Yao-Yuan Yang

9500 Gilman Dr, La Jolla, CA 92093 | Phone: +1-858-900-4376
http://yyyang.me | https://github.com/yangarbiter | yaoyuany@deepmind.com

Skills

- 8+ years of machine learning experiences, knowledgeable in state-of-the-art algorithms and capable of developing new ones for specific application
- 10+ years of programming experiences on Python, C, C++ on Linux, Unix, Windows system
- Experienced in the Python package development as well as its interface with C for faster implementation, capable of modifying existing packages or developing a new one

Education

University of California San Diego (UCSD) 09.2017 – 06.2022

Ph.D. in Computer Science and Engineering
Dissertation: A Principled Approach to Trustworthy Machine Learning [link]
Committee members: Kamalika Chaudhuri, Sanjoy Dasgupta, Sicun Gao, Tara Javidi, Lawrence Saul

National Taiwan University (NTU) 09.2012 – 06.2016

B.S. in Computer Science and Information Engineering

Research Interest

My research focuses on topics in trustworthy machine learning, including adversarial examples, interpretability, out-of-distribution data, spurious correlation, and distribution robustness. Grounded on both theoretical and empirical work, my research offers insights into how machine learning models work and how they can be applied to design trustworthy algorithms in practical domains, such as safety and security.

Keywords: trustworthy machine learning, adversarial examples, out-of-distribution, spurious correlation

Major Experiences

Research Engineer Google DeepMind, CA 08.2022 - Present

- Developing methods for multi-lingual transfer learning in low-data regime.

Graduate Student Researcher UCSD, CA (Advisor: Kamalika Chaudhuri) 09.2017 - 06.2022

- Researched when and how spurious correlations are learned in machine learning models, the consequences of learning rare spurious correlations, and methods (including regularization, data deletion, and outlier detection) to mitigate spurious correlations
- Researched methods to quantitatively measure creativity
- Studied the effect of adversarial robust training algorithms on out-of-distribution data; discovered that robust neural networks tend to generalize better on some natural manifold
- Identified that the trade-off between adversarial robustness and accuracy is not an intrinsic property for many datasets, but instead, is caused by the increase in the generalization gap when applying robust training
- Designed adversarial attack/defense algorithms that work well across many non-parametric classifiers including decision trees, k-nearest neighbors, and random forest; the implementation of the algorithm is accelerated with customized C-extension
- Implemented a machine learning approach for Python error localization

Machine Learning Intern (`PyTorch` team) Facebook, Remote 06.2021 - 09.2021

- Contributed $30+$ commits and $6000+$ lines of code in `torchaudio`
- Implemented a text-to-speech pipeline including text-preprocessing with phonemes, spectrogram generation with Tacotron2, and time-domain conversion with WaveRNN with `PyTorch`
- Benchmarked and compared the performance of multiple operations and models with other related works

Ph.D. Intern (video intelligence team) Yahoo, NY 06.2018 - 09.2018

- Designed the pipeline to insert ads realistically into a video
- Developed an algorithm to identify good places in a video to insert ads, and adopted Mask R-CNN for the segmentation of the foreground and background to perform the insertion

Research Assistant NTU, Taipei, Taiwan (Advisor: Hsuan-Tien Lin) 09.2013 - 06.2016

- Developed a novel multi-label classification (MLC) algorithm that utilizes the memory structure within recurrent neural networks to extract the hidden correlation between labels (with `keras` and `tensorflow`)
- Proposed the first cost-sensitive error-correcting-code, and adopted it for solving MLC and active learning problems
- Led the development of the open-source package (with $700+$ stars), `libact`, which provides a unified interface for active learning algorithms in Python; setup the infrastructures including continuous integration, documentation generation, and PyPI installation; contributed $350+$ commits and $25,000+$ lines of code

Person Identification with EEG NTU, Taipei, Taiwan (Advisor: Tsung-Ren Huang) 02.2014 - 06.2016

- Designed and carried out $50+$ sessions of Electroencephalography (EEG) experiments; studied the characteristic of EEG including neural oscillations and artifacts that contaminates the data
- Conducted twin study to understand the properties of using EEG as biometrics, such as how the personality and value of a person correlate with the identification rate

Gesture recognition through electromyography (EMG) NTU, Taipei, Taiwan 04.2014 - 06.2014

- Built a device that measures muscle signal and sends the signal to a laptop
- Applied support vector machine for the recognition of $4+$ hand gestures

Intern IKV-Tech, Taipei, Taiwan 07.2013 - 08.2013

- Implemented FAT32 file system component through UCB protocol for Windows and embedded device (ARM922t Mcu); designed C APIs for encrypted file storage between PC and the embedded device

Cryptanalysis on Mifare Crypto-1 NTU, Taipei, Taiwan (Advisor: Chen-Mou Cheng) 02.2013 - 07.2013

- Designed and analyzed cryptanalytic time-memory trade-off (rainbow table) on the Crypto-1 encryption
- Implemented Crypto-1 rainbow table generation with OpenCL and optimized it on NVIDIA GPU, which results in 10 times faster implementation

PUBLICATIONS (* EQUAL CONTRIBUTION)

Preprints

- **Y.-Y. Yang**, C.-N. Chou, K. Chaudhuri. Understanding Rare Spurious Correlations in Neural Networks, ICML SCIS Workshop, 2022 (link)
- **Y.-Y. Yang**, S.-C. Lee, Y.-A. Chung, T.-E. Wu, S.-A. Chen, H.-T. Lin. libact: Pool-based Active Learning in Python. 2017 (link)

Conference Papers

- B. Kulynych*, **Y.-Y. Yang***, Y. Yu, J. Błasiok, P. Nakkiran. What You See is What You Get: Principled Deep Learning via Distributional Generalization, NeurIPS; ICML PODS Workshop, 2022 (link)
- **Y.-Y. Yang***, M. Hira*, Z. Ni*, A. Chourdia, A. Astafurov, C. Chen, C.-F. Yeh, C. Puhrsch, D. Pollack, D. Genzel, D. Greenberg, E. Z. Yang, J. Lian, J. Mahadeokar, J. Hwang, J. Chen, P. Goldsborough, P. Roy, S. Narenthiran, S. Watanabe, S. Chintala, V. Quenneville-Bélair, Y. Shi, TorchAudio: Building Blocks for Audio and Speech Processing, ICASSP, 2022 (link)
- A. H.-C. Hwang, C. Y. Wang, **Y.-Y. Yang**, and A. S. Won Hide and seek: Choices of Virtual Backgrounds in Video Chats and Their Effects on Perception, CSCW, 2021 (link)
- M. Moshkovitz, **Y.-Y. Yang**, and K. Chaudhuri. Connecting Interpretability and Robustness in Decision Trees through Separation, ICML, 2021 (link)
- **Y.-Y. Yang***, C. Rashtchian*, H. Zhang, R. Salakhutdinov, K. Chaudhuri. A Closer Look at Accuracy vs. Robustness, NeurIPS 2020; ICML UDL workshop 2020 (spotlight presentation) (link)
- **Y.-Y. Yang***, C. Rashtchian*, Y. Wang, K. Chaudhuri. Robustness for Non-Parametric Classification: A Generic Attack and Defense, AISTATS, 2020 (link)

- B. Cosman, M. Endres, G. Sakkas, L. Medvinsky, **Y.-Y. Yang**, R. Jhala, K. Chaudhuri, W. Weimer, PABLO: Helping Novices Debug Python Code Through Data-Driven Fault Localization, SIGCSE 2020 (link)
- **Y.-Y. Yang**, Y.-A. Lin, H.-M. Chu, H.-T. Lin. Deep Learning with a Rethinking Structure for Multi-label Classification, ACML, 2019. (link)
- **Y.-Y. Yang**, K.-H. Huang, C.-W. Chang, H.-T. Lin. Cost-Sensitive Reference Pair Encoding for Multi-Label Learning, PAKDD, 2018 (link)

Journal Articles

- **Y.-Y. Yang**, C. Rashtchian, R. Salakhutdinov, K. Chaudhuri. Probing Predictions on OOD Images via Nearest Categories. Transactions on Machine Learning Research, 2023 (link)
- **Y.-Y. Yang**, A. H.-C. Hwang, C.-T. Wu, T.-R. Huang. Person-identifying brainprints are stably embedded in EEG mindprints, Scientific Reports, 2022 (link)

SELECTED TALKS

What You See is What You Get: Distributional Generalization for Algorithm Design in Deep Learning 03.2022

Apple, Virtual

CSE Research Open House 2022, UCSD, La Jolla

A Closer Look at Accuracy vs. Robustness 06.2020-10.2021

INFORMS annual meeting, Anaheim, CA

SoCal ML Symposium, Virtual

G-Research, Virtual

NeurIPS 2020, Virtual

ICML UDL 2020, Virtual

In- and Out-of-Distribution Generalization Properties of Adversarially Robust Models 08.2021

Science of Deep Learning, Facebook AI Research, Virtual

Close Category Generalization for Out-of-Distribution Classification 03.2021

SoCal ML Symposium, Virtual

Robustness for Non-Parametric Classification: A Generic Attack and Defense 08.2020

AISTATS 2020, Virtual

Deep Learning with a Rethinking Structure for Multi-label Classification 11.2019

ACML, Nagoya, Japan

Cost-Sensitive Reference Pair Encoding for Multi-Label Learning 06.2018

PAKDD, Melbourne, Australia

Near-uniform Aggregation of Gradient Boosting Machines for KDD Cup 2015 08.2015

KDD, Sydney, Australia

SERVICE

UCSD ML Group Blog: `https://ucsdml.github.io/` Chief Editer w/ Cyrus Rashtchian 2020-Present

UCSD CSE PhD Admissions Student Committee Committee Member 2019

PEER REVIEW

International Conference on Learning Representations Reviewer 2022

Applied Artificial Intelligence Reviewer 2022

International Conference on Machine Learning Reviewer 2022

IEEE International Symposium on Information Theory Reviewer 2022

International Conference on Artificial Intelligence and Statistics Reviewer 2022

Conference on Neural Information Processing Systems Reviewer 2021

International Conference on Artificial Intelligence and Statistics Reviewer 2021

International Conference on Machine Learning Reviewer 2021

IEEE Transactions on Neural Networks and Learning Systems Reviewer 2020

International Conference on Artificial Intelligence and Statistics Reviewer 2020

International Conference on Machine Learning Reviewer 2020

IEEE Transactions on Neural Networks and Learning Systems Reviewer 2019

Journal of Machine Learning Research Reviewer 2019

IEEE Transactions on Pattern Analysis and Machine Intelligence Reviewer 2019

Journal of Machine Learning Research Reviewer 2018

TEACHING

CSE 151A: Introduction to AI: A Statistical Approach Teaching Assistant Winter 2021/2020; Spring 2018/2019

AWARDS

Fourth Place (out of 800+ teams), Predicting MOOC dropouts (ACM KDD Cup) 08.2015

with NTU team, ACM KDD

- Led the team in the ensemble of 70+ models, and contributed to feature generation and model tuning

First Place, Big data analytics for semiconductor manufacturing 02.2015

with P.-H. Chu and Y.-A. Chung, Taiwan Semiconductor Manufacturing Company

- Proposed a solution for detecting failures in semiconductor manufacturing process using ridge regression