

Sin-Han Yang

[✉ sinhan.yang.tw@gmail.com](mailto:sinhan.yang.tw@gmail.com) | [🏠 sinhanyang.github.io](https://sinhanyang.github.io) | [🐦 SinHanYang](https://twitter.com/SinHanYang) | [in sinhanyang](https://www.linkedin.com/in/sinhanyang) | [📷 Sin-Han Yang](https://www.instagram.com/Sin-HanYang)

Research Interests

Driven by curiosity to explore AI systems, seeking deeper understanding and actionable insights. So far, I have explored algorithms, calibration, and phenomena in AI systems. My long-term goal is to make AI beneficial to everyone.

Education

National Taiwan University (NTU)

B.S. in *Computer Science and Information Engineering and Physics*

Taipei, Taiwan

Sep. 2019 - Jun. 2024

University of Birmingham

Exchange Student in Computer Science

Birmingham, UK

Sep. 2022 - Jan. 2023

Selected Experiences

Appier

Taipei, Taiwan

AI Research Scientist

Sep. 2025 - NOW

AI Research Scientist Intern

May 2025 - Sep. 2025

- Advisor: [Prof. Shao-Hua Sun](#), [Prof. Hung-yi Lee](#), [Prof. Yun-Nung \(Vivian\) Chen](#)
- **Research 1** : Rethinking the standard model merging pipeline **[P1]**
 - Identified an implicit regularization in the model merging pipeline caused by coefficient optimization.
 - Removed this regularization simply via gradient descent or random sampling of model weights, consistently improving model merging performance.
 - Conducted controlled experiments to investigate the role of the bias-variance tradeoff in model merging.
- **Research 2** : A new LLM confidence calibration framework **[P3]**
 - Proposed a new confidence calibration framework that addresses LLM stochasticity; theoretically and empirically demonstrated its differences from the traditional confidence calibration framework.
 - Demonstrated the practical utility of the proposed calibration framework across several downstream applications.
 - Benchmarked confidence estimators under the proposed framework to understand its practical feasibility.
- **Product 1** : Developing LLM evaluation for enterprise-scale model routing.
- **Product 2** : Automated sensitive product detection in advertisements by finetuning a pretrained multimodal model.

RIKEN Center for Advanced Intelligence Project

Tokyo, Japan

Technical Staff I (Research Assistant)

Aug. 2024 - Mar. 2025

- Advisor: [Prof. Emtiyaz Khan](#)
- **Research** : Variational learning induces adaptive label smoothing **[P4]**
 - Derived the implicit label noise in Variational Learning (VL), which is instance-specific.
 - Empirically showed that VL's adaptive label noise is similar to an existing adaptive label smoothing method.
 - Evaluated VL on noisy-label classification tasks and showed that it outperforms traditional label smoothing.
- Maintained the [team website](#).

Natural Language Processing Laboratory, NTU

Taipei, Taiwan

Undergraduate Researcher

Nov. 2021 - Feb. 2024

- Advisor: [Prof. Hsin-Hsi Chen](#)
- **Research** : An interpretable model for fake news detection **[P6]**
 - Developed a new attention-based architecture for interpretability.
 - Empirically showed that the new architecture outperforms baselines in standard benchmarks.
 - Utilized model's interpretability to analyze the key words and sentences for final predictions.

Papers

*: equal contribution, †: equal advisory contribution

- [P1] **Sin-Han Yang**^{*}, Shih-Cheng Huang^{*}, Chieh-Yeh Lin, Yun-Nung Chen[†], Shao-Hua Sun[†], Hung-yi Lee[†], [A Broader Look at Model Merging: Rethinking Implicit Regularization Induced By Task Arithmetic](#), *arXiv preprint arXiv:2610.07990*, 2026.
- [P2] Zhi Rui Tam, Chao-Chung Wu, **Sin-Han Yang**, Peyton Ku, Brendan Kuang, Tzu-Ting Hsieh, Min-Fang Hsu, Fang-Ling Tsai, Yun-Nung Chen, Wei-Chiu Ma, Chieh-Yen Lin, [TasteRoute: Personalized Routing for Video Generation](#), *arXiv preprint arXiv:2610.05896*, 2026.
- [P3] **Sin-Han Yang**^{*}, Cheng-Kuang Wu^{*}, Chieh-Yeh Lin, Yun-Nung Chen, Hung-yi Lee, Shao-Hua Sun, [On Calibration of Large Language Models: From Response To Capability](#), *arXiv preprint arXiv:2602.13540*, 2026.
- [P4] **Sin-Han Yang**, Zhedong Liu, Gian Maria Marconi, Mohammad Emtiyaz Khan, [Variational Learning Induces Adaptive Label Smoothing](#), *The 7th Symposium on Advances in Approximate Bayesian Inference – Workshop Track*, 2025. (**Best Paper Runners-Up**)
- [P5] **Sin-Han Yang**, Tuomas Oikarinen, Tsui-Wei Weng, [Concept-Driven Continual Learning](#), *Transactions on Machine Learning Research (TMLR)*, 2024.
- [P6] **Sin-Han Yang**, Chung-Chi Chen, Hen-Hsen Huang and Hsin-Hsi Chen, [Entity-Aware Dual Co-Attention Network for Fake News Detection](#), *Findings of the Association for Computational Linguistics: EACL*, 2023.

Honors & Awards

Best Paper Runners-Up	The 7th Symposium on Advances in Approximate Bayesian Inference Workshop Track	Singapore
		2025
Best Reviewers Award	The 28th International Conference on Artificial Intelligence and Statistics (AISTATS)	
		2025
Appier Best R&D Award and CSIE 3rd Place	NTU CSIE Undergraduate Research Exhibition	Taipei, Taiwan
		2023
College Student Research Scholarship	National Science Council (NSC)	Taipei, Taiwan
• NSC scholarship for students based on written research proposal		2023
NTU Y.L.LIN Scholarship		Taipei, Taiwan
• For exchange students		2022
NTU Dean's List Award		Taipei, Taiwan
		2021
Final Selection and Training Camp	IPhO (International Physics Olympics) Taiwan team	Taipei, Taiwan
		2018

Professional Activities

- **Reviewer** ICLR 2025-2027, NeurIPS 2026, AISTATS 2025
IEEE Transactions on Knowledge and Data Engineering
NeurIPS 2024 Safe Generative AI Workshop
- **Talk** LLM Confidence Calibration, [NTU Machine Learning 2026 Spring](#), May 15th 2026

Skills

- **Programming** Python, MATLAB, C/C++
- **Tools** PyTorch, WandB, Git, LaTeX
- **Languages** Mandarin (Native), Taiwanese (Native), English (TOEFL 5.5/6.0), Japanese (JLPT N3)