

EDUCATION

- Peking University**, Beijing, China
M.S. in Computer Science. Advisor: Prof. Junfeng Hu Sept. 2016–June 2019
- Peking University**, Yuanpei Honors Program, Beijing, China
B.S. in Physics. Advisor: Dr. Qibo Chen Sept. 2012–June 2016

RESEARCH INTERESTS

My research interests lie in the general area of machine learning, particularly in **representation learning** and **generative modeling**. Currently, I focus on applications on discrete and structured data settings such as **natural language processing** (*e.g.*, text generation and representation) and **computational biology** (*e.g.*, molecular generation for drug discovery).

RESEARCH EXPERIENCE

Engineer and Researcher at ByteDance AI Lab Shanghai, China
Machine Learning and Natural Language Computing Group, advisor: Dr. Hao Zhou and Dr. Lei Li. 2018–Present
At ByteDance, I mainly worked on designing machine learning algorithms for interpretable and controllable structured data generation.

- **Knowledge distillation via exploring local targets.** (submitted to ICLR 2021)
- Propose a simple yet effective knowledge distillation method which projects the supervision signals of a teacher model into the student’s parameter space.
- Outperform strong knowledge distillation baselines and achieve the **state-of-the-art** performance in various **image and text classification** tasks.
- **Dispersed Exponential Family Mixture VAEs for interpretable text generation.** (accepted by ICML 2020.)
- Introduce VAEs with a **mixture of exponential family distribution** for **interpretable** text generation.
- Theoretically analyze the general **mode-collapse problem** in mixture of exponential family VAEs and propose an effective method to fix it.
- Experimental results show that our approach does induce a more **meaningful space**, and it outperforms strong baselines in text generation and dialog generation.
- **Variational Template Machine for data-to-text generation.** (accepted by ICLR 2020)
- Propose a novel method “variational template machine (VTM)” to generate text descriptions from **tabular data**.
- Construct a disentangled **template** and semantic latent space, and provide additional signals for the template space with large-scale descriptions data without tables.
- Experiments show that VTM is able to generate **more diverse** outputs while maintaining fluency and quality.

Student Researcher at Peking University Beijing, China
Key Laboratory of Computational Linguistics, advisor: Prof. Junfeng Hu 2016–2019

I worked on extracting information from structural Chinese academic data.

- **Sentence pattern extraction from the titles of Chinese natural science papers.** (thesis)
- Design a pipeline to extract the sentence patterns from the titles of Chinese natural science papers.

- Define and extract specific semantic roles for the scientific corpus and extract basic sentence patterns of the titles. Construct semantic graphs to extend the patterns.
- Sentence patterns could be used in downstream tasks, such as searching system customized by user preferences.

Student Researcher at Peking University

Beijing, China

State Key Laboratory of Nuclear Physics and Technology, advisor: Dr. Qibo Chen

2014–2015

I worked on investigating the motion of nuclei by quantal models.

- **Wobbling geometry in a simple triaxial rotor.** (accepted by CPC 2015)
- We investigated the spectroscopy properties and angular momentum geometry for the **wobbling motion** of a **simple triaxial rotor**.
- A specific evolutionary track can be used to depict the motion of a triaxial rotating nuclei.

PUBLICATIONS

- [1] **W. Shi**, H. Zhou, N. Miao, and L. Li, “Dispersing Exponential Family Mixture VAEs for Interpretable Text Generation”, in *Proceedings of the 37th International Conference on Machine Learning (ICML)*, 2020.
- [2] R. Ye, **W. Shi**, H. Zhou, Z. Wei, and L. Li, “Variational Template Machine for Data-to-Text Generation”, in *International Conference on Learning Representations (ICLR)*, 2020.
- [3] N. Miao, H. Zhou, C. Zhao, **W. Shi**, and L. Li, “Kernelized Bayesian Softmax for Text Generation”, in *the 33rd Conference on Neural Information Processing Systems (NeurIPS)*, 2019.
- [4] **W. Shi** and Q. Chen, “Wobbling geometry in a simple triaxial rotor”, *Chinese Physics C (CPC, impact factor 5.861 in 2018)*, vol. 39, no. 5, p. 054 105, 2015.

TEACHING

- **Teaching Assistant** at Peking University Autumn 2018
Introduction to Computing
- **Teaching Assistant** at Peking University Spring 2017
Data Structures and Algorithms

SCHOLARSHIPS AND AWARDS

- FuGuang Fellowship 2012–2016
- Merit Student Award of Peking University 2014
- Excellent Student of Academic Records of Peking University 2013
- Model Student of Social Work of Peking University 2012