# MAOJIANG SU

✉ sumaojiang@mail.ustc.edu.cn 🔗 Personal Homepage 🏛 Google Scholar

## EDUCATION

**University of Science and Technology of China (USTC)**                    Hefei, Anhui, China
**School of the Gifted Young**
**Bachelor of Information and Computational Science**                         Sep 2021 - Jun 2025
- *GPA*: 3.72/4.3 (Overall/Scale) | *Ranking*: 20/95

## RESEARCH INTERESTS

**Keywords:** Statistical Machine Learning; Foundation Model; Large Language Model; AI/ML for Science.

My research focuses on machine learning theory, with an emphasis on the statistical and computational capabilities of foundation models and generative AI. My goal is to develop new strong models and methods for important scientific and societal problems, with robust (i.e., low-assumption and realistic) theoretical guarantees to ensure their empirical performance.

With this aim, my research interests span machine learning, optimization, statistics, and computational complexity theory.

## PUBLICATIONS

Google Scholar | arXiv | $^\star$Equal contribution. $^{\alpha\beta}$Alphabetical authorship.

[5] Pretrained Transformers are Deep Optimizers: Provable In-Context Learning for Deep Model Training,
Weimin Wu$^\star$, **Maojiang Su**$^\star$, Jerry Yao-Chieh Hu$^\star$, Zhao Song, Han Liu ,
Under review, ArXiv preprint forthcoming, 2024.

[4] SPARQ: Outlier-Efficient SpeechLM with Fast Adaptation and Robust Quantization,
Shang Wu$^\star$, Yen-Ju Lu$^\star$, Haozheng Luo$^\star$, **Maojiang Su**, Jerry Yao-Chieh Hu, Jiayi Wang, Jing Liu, Najim Dehak,
Jesus Villalba, Han Liu,
Under review, ArXiv preprint forthcoming, 2024.

[3] Making Genomic Foundation Models more Foundational Requires Outlier Removal: A Case Study on DNABERT-2,
Haozheng Luo$^\star$, Chenghao Qiu$^\star$, **Maojiang Su**, Zhihan Zhou, Jerry Yao-Chieh Hu, Zoe Mehta, Guo Ye, Han Liu,
Under review, ArXiv preprint forthcoming, 2024.

[2] Computational Limits of Low-Rank Adaptation (LoRA) for Transformer-Based Models,
Jerry Yao-Chieh Hu, **Maojiang Su**, En-Jui Kuo, Zhao Song, Han Liu,
Under review, ArXiv Preprint https://arxiv.org/abs/2406.03136, 2024.

[1] Boolean operations on generalized polygons with arcs,
Xiaolong Feng, **Maojiang Su**, Weihua Tong, Falai Chen,
Under review, ArXiv preprint forthcoming, 2024.

## RESEARCH EXPERIENCES

**Statistical Machine Learning Group, Northwestern University**
*Research Intern, with Prof.* Han Liu                                      Feb 2024 - Present
**Computational Limits of Low-Rank Adaptation (LoRA) for Transformer-Based Models**
- Explored the computational limits of Low-Rank Adaptation (LoRA) update for fine-tuning transformer-based models.
- Identified a phase transition behavior in LoRA adaptation on transfomer-based model.
- Proved the existence of nearly linear algorithms for gradient computation by controlling the LoRA update term by term.

**Pretrained Transformers are Deep Optimizers: Provable In-Context Learning for Deep Model Training**
- Investigated the strong in-context learning capabilities of transformers model.
- Provided an construction of a pretrained transformer capable of simulating gradient descent steps on deep NN.
- Provided theoretical guarantees for the approximation within any given error.
- Extended our analysis to the more practical setting using Softmax-based transformers.

**Outlier-Free SpeechLM with Fast Adaptation and Robust Quantization**

- Proposed SpARQ, to tackle the outlier problem of Speech and Language multi-modal Models.
- Investigated how outliers stemming from cross-modal (speech and text) low-rank adaptation and post-training quantization stages affect the performance of the current Speech and Language multi-modal Models.
- Provided the expressive guarantee of Low-Rank Adaption for modified transformer model.

**Making Genomic Foundation Models more Foundational Requires Outlier Removal: A Case Study on DNABERT-2**
- Introduced an efficient genomic foundation model optimized for accessibility and adaptability.
- Provided the expressive guarantee of Low-Rank Adaption for modified transformer model.
- Identified the conditions for the existence of low-rank adapters.

**Graphics & Geometric Computing Laboratory, University of Science and Technology of China**
*Thesis Research, with Prof.* Falai Chen                                                                 Sept 2023 - Aug 2024
**Boolean operations on generalized polygons with arcs**
- Extended the Vatti algorithm to enable plane Boolean operations on polygons with arcs.
- Implemented and compared multiple algorithms and encapsulated the extended Vatti algorithm into a robust program.

## HONORS

- **Anhui Province University Student Mathematics Competition Second Prize**                    2021
- **USTC Outstanding Freshman Scholarship (Top 25%)**                                           2021
- **USTC Outstanding Student Scholarship (Bronze, Top 20%)**                                    2022

## SKILLS

- **Programming Languages:** C/C++, Python, LaTeX.
- **Tool:** Pytorch, Tensorflow, Github.
- **Mathematical Skills:** Probability, Statistics, Linear Algebra, ODEs, PDEs, Mathematical Analysis, Complex Analysis.
- **Languages:** TOEFL score: 101 (R:27, L:28, S:22, W:24)