ANDREW CHING-YUAN BAI

andrewbai@cs.ucla.edu

I am currently a third year PhD student at UCLA Computer Science Department advised by Prof. Cho-Jui Hsieh. My research interests involve understanding how machine learning models function. My current focus is on developing practical model interpretation methods to facilitate adoption of machine learning models in critical domains, where explaining *how* a model make decisions is crucial.

EDUCATION

University of California, Los Angeles (Los Angeles, USA) Sep 2021 – Exp. Jun 2027
Ph.D. in Computer Science (advised by Prof Cho-Jui Hsieh).

National Taiwan University (Taipei, Taiwan) Sep 2016 – Jan 2021
B.S. in Computer Science and Information Engineering. (GPA: 4.2/4.3)
Minor in Mechanical Engineering.

RWTH Aachen (Aachen, Germany) Oct 2019 – Mar 2020
Undergraduate exchange student

PUBLICATIONS

- A. Bai, C.-K. Yeh, C.-J. Hsieh, and A. Taly. **Which Pretrain Samples to Review when Finetuning Pretrained Models?** Under submission.
- A. Bai, C.-K. Yeh, P. K. Ravikumar, N. Lin, and C.-J. Hsieh. **Concept Gradient: Concept-based Interpretation Without Linear Assumption.** In *Proceedings of the 11th International Conference on Learning Representations (ICLR)*, May 2023.
- A. Bai, H.-T. Lin, C. Raffel, and W. Kan. **On training sample memorization: Lessons from benchmarking generative modeling with a large-scale competition.** In *Proceedings of the 27th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, Aug 2021.
- S.-L. Wu*, C.-Y. Bai*, K.-C. Chang, Y.-T. Hsieh, C. Huang, C.-W. Lin, E. Kang, and Q. Zhu. **Efficient system verification with multiple weakly-hard constraints for runtime monitoring.** In *Proceedings of the International Conference on Runtime Verification*, Oct 2020.
- C.-Y. Bai, B.-F. Chen, and H.-T. Lin. **Benchmarking Tropical Cyclone Rapid Intensification with Satellite Images and Attention-based Deep Models.** In *Proceedings of the The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML/PKDD)*, Sep 2020.

WORK EXPERIENCE

Google Cloud (Remote) Apr 2023 – Aug 2023
Student Researcher

- Designed finetuning schemes to mitigate catastrophic forgetting without incurring additional computation overhead by identifing "useful samples" with loss filtering.
- Investigated targeted active learning setting on improving specific data slice while maintaining performance on the overall data distribution.

Amazon (Palo Alto, California) Jun 2022 – Sep 2022
Applied Scientist Intern

* These authors contributed equally to this work

· Implemented and optimized factorization machine training and inferencing in C++, increasing the training speed by 43x compared to `libffm`.
· Investigated the impact of replacing inner product search with cross-attention methods (e.g. factorization machine) in two-tower deep neural network retrieval models.

WorldQuant Research LLC, Taiwan Branch (Taipei, Taiwan) Aug 2020 – Feb 2021
Quantitative Analysis Research Intern

· Predicted close-form expression for analyst data using fundementals by iterative feature selection with deep neural networks, achieving 95%+ accuracy.
· Designed techniques for statistical learning to rank under non-stationary data distribution.

Taiwan Artificial Intelligence Lab (Taipei, Taiwan) Jan 2019 – Aug 2019
Machine Learning Research Intern

· Adapted sparse RGB natural image segmentation model (CurveGCN) across domain to monochrome CT scans for detecting lung nodules.
· Built full-stack infrastructure to allow interactive labeling of segmentation tasks for medical experts and deployed in the largest hospital in Taiwan.

RESEARCH EXPERIENCE

Dept. of Computer Science, UCLA (Los Angeles, CA) Sep 2021 – Present
Graduate Student Researcher (advised by Prof. Cho-Jui Hsieh)

· Designed concept-based interpretability methods for general differentiable models (e.g. neural networks) by propagating gradients through shared input feature representation.

Dept. of Computer Science and Engineering, NTU (Taipei, Taiwan) Jun 2018 – Jan 2021
Research Assistant (advised by Prof. Hsuan-Tien Lin)

· Collaborated with Kaggle (now a subsidiary of Google) on generative modeling metric design and held the first-ever public large-scale generative modeling competition with 900+ participating teams.
· Designed the first algorithm to reduce training sample memorization during Generative Advesarial Networks (GANs) training with rejection sampling.
· Designed the first deep neural network model to predict tropical cyclone rapid intensification using satellite image data and establish strong baseline for our proposed benchmark.

TEACHING AND MENTORSHIP

Undergraduate Research Mentorship Jan 2023 – Present

· Volunteered to mentor 8 undergraduate students on investigating the memorization of diffusion models.
· TAed COM SCI 180 (Introduction to Algorithms and Complexity) for two academic quarters.

GRANTS AND FELLOWSHIP

Kaggle, Alphabet Inc. Jul 2019 – Aug 2019
Generative Adversarial Network Research Grant

· Funding for holding the Kaggle Generative Dog Images competition

Taiwan Ministry of Science and Technology (MOST) Jul 2019 – Feb 2020
MOST Research Grant for University Students

· Funding for tropical cyclone rapid intensification prediction resesarch