Nathan Ng

+1 (858)-722-0654 · nathanng@mit.edu
https://github.com/nng555

Research Interests

NLP: semi-supervised learning, self-training, data augmentation
ML for Healthcare: differential privacy, clinical NLP, molecular design

Education

Massachusetts Institute of Technology — Cambridge, MA
Visiting Scholar — *Sept 2021 – present*
Advisor: Prof. Marzyeh Ghassemi

University of Toronto — Toronto, Ontario
Ph.D. Machine Learning — *Sept 2019 – present*
Advisor: Prof. Marzyeh Ghassemi

University of California San Diego — San Diego, California
BS Computer Science (Summa Cum Laude) — *Sep 2014 – Jun 2018*
Advisor: Prof. Zachary Lipton and Prof. Julian McAuley

Professional Experience

Prescient Design — New York, New York
Research Intern (Kyunghyun Cho) — *Summer 2022*
Blind Denoising of Long Read DNA Sequences with Self-Supervised Set Learning

Facebook — New York, New York (Virtual)
Research Intern (Naman Goyal) — *Summer 2021*
Growing Switch Transformers for Multilinguality

Google — Mountain View, California (Virtual)
Research Intern (Qi Guo) — *Summer 2020*
Improving Dialogue Breakdown Detection Models with Semi-Supervised Learning

Facebook — Menlo Park, California
Research Engineer — *Sep 2018 – Sep 2019*

Facebook — Menlo Park, California
Software Engineering Intern — *Summer 2016 / Summer 2017*

Qualcomm — San Diego, California
Software Engineering Intern — *Summer 2015*

Refereed Publications

1. **N. Ng**, K. Cho, and M. Ghassemi. "Predicting Out-of-Domain Generalization with Local Manifold Smoothness". In: *Proc. of ICLR*. 2023.
2. **N. Ng**, J. W. Park, J. H. Lee, R. Kelly, and K. Cho. "Blind Denoising of Long Read DNA Sequences with Self-Supervised Set Learning". In: *TMLR*. 2023.
3. J. Bae, **N. Ng**, A. Lo, M. Ghassemi, and R. Grosse. "If Influence Functions are the Question, What is the Answer?" In: *Proc. of NeurIPS*. 2022.
4. **N. Ng**, K. Cho, and M. Ghassemi. "SSMBA: Self-Supervised Manifold Based Data Augmentation for Improving Out-of-Domain Robustness". In: *Proc. of EMNLP*. 2020.
5. T. Lau, **N. Ng**, J. Gingold, N. Desai, J. McAuley, and Z. C. Lipton. "Embryo staging with weakly-supervised region selection and dynamically-decoded predictions". In: *Proc. of Machine Learning for Healthcare*. 2019.

6. **N. Ng**, K. Yee, A. Baevski, M. Ott, M. Auli, and S. Edunov. "Facebook FAIR's WMT19 News Translation Task Submission". In: *Proc. of WMT.* 2019.

7. K. Yee, **N. Ng**, Y. Dauphin, and M. Auli. "Simple and Effective Noisy Channel Modeling for Neural Machine Translation". In: *Proc. of EMNLP.* 2019.

8. **N. Ng**, R. Gabriel, J. McAuley, C. Elkan, and Z. Lipton. "Predicting surgery duration with neural heteroscedastic regression". In: *Proc. of Machine Learning for Healthcare.* 2017.

Workshop Publications

1. **N. Ng**, N. Thangarajan, J. Pan, M. Ghassemi, and Q. Guo. "Improving Dialogue Breakdown Detection with Semi-Supervised Learning". In: *Proc. of Workshop on Human in the Loop Dialogue Systems at NeurIPS.* 2020. Oral.

2. M. Ott, S. Edunov, A. Baevski, A. Fan, S. Gross, **N. Ng**, D. Grangier, and M. Auli. "fairseq: A fast, extensible toolkit for sequence modeling". In: *Proc. of NAACL-HLT 2019: Demonstrations.* 2019.

3. **N. Ng**, J. McAuley, Z. Lipton, and N. Desai. "Predicting Embryo Morphokinetics in Videos with Late Fusion Nets Dynamic Decoders". In: *Proc. of ICLR.* 2018.

Shared Tasks

1st in Dialogue Breakdown Detection Challenge English task *2020*
1st in WMT News Translation English ↔ German task *2019*
1st in WMT News Translation English ↔ Russian task *2019*

Teaching and Talks

Facebook Internal Lecturer

Special Topics in Deep Learning: NLP and Translation *Feb 2019, Sep 2019*

University of Toronto Teaching Assistant

CSC 2515: Introduction to Machine Learning (Graduate Level) *Fall 2020*
CSC 2541: Topics in Machine Learning: Machine Learning for Health *Winter 2020*
CSC 311: Introduction to Machine Learning *Fall 2019*

University of California, San Diego Teaching Assistant

CSE 101: Design and Analysis of Algorithms *Winter 2018*
CSE 158: Web Mining and Recommender Systems *Fall 2017*
CSE 21: Mathematics for Algorithms and Systems *Winter 2017*
CSE 11: Introduction to Object-Oriented Programming *Fall 2015*

Professional Activities

Organizer

Workshop on Robustness in Sequence Modeling at NeurIPS *2022*

Reviewer

NeurIPS *2022*
Machine Learning for Healthcare *2020*

Honors and Awards

- Jacobs Scholarship, University of California San Diego *2014*
- Regents Scholarship, University of California San Diego *2014*