SHUNING SHANG

Zhejiang University ◇ Hangzhou, Zhejiang, 310027, P.R.China
(+86) 158 6916 0623 ◇ shuningshang@gmail.com ◇ https://nooraovo.github.io/

EDUCATION

Zhejiang University

B.Eng in Computer Science(Honors Program "Mixed Class") *Aug, 2021 – Jul, 2025 (expected)*

- **GPA: 3.99/4.0, Ranking: 1^{st}/304**
- **Minor** in Statistics

Core Courses

- **Mathematics**: Calculus (4.0/4.0), Linear Algebra (4.0/4.0), Mathematical Analysis (4.0/4.0), Probability Theory (4.0/4.0), Discrete Mathematics (4.0/4.0), Mathematical Statistics (4.0/4.0), Computation and Optimization for Statistics(4.0/4.0)
- **CS**: Lectures on C Programming (4.0/4.0), Fundamental of Data Structures (4.0/4.0), Advanced Data Structure and Algorithm Analysis (4.0/4.0), Object-Oriented Programming (4.0/4.0), Modern Statistical Prediction and Machine Learning(4.0/4.0)

University of California, Berkeley, CA, U.S.

International Exchange Student in Spring Semester *Jan, 2024 – May, 2024*

University of California, Los Angeles, CA, U.S.

International Exchange Student in Summer Session *Aug, 2023 – Sep, 2023*

SCHOLARSHIPS & AWARDS

- **2022,2023 National Scholarship**(Highest scholarship awarded by the Chinese government, $<$ **0.1%**)
- **2022,2023 The First Prize** Scholarship of Zhejiang University (Highest scholarship awarded by Zhejiang University, $<$ **3%**)
- **2022 Top Ten** First-year Students of 2022 (**10/731**)

RESEARCH INTEREST

Machine Learning Theory. Science of deep learning and foundational models. optimization. High-dimensional statistics.

PUBLICATIONS & MANUSCRIPTS

1 Roey Magen*, **Shuning Shang***, Zhiwei Xu, Spencer Frei, Wei Hu, Gal Vardi. Benign Overfitting in Single-Head Attention. Submitted to *2025 The Thirteenth International Conference on Learning Representations (ICLR)*. Under review. [arXiv]

2 **Shuning Shang**, Xuran Meng, Yuan Cao, Difan Zou. Initialization Matters: On the Benign Overfitting of Two-Layer ReLU CNN with Fully Trainable Layers. Submitted to *Journal of Machine Learning Research(JMLR)*. Under review. [arXiv]

RESEARCH EXPERIENCE

The University of Michigan, MI, US
Department of Computer Science and Engineering *May, 2024 – Present*
Research Assistant, Advisor: Prof. Wei Hu
Researcg Topic: Benign Overfitting in Single-Head Attention Model

- Analyzed the dynamics of attention mechanisms on training signal-noise data models using max-margin interpolators, from both empirical and theoretical perspectives.
- Investigated benign overfitting phenomenon in attention models by extending the "Optimal Token Selection" mechanism to non-asymptotic scenarios, enhancing its practical applicability.
- Established a precise signal-to-noise ratio (SNR) threshold to distinguish between benign and harmful overfitting, and estimated the corresponding test errors.

The University of Hong Kong, HongKong, P.R.China
Department of Computer Science *Jul, 2023 – Present*
Research Assistant, Advisor: Prof. Difan Zou
Research Topic: Two-layer ReLU CNN with Fully Trainable Weights

- Extended two-layer CNN models to a more practical setting where both layers are trainable, revealing that the initialization scale of the output layer is crucial to training dynamics.
- Applied two-stage analysis method to better track the dynamic of the training process and discovered a phase transition occurs when the output layer initialization scale is small.
- Characterized benign overfitting behavior in two-layer CNNs and established nearly matching upper and lower bounds on the test error.

SKILLS

Programming	Python, PyTorch, C/C++, MATLAB, LaTeX, HTML, etc.
Language	TOEFL iBT 108/120 (Reading 26, Listening 28, Speaking 25, Writing 29)