

NING SHI

LinkedIn | Twitter | GitHub | Homepage | Google Scholar

Phone: +86 13917068481 | Email: ning.shi@gatech.edu & mrshininnnnn@gmail.com

Address: Pudong New Area, Shanghai, China, 201201

RESEARCH INTERESTS

My research interests lie in the broad area of deep learning, natural language processing (NLP), reinforcement learning (RL), and their applications. In particular, I'm interested in the extent to which neural networks can do the same as humans to understand the world systematically from the perspective of cognition.

EDUCATION

Georgia Institute of Technology Master of Science in Computer Science	<i>Aug. 2019 - Aug. 2022 (Expected)</i> <i>GPA: 4.0/4.0</i>
Syracuse University Master of Science in Applied Data Science	<i>Oct. 2018 - May 2021</i> <i>GPA: 4.0/4.0</i>
New York University Master of Science in Management and Systems	<i>Sep. 2016 - May 2018</i> <i>GPA: 3.9/4.0</i>
Donghua University Bachelor of Management in Electronic Commerce	<i>Sep. 2010 - Jul. 2014</i>

PUBLICATIONS

Peer-reviewed Conference

- [C.1] Wei, W., Boxin, W., **Ning, S.**, Jinfeng, L., Bingyu, Z., Xiangyu, L., & Rong, Z. (2021). Counterfactual Adversarial Learning with Representation Interpolation. Findings of the Association for Computational Linguistics: EMNLP 2021.
- [C.2] **Ning, S.**, Wei, W., Boxin, W., Jinfeng, L., Xiangyu, L., & Zhouhan, L. (2021). Incorporating External POS Tagger for Punctuation Restoration. Conference of the International Speech Communication Association (INTERSPEECH) 2021.
- [C.3] **Ning, S.**, Ziheng, Z., Haotian, Z., & Yichen, G. (2020). Recurrent Inference in Text Editing. Findings of the Association for Computational Linguistics: EMNLP 2020.
- [C.4] Keran, Z., Yingda, L., **Ning, S.**, & Yuheng, H. (2019). The Impact of Emotion On Social Media Content Consumption: Evidence From Video Analysis On Live Streaming Platforms. Institute for Operations Research and the Management Sciences (INFORMS) 2019.

Preprint

- [P.1] **Ning, S.**, Boxin, W., Wei, W., Xiangyu, L., Rong, Z., Hui, X., Xinbing, W., & Zhouhan, L. (2021). From SCAN to Real Data: Systematic Generalization via Meaningful Learning.
- [P.2] Zihan, Z., Jinfeng L., **Ning S.**, Bo Y., Rong Z., Hui X., Chao Z., Donghong S. (2021). RoBERT: Towards Robust Chinese BERT-based Classifiers.
- [P.3] Wei W., Bingyu Z., **Ning S.**, Xiangyu L., Bo Y., Hui X. (2021). IV-FAIR: Disentangling Sensitive Information through Instrumental Variables

RESEARCH AND INDUSTRY EXPERIENCE

Alibaba Group, Turing Security Lab
Senior Algorithm Engineer

Hangzhou, China
Nov. 2020 - Present

- Explored cool applications of NLP in Security.

Shining Lab*Founder, Supervised by Prof. Zhouhan Lin*

Shanghai, China

Aug. 2019 - Present

- Directed the team to research cutting-edge innovation in NLP.
- 3 papers accepted to the top conference and 3 works in process.

Learnable, Inc.*Machine Learning Engineer*

Boston, United States

Sep. 2018 - Sep. 2020

- Supervised the NLP team to define problems, conduct researches, and provide solutions.
- Led several studies such as masked language modeling, semantic parsing, and text post-processing.
- Followed, reproduced, and applied the state-of-the-art to fulfill specific production requirements.

Pythagoras Investment Management LLC*Algorithm Engineer Intern*

New York City, United States

Dec. 2017 - May 2018

- Adopted best practices to design and develop an automatic digital currency trading system.
- Managed over 10+ exchanges' APIs for various quantitative trading strategies.

China Net Cloud (Shanghai) Co., Ltd*Project Manager*

Shanghai, China

Mar. 2014 - Jun. 2016

- Managed projects from end to end and took ultimate responsibility for projects' success.

TALKS**Recurrent Inference in Text Editing***Fourth Workshop on Structured Prediction for NLP (SPNLP) 2020**Nov. 20, 2020***SERVICES**

Reviewer of ARR 2021/2022, ICML 2022; Secondary Reviewer of EMNLP-IJCNLP 2019; DataKind Volunteer

PROJECTS**Proactive Conversational AI, Alibaba Group***Nov. 2020 - Present*

- Worked as the owner of the inspection services chatbot for violations detection.
- Directed the team to apply the forefront studies to improve the dialogue system.
- Identified thousands of hidden illegal commodity transactions over 12+ risk points.

Reinforcement Learning in Text Editing, Alibaba Group*Aug. 2021 - Present*

- Defined text editing as a solvable Markov decision process to involve RL methods.
- Investigated offline imitation learning to deal with text editing.

Chinese-LaTeX-mixed Mathematics Language Modeling, Learnable, Inc.*Jun. 2019 - Nov. 2019*

- Cleaned, tokenized, and managed math corpus from HTML-embedded raw data.
- Applied and improved BERT-based MLM for Chinese-LaTeX-mixed corpus representation.

Post-Processing for Optical Character Recognition, Learnable, Inc.*Apr. 2019 - Aug. 2019*

- Researched on involving language modeling into a real-time OCR system.
- Established a post-processor to improve OCR results and a simulator to reproduce OCR errors.

Knowledge Points Tagging for K12 Examination, Learnable, Inc.*Jul. 2018 - Apr. 2019*

- Developed SQL pipeline to manage the 14 Million Question Bank.
- Investigated multi-label multi-class classification via hierarchical inference.

The Relationship of Academic Capital and Mendeley Altmetrics, New York University*Advised by Farid Razzak and Thomas Christo.**Sep. 2017 - Jul. 2018*

- A thesis submitted in partial fulfillment of the requirements for the Master's degree at NYU.