

Fufangchen Zhao, Ph.D.

✉ zhaofufangchen@bupt.edu.cn

🌐 **Phone:** +86 137 0090 3801



Education

- 2023 – **Ph.D. Beijing University of Posts and Telecommunications**
Computer Science and Technology, successive postgraduate and doctoral programs of study. GPA: 3.63/4.00
- 2021 – 2023 **M.Sc. Beijing University of Posts and Telecommunications**
Intelligent Science and Technology. GPA: 3.63/4.00
- 2017 – 2021 **Bachelor Hainan University**
Computer Science and Technology. GPA: 3.82/4.00

Professional Experience

- 2023.4 – 2024.7 **NLP Intern Researcher, SenseTime Research**
Project 1: Development of consistency detection tools for cross team research on industrial large-scale models.
Independently design and develop (from dataset to overall tool implementation) a cross team consistency detection tool for industrial large-scale language models, used for consistency detection of cross team delivery of large models in industrial research. The research results have been promoted in all major language model development departments in SenseTime Research.
Project 2: Development of Embedding Model for RAG Capability in the Financial Sector Based on *Sensechat*.
With comparative learning training data from the financial field, use the hidden embedding of the EOS token at the end of the sentence as the sentence embedding, and directly fine tune mistral-7b as the embedding model. And ranked **first** on the *Chinese MTEB leaderboard*. (March 2024)
Project 3: Provide data support for *Sensechat* R&D.
Developed a pipeline for constructing training data and benchmarks, generating **300 B** high-quality pre-training corpora and **20 B** high-quality SFT-labeled corpora, providing a solid foundation for the team's model training and evaluation.
Project 4: Research on Enhancing Large Models for Edge Deployment
Collaborated with full-time staff to utilize the pruning method in **Progressive Wanda**, complemented by high-quality calibration samples. This approach ensured that the model's performance on specific tasks (high-frequency tasks in actual application scenarios) remained essentially unchanged while reducing the model's parameter count by 75%.

Professional Experience (continued)

2021.12 – 2023.3

📖 **Intern Researcher**, XiZhi AI

Projection: Math Word Problem Processing (MWP) Development and Research Implementation.

Utilized COT (Chain-of-Thought) to prompt ChatGPT with 8 original prompt templates and self-designed templates, enabling ChatGPT to generate reasonable step-by-step solutions for math problems.

Generated code using ChatGPT and GPT-4 to solve plotting problems, focusing on zero-shot and few-shot approaches.

Research Publications (Under Review)

- 1 F. Zhao, G. Jian, and D. Yan, "Recse: Portable reshaping features for sentence embedding in self-supervised contrastive learning," *arXiv preprint arXiv:2408.04975 (NAACL 2025 under review)*, 2024. 🔗 URL: <https://arxiv.org/abs/2408.04975>.
- 2 F. Zhao, G. Jin, R. Zhao, J. Huang, and F. Tan, "Simct: A simple consistency test protocol in llms development lifecycle," *arXiv preprint arXiv:2407.17150 (NAACL 2025 under review)*, 2024. 🔗 URL: <https://arxiv.org/abs/2407.17150>.
- 3 F. Zhao, H. Zhang, J. Jiang, and D. Yan, "Bscse: Bidirectional soft samples for contrastive learning of unsupervised sentence embedding," *2025 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2025 under review)*, 2024.

Skills

- | | |
|-----------|--|
| Languages | 📖 Strong reading, writing and speaking competencies for English (CET-6) |
| Coding | 📖 Python, SQL, XML/XSL, L ^A T _E X, Pytorch, git, ... |