

# Yizhe Zhang

One Microsoft Way, Redmond, WA, 98052

Email: [jeremy071242044@gmail.com](mailto:jeremy071242044@gmail.com) | Telephone: +1 919 536-8858

Website: <http://yizhezhang.net>

## Research Interest

My recent research focuses on several topics that are highly related and entangled: 1) progressive/non-autoregressive text generation; 2) sequence-level objective for end-to-end training; 3) long-form text generation and representation learning; 4) reasoning and planing in NLP.

From methodology perspective, I hope to develop generic architectures/methods that can be applied to various NLG tasks, by leveraging various techniques from NLP, deep generative models (DGM), MCMC and Reinforcement learning (RL).

## Experience

02/2018 – Present

**Senior Researcher | Natural Language Processsing @ Microsoft Research AI, Redmond, WA**

- Manager: Bill Dolan
- Research topics: text generation, conversation modeling, interplays between DGM, RL and NLP.

## Education

08/2013 – 02/2018

**Ph.D. in machine learning | Duke University, Durham, NC**

- Advisor: Dr. Lawrence Carin

08/2015 – 02/2018

**M.Sc. in statistics | Duke University, Durham, NC**

- Advisors: Drs. David Dunson, Scott Schmidler and Katherine Heller

08/2007 – 06/2011

**B.Sc. in physics | Nanjing University, Nanjing, China**

- Department of Intensive Institution, Kuang Yamin Honors School

## Publications

**Peer-reviewed Conferences and Journals** (\* equally contributed)

- **Yizhe Zhang**, Siqi Sun, Michel Galley, Yen-Chun Chen, Chris Brockett, Xiang Gao, Jianfeng Gao, Jingjing Liu, Bill Dolan. DialoGPT: Large-Scale Generative Pre-training for Conversational Response Generation. *demo track, ACL (2020)*
- Yichen Huang\*, **Yizhe Zhang**\*, Oussama Elachqar, Yu Cheng. INSET: Sentence Infilling with Inter-sentential Generative Pre-training. *ACL (2020)*
- Pengyu Cheng, Renqiang Min, Dinghan Shen, Christopher Malon, **Yizhe Zhang**, Yitong Li and Lawrence Carin. Improving Disentangled Text Representation Learning with Information Theoretical Guidance. *ACL (2020)*
- Xinjie Fan, **Yizhe Zhang**, Zhendong Wang, Mingyuan Zhou. Adaptive Correlated Monte Carlo for Contextual Categorical Sequence Generation. *ICLR (2020)*
- Yuan Li, Chunyuan Li, **Yizhe Zhang**, Xiujuan Li, Guoqing Zheng, Lawrence Carin, Jianfeng Gao. Complementary Auxiliary Classifiers for Label-Conditional Text Generation. *AAAI (2020)*

- Liqun Chen, Ke Bai, Chenyang Tao, **Yizhe Zhang**, Guoyin Wang, Wenlin Wang, Ricardo Henao, Lawrence Carin. Sequence Generation with Optimal-Transport-Enhanced Reinforcement Learning. *AAAI* (2020)
- Xiang Gao, **Yizhe Zhang**, Sungjin Lee, Michel Galley, Chris Brockett, Jianfeng Gao and Bill Dolan. Structuring latent spaces for stylized response generation. *EMNLP* (2019)
- Dianqi Li, **Yizhe Zhang**, Zhe Gan, Yu Cheng, Chris Brockett, Ming-Ting Sun and Bill Dolan. Domain Adaptive Text Style Transfer. *EMNLP* (2019)
- Xinnuo Xu, **Yizhe Zhang**, Lars Liden and Sungjin Lee. Unsupervised Dialogue Spectrum Generation for Log Dialogue Ranking. *SIGDIAL* (2019), *Best paper nomination*
- Liqun Chen, Guoyin Wang, Chenyang Tao, Dinghan Shen, **Yizhe Zhang** and Lawrence Carin. Improving Textual Network Embedding with Global Attention via Optimal Transport. *ACL* (2019)
- Dinghan Shen, Asli Celikyilmaz, **Yizhe Zhang**, Liqun Chen, Xin Wang, Jianfeng Gao, Lawrence Carin. Towards Generating Long and Coherent Text with Multi-Level Latent Variable Models. *ACL* (2019)
- Xiang Gao, Sungjin Lee, **Yizhe Zhang**, Chris Brockett, Michel Galley, Jianfeng Gao, Bill Dolan. Jointly Optimizing Diversity and Relevance in Neural Response Generation. *NAACL* (2019)
- Liqun Chen, **Yizhe Zhang**, Ruiyi Zhang, Chenyang Tao, Zhe Gan, Haichao Zhang, Bai Li, Dinghan Shen, Changyou Chen, Lawrence Carin. Improving Sequence-to-Sequence Learning via Optimal Transport. *ICLR* (2019)
- **Yizhe Zhang**, Michel Galley, Jianfeng Gao, Zhe Gan, Xiujuan Li, Chris Brockett, Bill Dolan. Generating Informative and Diverse Conversational Responses via Adversarial Information Maximization. *NeurIPS* (2018)
- Liqun Chen, Shuyang Dai, Chenyang Tao, Dinghan Shen, Zhe Gan, Haichao Zhang, **Yizhe Zhang**, Lawrence Carin. Adversarial Text Generation via Feature-Mover's Distance. *NeurIPS* (2018)
- Yunchen Pu, Shuyang Dai, **Yizhe Zhang**, Zhe Gan and Lawrence Carin. Multi-Domain Joint Distribution Learning with Generative Adversarial Nets. *ICML* (2018)
- Dinghan Shen, Guoyin Wang, Wenlin Wang, Martin Renqiang Min, Qinliang Su, **Yizhe Zhang**, Chunyuan Li, Ricardo Henao and Lawrence Carin. On Simple Word-Embedding-Based Models and Associated Pooling Mechanisms. *ACL* (2018)
- Guoyin Wang, Chunyuan Li, Wenlin Wang, **Yizhe Zhang**, Dinghan Shen, Xinyuan Zhang, Ricardo Henao and Lawrence Carin. Joint Embedding of Words and Labels for Text Classification. *ACL* (2018)
- Dinghan Shen, **Yizhe Zhang**, Ricardo Henao, Qinliang Su, Lawrence Carin. Deconvolutional Latent-Variable Model for Text Sequence Matching. *AAAI* (2018).
- Wenlin Wang, Piyush Rai, Yunchen Pu, Kai Fan, **Yizhe Zhang**, Ricardo Henao, Lawrence Carin. A Flexible Probabilistic Framework for Learning to Predict Unseen Classes. *AAAI* (2018).
- Zhe Gan, Liqun Chen, Weiyao Wang, Yunchen Pu, **Yizhe Zhang**, Lawrence Carin. Triangle Generative Adversarial Networks. *NIPS* (2017).
- **Yizhe Zhang**, Changyou Chen, Zhe Gan, Lawrence Carin. Stochastic Gradient Monomial Gamma Sampler. *ICML* (2017).
- **Yizhe Zhang**, Zhe Gan, Zhi Chen, Lawrence Carin. Adversarial Feature Matching for Text Generation. *ICML* (2017).
- **Yizhe Zhang**, Xiangyu Wang, Changyou Chen, Lawrence Carin. Towards Unifying Hamiltonian Monte Carlo and Slice Sampling. *NIPS* (2016).
- Changyou Chen, Nan Ding, Chunyuan Li, **Yizhe Zhang**, Lawrence Carin. Distributed Bayesian Learning with Stochastic Gradient MCMC. *NIPS* (2016).
- **Yizhe Zhang**, Ricardo Henao, Lawrence Carin. Dynamic Poisson Factor Analysis. *ICDM* (2016).
- Kai Fan, **Yizhe Zhang**, Katherine Heller. Triply Stochastic Variational Inference for Non-linear Beta Process Factor Analysis. *ICDM* (2016).

- **Yizhe Zhang**, Ricardo Henao, Jianling Zhong, Lawrence Carin, Alexander Hartemink. Learning a Hybrid Architecture for Sequence Regression and Annotation. *AAAI (2016)*.
- **Yizhe Zhang**, Ricardo Henao, Chunyuan Li, Lawrence Carin. Bayesian Dictionary Learning with Gaussian Processes and Sigmoid Belief Networks. *IJCAI (2016)*.
- **Yizhe Zhang**, Changyou Chen, Ricardo Henao, Lawrence Carin. Laplacian Hamiltonian Monte Carlo. *ECML (2016)*.
- **Yizhe Zhang**, Yupeng He and Chaochun Wei (2015). MOST+: a Motif Finding Approach Combining Genomic Sequence and Heterogeneous Genome-wide Signatures. *BMC Genomics*.
- Yupeng He, **Yizhe Zhang**, Guangyong Zheng and Chaochun Wei (2012). CRF-based Transcription Factor Binding Site Finding System. *BMC Genomics*.
- Jiemeng Liu, Haifeng Wang, Hongxing Yang, **Yizhe Zhang**, Jinfeng Wang, Fangqing Zhao and Ji Qi. (2012). Composition-based Classification of Short Metagenomic Sequences Elucidates the Landscapes of Taxonomic and Functional Enrichment of Microorganisms. *Nucleic Acids Research*.

## Workshop

- Woon Sang Cho, **Yizhe Zhang**, Sudha Rao, Chris Brockett and Sungjin Lee. Generating a Common Question from Multiple Documents using Multi-source Encoder-Decoder Models. *WNGT, EMNLP (2019)*
- Woon Sang Cho, Pengchuan Zhang, **Yizhe Zhang**, Xiujun Li, Michel Galley, Chris Brockett, Mengdi Wang, Jianfeng Gao. Towards coherent and cohesive long-form text generation. *Workshop on Narrative Understanding, NAACL (2019)*
- **Yizhe Zhang**, Zhe Gan and Lawrence Carin. Generating Text with Adversarial Training. *Workshop on Adversarial Training, NIPS (2016)*.
- **Yizhe Zhang**, Lawrence Carin. Learning Dictionary with Spatial and Inter-dictionary Dependency. *Workshop on representation learning, NIPS (2015)*.

## Preprint

- **Yizhe Zhang**, Xiang Gao, Sungjin Lee, Chris Brockett, Michel Galley, Jianfeng Gao, Bill Dolan. Consistent Dialogue Generation with Self-supervised Feature Learning.
- Woon Sang Cho, **Yizhe Zhang**, Sudha Rao, Asli Celikyilmaz, Chenyan Xiong, Jianfeng Gao, Mengdi Wang, Bill Dolan. Unsupervised Common Question Generation from Multiple Documents using Reinforced Contrastive Coordinator.

## Teaching

- Advanced Machine Learning (STA571). Instructor: *Katherine Heller*
- Probabilistic Machine Learning (CS571). Instructor: *Cynthia Rudin*

## Rewards

- Department Fellowship (2008-2011)
- National Excellent Graduate Scholarship (top 1%) (2012)
- Travel award: NIPS (2015, 2016), ICML (2017), ICDM (2016), IJCAI (2016), AAAI (2016)

## Professional Services

Senior Program committee: AAAI

Program committee: NeurIPS, ICML, ICLR, AAAI, IJCAI, ACL, EMNLP, NAACL, CoNLL, SIAM

Website chair for ACL 2020

## Proficiency

- Theano, Tensorflow, C/C++, Python, Java, Lua, MATLAB and R.