workshop Program

Thursday, June 10, 2021

Poster Session 1

\+ ***Long Papers : Courses and Curricula***

Pedagogical Principles in the Online Teaching of Text Mining: A Retrospection
Rajkumar Saini, György Kovács, Mohamadreza Faridghasemnia, Hamam Mokayed, Oluwatosin Adewumi, Pedro Alonso, Sumit Rakesh and Marcus Liwicki

Teaching a Massive Open Online Course on Natural Language Processing
Ekaterina Artemova, Murat Apishev, Denis Kirianov, Veronica Sarkisyan, Sergey Aksenov and Oleg Serikov

Natural Language Processing 4 All (NLP4All): A New Online Platform for Teaching and Learning NLP Concepts
Rebekah Baglini and Hermes Hjorth

A New Broad NLP Training from Speech to Knowledge
Maxime Amblard and Miguel Couceiro

\+ ***Teaching Materials Short Papers : Courses and Curricula***

Applied Language Technology: NLP for the Humanities
Tuomo Hiippala

A Crash Course on Ethics for Natural Language Processing
Annemarie Friedrich and Torsten Zesch

A dissemination workshop for introducing young Italian students to NLP
Lucio Messina, Lucia Busso, Claudia Roberta Combei, Alessio Miaschi, Ludovica Pannitto, Gabriele Sarti and Malvina Nissim

\+ ***Teaching Materials Short Papers: Tools and Assignments***

MiniVQA - A resource to build your tailored VQA competition
Jean-Benoit Delbrouck

Thursday, June 10, 2021 (continued)

From back to the roots into the gated woods: Deep learning for NLP
Barbara Plank

Learning PyTorch Through A Neural Dependency Parsing Exercise
David Jurgens

Panel 1 : "What Should We Be Teaching?" Panelists : Isabelle Augenstein (U of Copenhagen), Emily M. Bender (U of Washington), Yoav Goldberg (Bar Ilan U), and Dan Jurafsky (Stanford U)

Poster Session 2

\+ ***Long Papers : Courses and Curricula***

A Balanced and Broadly Targeted Computational Linguistics Curriculum
Emma Manning, Nathan Schneider and Amir Zeldes

Gaining Experience with Structured Data: Using the Resources of Dialog State Tracking Challenge 2
Ronnie Smith

The Flipped Classroom model for teaching Conditional Random Fields in an NLP course
Manex Agirrezabal

\+ ***Teaching Materials Short Papers: Courses and Curricula***

Flamingos and Hedgehogs in the Croquet-Ground: Teaching Evaluation of NLP Systems for Undergraduate Students
Brielen Madureira

An Immersive Computational Text Analysis Course for Non-Computer Science Students at Barnard College
Adam Poliak and Jalisha Jenifer

Introducing Information Retrieval for Biomedical Informatics Students
Sanya Taneja, Richard Boyce, William Reynolds and Denis Newman-Griffis

Thursday, June 10, 2021 (continued)

+ ***Teaching Materials Short Papers : Tools and Assignments***

Contemporary NLP Modeling in Six Comprehensive Programming Assignments
Greg Durrett, Jifan Chen, Shrey Desai, Tanya Goyal, Lucas Kabela, Yasumasa Onoe and Jiacheng Xu

Interactive Assignments for Teaching Structured Neural NLP
David Gaddy, Daniel Fried, Nikita Kitaev, Mitchell Stern, Rodolfo Corona, John DeNero and Dan Klein

Learning about Word Vector Representations and Deep Learning through Implementing Word2vec
David Jurgens

Naive Bayes versus BERT: Jupyter notebook assignments for an introductory NLP course
Jennifer Foster and Joachim Wagner

Oral Presentations 1

+ ***Long Papers : Courses and Curricula***

Natural Language Processing for Computer Scientists and Data Scientists at a Large State University
Casey Kennington

On Writing a Textbook on Natural Language Processing
Jacob Eisenstein

Learning How To Learn NLP: Developing Introductory Concepts Through Scaffolded Discovery
Alexandra Schofield, Richard Wicentowski and Julie Medero

Thursday, June 10, 2021 (continued)

Keynote Address : Ines Montani (Explosion)

Friday, June 11, 2021

Oral Presentations 2

\+ ***Long Papers : Courses and Curricula***

The Online Pivot: Lessons Learned from Teaching a Text and Data Mining Course in Lockdown, Enhancing online Teaching with Pair Programming and Digital Badges
Beatrice Alex, Clare Llewellyn, Pawel Orzechowski and Maria Boutchkova

Teaching NLP outside Linguistics and Computer Science classrooms: Some challenges and some opportunities
Sowmya Vajjala

Teaching NLP with Bracelets and Restaurant Menus: An Interactive Workshop for Italian Students
Ludovica Pannitto, Lucia Busso, Claudia Roberta Combei, Lucio Messina, Alessio Miaschi, Gabriele Sarti and Malvina Nissim

Panel 2 : "What Does Industry Need?" Panelists : Lenny Bronner (Washington Post), Delip Rao (AI2), Frank Rudzicz (U of Toronto), and Rachel Tatman (Rasa)

Keynote Address : Jason Eisner (Johns Hopkins University)