Jianzhu Yao

https://yao-jz.github.io

Email : cnyaojz@gmail.com
Mobile : +86-139-3693-8133

Research Interest

My primary research interests are in the fields of **Natural Language Processing and Machine Learning**, including conversational systems, consistent and faithful language generation, and biomedical domain application.

Education

Tsinghua University **Sep 2019 – Present**

Bachelor of Computer Science and Technology; GPA: 3.86/4.00 *Beijing, China*

Courses: Linear Algebra(4.0), Object-Oriented Programming(4.0), Calculus(4.0), Discrete Mathematics(4.0), Introduction to Artificial Intelligence(4.0), Computer Graphics(4.0), Operating Systems(4.0), Compiler(4.0), Organization(4.0), Software Engineering

Papers

- **A Benchmark for Understanding and Generating Dialogue between Characters in Stories**
  Jianzhu Yao, Ziqi Liu, Jian Guan, Minlie Huang
  AAAI 2023, under review

- **EVA2.0: Investigating Open-Domain Chinese Dialogue Systems with Large-Scale Pre-Training**
  Yuxian Gu, Jiaxin Wen, Hao Sun, Yi Song, Pei Ke, Chujie Zheng, Zheng Zhang, ***Jianzhu Yao***, Xiaoyan Zhu, Jie Tang, Minlie Huang
  arXiv:2203.09313

Research Experience

Biomedical Terminology Definition Generation and Faithful Evaluation Metrics **Jul 2022 – Present**

Advisor: Prof. Sheng Wang, at Wang Lab, University of Washington, Seattle *Beijing, China(Remote)*

- Utilized biomedical knowledge base Unified Medical Language System(UMLS) to retrieve relative information, and implemented Graph Convolutional Network on the knowledge graph in order to generate more faithful and accurate terminology definition.
- Designed terminology definition generation evaluation metrics focused on the faithfulness and factuality, considering existing automatic evaluation metrics' bad performance on keyword precision.

A Benchmark for Dialogue Understanding and Generation in Stories **Jul 2021 – Mar 2022**

Advisor: Prof. Minlie Huang, at Conversational AI Group, Tsinghua University *Beijing, China*

- Designed a benchmark for dialogue understanding and generation between characters in stories, which put forth a higher requirement for understanding character relationships and the story lines. The paper is submitted to AAAI 2023.
- Built a new story dataset with marked dialogue and speakers, and tested the performance of existing baselines on the benchmark, to investigate the machine's ability in dialogue understanding and generation.
- Proposed to learn explicit character representation to guide the generation and understanding. The automatic and manual evaluation shows that our approach outperforms the strong baselines by 30% and can generate more coherent and informative dialogue.

Open-Domain Chinese Dialogue System EVA1.0 and EVA2.0 **Jul 2021 – Mar 2022**

Advisor: Prof. Minlie Huang, at Conversational AI Group, Tsinghua University *Beijing, China*

- Collaborated with team members on the development of the Open-Domain Chinese Dialogue System EVA1.0 and EVA2.0, and responsible for decoding strategies. The paper is submitted to EMNLP 2022.
- Trained and implemented a contradiction detection classifier (RoBERTa) for Chinese Dialogue Systems, and designed the regeneration pipeline in order to avoid the inconsistent generation, and proven the effectiveness using case studies. The detector is also used in other emotional chat bots.

Mixture of Experts Model and Expert Pruning Technique **Mar 2022 – Jun 2022**

Advisor: Dr. Tao Ge, at Natural Language Computing Group, Microsoft Research Asia *Beijing, China*

- Observed insightful patterns of computing cost trend of all2all communication by conducting profiling experiments of Mixture of Experts models with different numbers of experts on each device.
- Explored expert pruning algorithms on the Mixture of Experts language models in order to accelerate computation, lower computation latency, and reduce GPU memory usage.

Researched on Rotation Invariance of Image Local Features in Object Reconstruction **2021**

Advisor: Prof. Shimin Hu, at Graphics & Geometric Computing Group, Tsinghua University *Beijing China*

- Researched on rotation invariance of image local features in object reconstruction based on deep learning methods and traditional algorithms, which shows existing neural networks suffer from serious data bias and cannot outperform traditional algorithms on some data with rotation invariance.
- Observed the influence of image transformation (like rotation, style transformation, affine transformation, streching) on feature extraction and matching in SuperPoint, D2-Net, SuperGlue, and SIFT.

Selected Projects

- **Gym Reservation Script(Python, Selenium, PhantomJS, Shell, Wireshark)**: Developed a script using Python and PhantomJS to register for a timeslot at gym with captured cookie by Wireshark. (May 2021)
- **K-Means Clustering on MNIST(Python, K-Means, NumPy, Matplotlib, sklearn)**: Implemented K-Means Clustering Algorithm on MNIST dataset. (Apr 2022)
- **Ensemble Learning Experiment on Spam Classification(Python, SVM, Decision-Tree, Bagging, AdaBoost)**: Incorporated SVM and Decision-Tree with ensemble learning method Bagging and AdaBoost for spam classification. (May 2022)
- **Four-in-a-row AI Bot(C++, Monte Carlo Tree Search, Upper Confidence Bound Apply to Tree)**: Designed an AI bot with MCTS and UCB algorithm to compete with other AI bots of Four-in-a-row game. (Apr 2021)
- **Enterprise personnel permissions Management System(Vue, JavaScript, Front-end)**: Developed a front-end service for the management of personnel permissions in an enterprise(Kuaishou). (Nov 2021)
- **Education platform app based on knowledge graph(Java, SpringBoot, Android Studio, Full-stack)**: Developed the Front-end and Back-end educational application IntelEdu based on knowledge graph with a semantic similarity classifier(BERT). (Aug 2021)
- **KD-Tree based Stochastic Progressive Photon Mapping image rendering framework(C++)**: Implemented the SPPM algorithm, KD-Tree based Bounding Box for intersection acceleration, 3-d scene construction, anti-aliasing, motion blur, etc. (May 2021)

Services and Membership

- EMNLP 2022 Review Committee for the track Dialogue and Interactive Systems, Jul. 2022
- Member of the Bodybuilding Team of Tsinghua University, 2022 – Present
- Member of the Winter Swimming Association of Tsinghua University, 2021 – Present
- Member of Tsinghua University Admission Group in Heilongjiang Province, Jun. 2020
- Member of Student Association for Science and Technology, Dept. CST of Tsinghua University, 2020 – Present
- Member of Swimming Team of CST in Tsinghua University, 2019 – 2019

Workshops

Mutlimodal Machine Learning + Mixture of Experts **June 2022**

Advisor: Prof. Bowen Zhou, at Tsinghua University *Beijing, China*

- Did a literature survey, and gave a presentation on the multimodal machine learning: Hierarchical Overview of Multimodal Machine Learning — And potential ideas with personal insights. [slides]
- Inspired by multimodal + MoE, designed experiments for Multitask Multimodal MoE, and wrote a research proposal: MMMoE: Advancing Multimodal Mixture-of-Experts Architecture to Power Next-Generation Multimodal Paradigm. [pdf][slides]

Technical Skills

Programming Languages: Python, Latex, Java, C, C++, C#, Shell, HTML/CSS, MATLAB, Assembly(RISC-V, x86)
Developer Tools: VS Code, PyCharm, Git, Docker, Linux, Xcode, Unity Hub, Vim, Android Studio, Vivado, Quartus
Libraries/Frameworks: PyTorch, Transformers, Fairseq, spaCy, CoreNLP, NumPy, Matplotlib, Jupyter Notebook, Scikit, Keras, Django, SpringBoot, Vue

Honors / Awards

- Academic Excellence Award, Tsinghua University, 2020
- 37th National Physics Competition for College Students in Beijing, China, First Prize, 2021
- The first prize of Engineering Technology Challenge, Tsinghua University, 2019
- The first prize of the Province Degree in the national Physic competition (senior group), 2018