

Meng Cao

+1-438-722-6288  meng.cao@mail.mcgill.ca  mcao516  mcao516.github.io

EDUCATION

McGill University / Mila - Québec AI Institute

Sept. 2019 – Present

Ph.D. in Computer Science

GPA: 4.0/4.0

Advisor: Jackie Chi Kit Cheung

Northeastern University (China)

Sept. 2015 – June 2019

Major in Software Engineering

GPA: 3.96/4.0

Ranking: 1/59

RESEARCH INTERESTS

Natural Language Processing, Text Summarization, Text Generation, Reinforcement Learning

PUBLICATIONS

[1] Mechanisms of Non-Factual Hallucination in Language Models

Lei Yu*, **Meng Cao***, Jackie CK Cheung, Yue Dong

ArXiv Preprint

[2] DRLC: Reinforcement Learning with Dense Rewards from LLM Critic

Meng Cao, Lei Shu, Lei Yu, Yun Zhu, Nevan Wichers, Yinxiao Liu, Lei Meng

ArXiv Preprint

[3] Successor Features for Efficient Multi-Subject Controlled Text Generation

Meng Cao, Mehdi Fatemi, Jackie CK Cheung and Samira Shabanian

ICML 2024

[4] Systematic Rectification of Language Models via Dead-end Analysis

Meng Cao, Mehdi Fatemi, Jackie CK Cheung and Samira Shabanian

ICLR 2023

[5] Responsible AI Considerations in Text Summarization Research: A Review of Current Practices

Yu Lu Liu, **Meng Cao**, Su Lin Blodgett, Jackie Chi Kit Cheung, Alexandra Olteanu and Adam Trischler

Findings of EMNLP 2023

[6] Learning with Rejection for Abstractive Text Summarization

Meng Cao, Yue Dong, Jingyi He and Jackie C. K. Cheung

EMNLP 2022

[7] Hallucinated but Factual! Inspecting the Factuality of Hallucinations in Abstractive Summarization

Meng Cao, Yue Dong and Jackie C. K. Cheung.

ACL 2022

[8] Factual Error Correction for Abstractive Summarization Models

Meng Cao, Yue Dong, Jiapeng Wu and Jackie C. K. Cheung

EMNLP 2020

[9] TeMP: Temporal Message Passing for Temporal Knowledge Graph Completion

Jiapeng Wu, **Meng Cao**, Jackie Chi Kit Cheung, William L. Hamilton

EMNLP 2020

[10] Referring Expression Generation Using Entity Profiles

Meng Cao and Jackie Chi Kit Cheung

EMNLP-IJCNLP 2019

INDUSTRY EXPERIENCES

Google Research – Mountain View

Research Intern

Aug. 2023 – Nov. 2023

Advisor: Lei Meng, Shu Lei

- Tackling the reward sparsity problem in RLHF
- Leverage the critique ability of LLM to generate dense intrinsic rewards
- Extrinsic & intrinsic rewards combined significantly improve sample efficiency of PPO

Microsoft Research – Montréal

Machine Learning Research Intern

Apr. 2022 – Mar. 2023

Advisor: Samira Shabanian & Mehdi Fatemi

- Work on detoxification of large language models like GPT-3
- Our approach, *rectification*, utilizes a separate but significantly smaller value function for LM detoxification
- Paper accepted by ICLR 2023

Huawei Noah's Ark Lab – Montréal

Applied Machine Learning Research Intern

May 2021 – Mar. 2022

Advisor: Mehdi Rezagholizadeh

- Knowledge distillation for text generation tasks
- Data augmentation for knowledge distillation using cluster labels

Borealis AI – Toronto

Research Intern, Alan Team

May 2020 – Aug. 2020

Advisor: Yanshuai Cao

- Build a distributed data-parallel Text-to-SQL system
- Applied meta-learning algorithms (MAML, Reptile, MetaReg) in cross-database semantic parsing task
- Achieved promising results in the zero-shot domain transfer setting

IBM Research – Beijing

Research Assistant Intern, Information Analytics Team

Sept. 2018 – June 2019

Advisor: Shiwan Zhao

- Chinese word segmentation: responsible for training Chinese word segmentation model. Proposed a novel segmentation model based on n-gram and information entropy to segment the input sentences
- Neural Architecture Search: PyTorch implementation of Google's ENAS neural architecture algorithm
- Migrated a neural network model which is implemented in Keras into Java environment using Deeplearning4j

PRESENTATIONS AND TALKS

Learning with Rejection for Abstractive Text Summarization (Poster)
EMNLP 2022

Dec., 2022
Abu Dhabi, UAE

Factual Error Correction for Abstractive Summarization Models (Poster)
EMNLP 2020

Nov., 2020
Online

Referring Expression Generation Using Entity Profiles (Poster)
EMNLP 2019

Nov., 2019
Hong Kong, China

SCHOLARSHIPS

China National Scholarship (¥ 16,000)

2015-2016, 2017-2018

The Chinese Government, among top-0.2% all undergraduates in China

The Bao Gang Education Scholarship (¥ 10,000)

Sept. 2018

China BAOWU Steel Group, selected rate: 4/20000 in NEU

The First-Class Scholarship in Northeastern University (¥ 9,000)

2015-2016, 2016-2017, 2017-2018

Northeastern University, among top-3% undergraduates in NEU

AWARDS

"Top Ten Undergraduates" in Northeastern University

May 2018

Northeastern University, among top-0.2% undergraduates in NEU

The Excellent Performance Award for IBM Undergraduate Innovation Lab Program

Apr. 2017

IBM Research China, Beijing

The First Prize in the 8th National College Students Mathematics Competition Oct. 2016

Chinese Mathematical Society

The First Prize in the 9th National College Students Mathematics Competition Oct. 2017

Chinese Mathematical Society

The Meritorious Winner for the Mathematical Contest in Modeling (MCM/ICM), 2018 Feb. 2018

COMAP, the Consortium for Mathematics and Its Applications (USA)

The Honorable Mention for the Mathematical Contest in Modeling (MCM/ICM), 2017 Feb. 2017

COMAP, the Consortium for Mathematics and Its Applications (USA)

RESEARCH FUNDING

Mitacs Globalink Research Internship (\$19,500) June 2018 – Sept. 2018

Mitacs Globalink summer research internship program (\$4,500) & Mitacs graduate student funding (\$15,000)

SKILLS

Programming Languages: skilled in: Java, Python (especially NumPy, PyTorch); familiar with: Tensorflow, JavaScript, SQL

Extensive knowledge of: Natural Language Processing, Text Summarization, Deep Learning, Reinforcement Learning