Kevin Clark

http://cs.stanford.edu/~kevclark

kevclark@cs.stanford.edu

Education

Stanford University, Palo Alto, CA — September 2014 - Present
Ph.D. in Computer Science
Research Advisor: Christopher Manning
Main Research Interests: Self-supervised, semi-supervised, and multi-task learning applied to NLP.

University of Washington, Seattle, WA — September 2009 - June 2013
B.S. in Computer Science and Mathematics — GPA: 3.9

Research Experience

Student Researcher - Google Brain — June 2017 - Present
Advised by Quoc Le

- Developed ELECTRA, a new method for pre-training transformer networks on text. It uses a GAN-like loss function and trains much faster than prior models such as BERT, causing to perform well both at large scale (setting a new state-of-the-art for the SQuAD 2.0 question answering dataset) or even when trained on a single GPU.
- Built multi-task learning algorithms based on knowledge distillation to improve natural language understanding models, achieving consistent gains from multi-task learning across many tasks.
- Developed consistency-based semi-supervised learning algorithms and applied them to computer vision and natural language processing, achieving excellent results on multiple sequence labeling and image classification tasks.

Researcher Assistant - Stanford University — September 2014 - Present
Advised by Chris Manning

- Analyzed the attention mechanism of BERT and discovered it correlates remarkably well with aspects of dependency syntax and coreference.
- Analyzed thousands of counseling conversations from the nonprofit Crisis Text Line. Developed novel techniques to quantify various aspects of crisis counseling and determine which counseling strategies are the most effective.
- Developed methods to induce domain-specific sentiment lexicons from unlabeled corpora. Used the framework to analyze how word sentiments have changed over 150 years and how sentiment varies across 250 online communities.
- Applied reinforcement learning and clustering-based models to improve neural coreference resolution systems. Made substantial advancements to the current state-of-the-art for English and Chinese coreference and publicly released the models through Stanford's CoreNLP.

Researcher Assistant - University of Washington — September 2011 - June 2012
Advised by Oren Etzioni

- Built web interface and substantially improved the opinion mining component of RevMiner, a system that uses Natural Language Processing on reviews to concisely summarize and provide quality search of restaurants.

Publications

ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators.
Kevin Clark, Minh-Thang Luong, Quoc V. Le, and Christopher D. Manning.
Submitted to *International Conference on Learning Representations*, 2020.

What Does BERT Look At? An Analysis of BERT's Attention.
Kevin Clark, Urvashi Khandelwal, Omer Levy, and Christopher D. Manning.
BlackBoxNLP at Association for Computational Linguistics, 2019. **Best Paper Award.**

BAM! Born-Again Multi-Task Networks for Natural Language Understanding.
Kevin Clark, Minh-Thang Luong, Urvashi Khandelwal, Christopher D. Manning, and Quoc V. Le.
Association for Computational Linguistics, 2019.

Semi-Supervised Sequence Modeling with Cross-View Training.
Kevin Clark, Minh-Thang Luong, Christopher D. Manning, and Quoc V. Le.
Empirical Methods on Natural Language Processing, 2018.

Deep Reinforcement Learning for Mention-Ranking Coreference Models.
Kevin Clark and Christopher Manning. *Empirical Methods on Natural Language Processing*, 2016.

Inducing Domain-Specific Sentiment Lexicons from Unlabeled Corpora.
William L. Hamilton, Kevin Clark, Jure Leskovec, and Dan Jurafsky.
Empirical Methods on Natural Language Processing, 2016.

Improving Coreference Resolution by Learning Entity-Level Distributed Representations.
Kevin Clark and Christopher Manning. *Association for Computational Linguistics*, 2016.

Large-scale Analysis of Counseling Conversations: An Application of Natural Language Processing to Mental Health.
Tim Althoff*, Kevin Clark*, and Jure Leskovec.
Transactions of the Association for Computational Linguistics, 2016.
*equal contribution

Entity-Centric Coreference Resolution with Model Stacking.
Kevin Clark and Christopher Manning. *Association for Computational Linguistics*, 2015.

RevMiner: An Extractive Interface for Navigating Reviews on a Smartphone.
Jeff Huang, Oren Etzioni, Luke Zettlemoyer, Kevin Clark, and Christian Lee.
User Interface Software and Technology, 2012.

Teaching Experience

Head Teaching Assistant, NLP with Deep Learning Jan. 2018 - Mar. 2018

- Helped develop course material and manage TAs, gave three class lectures, mentored over 20 groups on their final projects.

Teaching Assistant, NLP with Deep Learning Jan. 2017 - Mar. 2017

- Wrote one of the three homework assignments, wrote lecture notes on computing neural network gradients, and mentored over 20 groups on their final projects.

Industry Experience

Software Engineering Intern, Sift Science August 2013 - May 2014

- Developed large-scale multi-task models for fraud detection.
- Built full-stack pipeline for detailed evaluation of fraud detection models.
- Improved infrastructure to scale the fraud detection system to hundreds of events per second.

Software Engineering Intern, Facebook June 2011 - August 2011

- Data mining to investigate which kinds of Facebook applications were creating unwanted posts.
- Trained machine learning models to identify apps contributing heavily to platform spam.

Achievements and Awards

- Google PhD Fellowship 2017
- Facebook PhD Fellowship Finalist 2016
- CRA Outstanding Undergraduate Researcher Honorable Mention 2013
- Topix Scholarship 2013
- Member of 1st place team out of 37 in the UW ACM Programming Competition 2012
- Meritorious Winner in the 2012 Mathematical Competition in Modeling 2012
- Microsoft Endowed Scholarship 2011
- University of Washington Research Training Group Grant 2009