Yuntian DENG

Google Scholar◇ Personal Website

yuntian@uwaterloo.ca

Appointment

University of Waterloo	Waterloo, Canada
Incoming Assistant Professor	Starting Fall 2024
Vector Institute	Toronto, Canada
Incoming Faculty Affiliate	Starting Fall 2024
Allen Institute for AI	Seattle, WA
Post-Doctoral Researcher (Young Investigator)	Starting Jul 2023

Education

Harvard University	Cambridge, MA
Ph.D. in Computer Science	May 2023
Advisors: Alexander Rush & Stuart Shieber	
Carnegie Mellon University	Pittsburgh, PA
Master in Language Technologies	Aug 2016
Advisor: Eric Xing	
Tsinghua University	Beijing, China
Bachelor of Engineering, Department of Automation	Jul 2014

Research Interests

My research focuses on structure modeling for language models, which aims to analyze and improve the modeling of various textual structures, including narrative progression, coreference resolution, and parse trees. To achieve this goal, my work develops deep generative models, which integrate the strengths of deep learning and probabilistic structure modeling techniques. You can find more about my research in my thesis here.

Awards

Argonne National Lab Impact Award	Jan 2023
University of Chicago Rising Stars in Data Science	Oct 2022
ACM Gordon Bell Special Prize for Covid Research	Nov 2022
Nvidia Fellowship (10 awardees)	Dec 2021
Harvard Certificates of Distinction in Teaching	Spring 2019, Fall 2020, Fall 2021

ICML expert reviewer	Feb 2021
Twitch Fellowship Finalist	Dec 2020
NeurIPS 2020 top 10% reviewer	Oct 2020
DAC 2020 Best Paper Award	Jul 2020
Baidu Fellowship (10 awardees)	Feb 2019
French American Doctoral Exchange Program (Embassy of France, 10 US awardees)	Jul 2018
ACL 2017 Best Demo Paper Award Runner-up	Aug 2017

Publications

*Equal Contribution.

Selected Papers

1 **Yuntian Deng**, Volodymyr Kuleshov, Alexander Rush. Model Criticism for Long-Form Text Generation. *EMNLP 2022*

2 **Yuntian Deng**, Anton Bakhtin, Myle Ott, Arthur Szlam, Marc'Aurelio Ranzato. Residual Energy-Based Models for Text Generation. *ICLR 2020*

3 **Yuntian Deng**, Alexander Rush. Cascaded Text Generation with Markov Transformers. *NeurIPS 2020*

4 Zachary Ziegler*, **Yuntian Deng***, Alexander Rush. Neural Linguistic Steganography. *EMNLP 2019 **oral*** (**Demo**: steganography.live)

5 **Yuntian Deng**, Noriyuki Kojima, Alexander Rush. Markup-to-Image Diffusion Models with Scheduled Sampling. *ICLR 2023* (Demo: huggingface.co/spaces/yuntian-deng/latex2im)

6 **Yuntian Deng***, Yoon Kim*, Justin Chiu, Demi Guo, Alexander M. Rush. Latent Alignment and Variational Attention. *NeurIPS 2018*

7 **Yuntian Deng**, Anssi Kanervisto, Jeffrey Ling, Alexander M. Rush. Image-to-Markup Generation with Coarse-to-Fine Attention. *ICML 2017* (Demo: im2markup.yuntiandeng.com)

Journal Papers

8 Anton Bakhtin*, **Yuntian Deng***, Sam Gross, Myle Ott, Marc'Aurelio Ranzato, Arthur Szlam. Residual Energy-Based Models for Text. *JMLR 2021*

Conference Papers

9 Richa Rastogi, Yair Schiff, Alon Hacohen, Zhaozhi Li, Ian Lee, **Yuntian Deng**, Mert R. Sabuncu, Volodymyr Kuleshov. Semi Parametric Inducing Point Networks. *ICLR 2023*

10 Maxim Zvyagin*, Alexander Brace*, Kyle Hippe*, **Yuntian Deng***, Bin Zhang, Cindy Orozco Bohorquez, Austin Clyde, Bharat Kale, Danilo Perez-Rivera, Heng Ma, Carla M. Mann, Michael Irvin,

J. Gregory Pauloski, Logan Ward, Valerie Hayot, Murali Emani, Sam Foreman, Zhen Xie, Diangen Lin, Maulik Shukla, Weili Nie, Josh Romero, Christian Dallago, Arash Vahdat, Chaowei Xiao, Thomas Gibbs, Ian Foster, James J. Davis, Michael E. Papka, Thomas Brettin, Rick Stevens, Anima Anandkumar, Venkatram Vishwanath, Arvind Ramanathan, GenSLMs: Genome-scale language models reveal SARS-CoV-2 evolutionary dynamics. ACM **Gordon Bell Special Covid Prize**

11 Justin Chiu*, **Yuntian Deng***, Alexander Rush. Low-Rank Constraints for Fast Inference in Structured Models. *NeurIPS 2021*

12 **Yuntian Deng**, Alexander Rush. Sequence-to-Lattice Models for Fast Translation. *EMNLP 2021 Findings*

13 Keyon Vafa, **Yuntian Deng**, David Blei, Alexander Rush. Rationales for Sequential Predictions. *EMNLP 2021 **oral***

14 Anton Bakhtin, Sam Gross, Myle Ott, **Yuntian Deng**, Marc'Aurelio Ranzato, Arthur Szlam. Real or Fake? Learning to Discriminate Machine from Human Generated Text. *arXiv 1906.03351*

15 Thierry Tambe, En-Yu Yang, Zishen Wan, **Yuntian Deng**, Vijay Janapa Reddi, Alexander Rush, David Brooks, Gu-Yeon Wei. Algorithm-Hardware Co-Design of Adaptive Floating-Point Encodings for Resilient Deep Learning Inference. *Design Automation Conference (DAC) 2020 **Best Paper Award***

16 **Yuntian Deng**, David Rosenberg, Gideon Mann. Challanges in end-to-end neural scientific table recognition. *ICDAR 2019*

17 Ran Zhao, **Yuntian Deng**, Mark Dredze, Arun Verma, David Rosenberg, Amanda Stent. Visual Attention Model for Cross-sectional Stock Return Prediction and End-to-End Multimodal Market Representation Learning. *FLAIRS 2019*

18 Sebastian Gehrmann, **Yuntian Deng**, Alexander M. Rush. Bottom-Up Abstractive Summarization. *EMNLP 2018*

19 Pengtao Xie, **Yuntian Deng**, Yi Zhou, Abhimanu Kumar, Yaoliang Yu, James Zou, Eric P Xing. Learning Latent Space Models with Angular Constraints. *ICML 2017*

20 Guillaume Klein, Yoon Kim, **Yuntian Deng**, Jean Senellart, Alexander M Rush. OpenNMT: Open-Source Toolkit for Neural Machine Translation. *ACL 2017 **Best Demo Paper Award Runner-Up***

21 Zichao Yang, Zhiting Hu, **Yuntian Deng**, Chris Dyer, Alex Smola. Neural Machine Translation with Recurrent Attention Modeling. *EACL 2017*

22 Xuezhe Ma, Yingkai Gao, Zhiting Hu, Yaoliang Yu, **Yuntian Deng**, Eduard Hovy. Dropout with expectation-linear regularization. *ICLR 2017*

23 Hao Zhang, Zhiting Hu, **Yuntian Deng**, Mrinmaya Sachan, Zhicheng Yan, Eric P. Xing. Learning Concept Taxonomies from Multi-modal Data. *ACL 2016*

24 Pengtao Xie, **Yuntian Deng**, Eric P. Xing. Diversifying Restricted Boltzmann Machine for Document Modeling. *KDD 2015*

25 Zhiting Hu, Poyao Huang, **Yuntian Deng**, Yingkai Gao, Eric P. Xing. Entity Hierarchy Embedding. *ACL 2015*

Teaching Experience

Harvard University

Head TA, Intro to Computational Linguistics and NLP (Stuart Shieber)	Aug 2021 - Dec 2021
Head TA, Intro to Computational Linguistics and NLP (Stuart Shieber)	Aug 2020 - Dec 2020
TA, Natural Language Processing (Alexander Rush)	Jan 2019 - May 2019
Lecture, Natural Language Processing - Translation (Alexander Rush)	Feb 13, 2019
TA, Advanced Machine Learning (Alexander Rush)	Aug 2017 - Dec 2017

Cornell University

Lab Material Preparation, Break Through Tech AI (Alexander Rush)	Aug 2021

Carnegie Mellon University

TA, Graduate Probabilistic Graphical Models (Matthew Gormley, Eric Xing)	Jan 2016 - May 2016
TA, Graduate Machine Learning (Ziv Bar Joseph, Eric Xing)	Aug 2015 - Dec 2015

Revere High School

Volunteer Instructor of AP CS (Microsoft Philanthropies)	Sep 2020 - May 2021

Professional Service

2023: Reviewer of ICML, ACL, ACL ARR, ACL Demo

2022: Reviewer of ACL ARR, EMNLP, ICML, NeurIPS, ICLR, EACL, ML Reproducibility Challenge, ACL Demo, NAACL Demo, and EMNLP GEM Workshop.

2021: Expert Reviewer of ICML. Oral Session Volunteer of ACL. Reviewer of NeurIPS, ACL, ACL ARR, ICLR, AAAI, EACL, NAACL, ML Reproducibility Challenge, ICLR EBM Workshop, IEEE TIIS, and NEJLT. Volunteer of EMNLP.

2020: Top 10% Reviewer of NeurIPS. Reviewer of ICML, ACL, EMNLP, AACL, AAAI, ACL Demo, COLING, IEEE TIFS, Transactions on Information Systems, and Journal of Computer Science and Technology. Volunteer of ICML. Volunteer moderator of ACL.

2019: Reviewer of NeurIPS, EMNLP, ICLR, NAACL, AAAI, IEEE TNNLS, ACM Computing Surveys, NeurIPS LIRE Workshop, NeurIPS Reproducibility Challenge, EMNLP Summarization Workshop, ICML GraphReason Workshop, and ICLR DeepGenStruct Workshop.

2018: Reviewer of ICLR.

2017: Reviewer of ICML TADGM workshop.

2016: Volunteer of ICML.

Community Service

Childcare Chair at NAACL 2022 D&I Committee	2022
Mentor of Harvard Women in CS (in charge of a bi-weekly reading group)	Sep 2020 - Present
Volunteer Instructor of AP CS at Revere High School	Sep 2020 - May 2021
Leader of Harvard English Language Table	Sep 2021 - Present

Internships

Nvidia / Argonne National Lab (Anima Anandkumar / Arvind Ramanathan)	May - Dec 2022
Facebook AI Research (Marc'Aurelio Ranzato)	May - Dec 2019
Bloomberg CTO Office (David Rosenberg, Gideon Mann)	Jan - Aug 2017
UCSD (Charles Elkan)	Jul - Sep 2013

Talks

NYU AI School 2023: Natural Language Processing	Jun 2023, New York, NY
AI2: Structural Coherence in Text Generation	Jun 2023, Seattle, Washington
UMass Amherst: Structural Coherence in Text Generation	May 2023, remote
Georgia Tech: Structural Coherence in Text Generation	Apr 2023, Atlanta, Georgia
Harvard Kempner Institute: Structural Coherence in Text Generation	Mar 2023, remote
U Alberta: Structural Coherence in Text Generation	Mar 2023, Edmonton, Canada
TTIC: Structural Coherence in Text Generation	Feb 2023, Chicago
U Waterloo ECE: Structural Coherence in Text Generation	Feb 2023, remote
Cornell Seminar in NLU: Structural Coherence in Text Generation	Feb 2023, New York
U Waterloo CS: Structural Coherence in Text Generation	Jan 2023, Waterloo, Canada

U Chicago Rising Star: Model Criticism for Long-Form Text Generation	Nov 2022, Chicago
Princeton NLP: Model Criticism for Long-Form Text Generation	Nov 2022, Princeton
EMNLP 2021 Findings: Sequence-to-Lattice Models for Fast Translation	Oct 2021, remote
OpenAI: Residual Energy-based Models for Text Generation	Apr 2021, remote
NeurIPS 2020: Cascaded Text Generation with Markov Transformers	Oct 2020, remote
Baidu: Cascaded Text Generation with Markov Transformers	Jun 2020, remote
ICLR 2020: Residual Energy-based Models for Text Generation	Apr 2020, remote
LinkedIn: Residual Energy-based Models for Text Generation	Mar 2020, remote
Wayfair: Neural Encoder-Decoder Models	Nov 2019, Boston
EMNLP 2019: Neural Linguistic Steganography	Nov 2019, Hong Kong
FAIR NLP Meeting: Energy-Based Models for Text Generation	Aug 2019, New York City
NeurIPS 2018: Variational Attention	Dec 2018, Montreal, Canada
The French National Center for Scientific Research: Variational Attention	Jul 2018, France
Association for Machine Translation in the Americas: OpenNMT	Mar 2018, Boston
Open-Source Neural Machine Translation Workshop: Image-to-Text	Mar 2018, Paris, France
Tencent Social Network Group TSAIC: OpenNMT	Dec 2017, Shenzhen, China
ICML: Image-to-Markup Generation	Aug 2017, Sydney, Australia

Open Source Projects

Cascaded Generation (120 Github stars)	PyTorch
Neural Lingustic Steganography (157 Github stars)	PyTorch
Variational Attention (322 Github stars)	PyTorch
Image-to-Markup Generation (1k Github stars)	LuaTorch
OpenNMT (5.8k Github stars)	PyTorch
Attention OCR (1.1k Github stars)	Tensorflow