# Jiayu Zheng

401-447-9534 | jiayu_zheng@brown.edu | Academic Homepage | github.com/JiayuZheng

## EDUCATION

**Brown University**                                                                  Providence, RI
*Sc.M. in Data Science*                                                      *Sept. 2022 – Present*
- GPA 4.0
- Relevant Courses: Machine Learning, Deep Learning, Artificial Intelligence, Computer Visions, Operating Systems, Parallel Computing

**Zhengjiang University**                                                              Zhejiang, China
*Sc.B. in Chemistry with Honors (Specialization in AI4Science)*              *Sept. 2018 – May 2022*
- Major GPA: 3.92/4.0, Cumulative GPA: 3.8/4.0, Rank: 2%
- Graduated from the Cho Kochen Honors College
- Certificate in **Computer Science**
- Relevant Courses: C&C++, Java, Mathematical Modeling, Digital Logic Design, Linux Application Technology, Advanced Data Structure and Algorithm Analysis

## PUBLICATIONS

A Molecular Stereostructure Descriptor Based On Spherical Projection                          2020
- Licheng Xu, Xin Li, Miaojiong Tang, Luotian Yuan, **Jiayu Zheng**, Shuoqing Zhang, Xin Hong

## RESEARCH EXPERIENCE

**Graduate Research Assistant**                                                      Providence, RI
BATS Lab, *Brown University*                                                  *Sept. 2023 – Present*
- Advised by Prof. Stephen Bach
- Built soft-prompted weak supervision systems
- Designed parameter-efficient soft prompts trained in a bootstrapping manner to improve the performance of LLMs on out-of-domain tasks

**Graduate Research Assistant**                                                      Providence, RI
Conversational AI Lab, *Brown University*                                      *June 2023 – Present*
- Co-advised by Prof. Shekhar Pradhan and Prof. Ritambhara Singh
- Built language models with semantically richer embeddings with additional supervision signals from paired images
- Designed relation-aware contrastive loss, which effectively encourages models to discriminate between relations in text
- Collected a hard-negative dataset for MSCOCO Captions by prompting large language models (LLM)
- Designed a weakly supervised, LLM-in-the-loop, labeling function to indicate whether alternative captions retain original semantics

**Undergraduate Research Assistant**                                                  Zhejiang, China
*Zhejiang University*                                                        *Apr. 2020 – July 2021*
- Mostly focused on finding appropriate representations of molecules that allow for efficient and precise prediction of properties
- Designed a numeric descriptor to transform the van der Waals force field of molecules to a sequence of 2D matrices, collected a dataset using the descriptor, trained a CNN to predict the stereo-selectivity of reactions
- Designed a distance-aware message-passing mechanism based on the intuition that most atoms only interact with nearby atoms within molecules, which greatly reduces the MSE of energy prediction

## PROJECTS

**GourmAIt** | *Python, PyTorch* July 2023 – Sept. 2023
- Implemented Google Research's Noisy Student Training, where pseudo labels generated by a teacher model are used to train a larger-or-equal-size student model, which is then used as the teacher in the next iteration
- Implemented stochastic depth, stepwise unfreezing scheduling, and learning rate scheduling to maximize the performance gain in finetuning ResNet

**WeenixOS** | *C, Unix kernel, X86_64 arch* Feb. 2023 – May 2023
- Semester-long project of the course Operating Systems instructed by Prof. Thomas Doeppner
- Developed a well-functioning, kernel-based, Unix operating system, the WeenixOS
- Implemented a thread pool, device drivers, a VFS, an S5FS, and a virtual memory system
- Implemented various system calls, like do_waitpid, do_read, do_write, do_brk, do_fork, etc.

**TransformerHub** | *Python, PyTorch* Sept. 2022 – Aug. 2023
- Implemented various encoder-only, decoder-only, and encoder-decoder Transformer models including Transformer, BERT, GPT, ViT, and CLIP
- Incorporated advanced designs like mixed precision training, rotary position embedding
- Evaluated these models on different tasks, including text summarization, language modeling, and sentiment analysis

## TEACHING AND MENTORING

**Teaching Assistant** Providence, RI
CS1410: Artificial Intelligence, *Brown University* *Sept. 2023 – Dec. 2023*
- Instructors: Thao Nguyen, Stefanie Tellex
- Mentored groups of students for the final projects, held debugging hours twice per week, answered miscellaneous questions on Ed, and graded assignments

**Mentor in Data101 boot camp** Providence, RI
*Brown University* *Apr. 2023*
- Introduced fundamental concepts of data science to a group of underrepresented high school students in the College-Day data science boot camp

**Mentor in WiDS at Brown** Providence, RI
*Brown University* *Feb. 2023 – Mar. 2023*
- Provided tutoring to prospective young women data scientists
- Rendered hands-on guidance of numpy and pandas, and demonstrated a classification task on a time-series weather dataset

**Head Teaching Assistant** Zhejiang, China
Structural Chemistry and Spectroscopy, *Zhejiang University* *Feb. 2020 – Dec. 2021*
- Prepared course materials, designed assignments, and hosted group presentations

## HONORS AND AWARDS

| | |
|---|---:|
| Outstanding Graduates of Zhejiang University | 2022 |
| Cho Kochen Honors College Scholarship for Innovation (publications or international contest awards) | 2020, 2021 |
| Outstanding Students of Zhejiang University | 2019, 2021 |
| Zhejiang University Scholarship, Third Prize (top 15%) | 2020 |
| Zhejiang University Scholarship, First Prize (top 3%) | 2019, 2021 |

## TECHNICAL SKILLS

**Programming**: Python, C/C++, CUDA C++, Java, SQL (MySQL), Latex
**Deep Learning**: Tensorflow, PyTorch, Huggingface Transformers, OpenAI Gym, Scikit-learn
**Development Tools**: Venv, Conda, Mamba, Docker, Git, ssh, Jupyter, CMake, SLURM, gdb, wandb
**Deep Learning Systems**: DeepSpeed(ZeRO), Accelerate, Flexgen, MegatronLM
**LLMs/VLMs**: GPT2/3/3.5/4, Claude, OPT, T0++, Llama1/2, Falcon, T5, Flan-T5, Mistral, CLIP