

# Chun-Hao Chang

+1-647-657-6119 • chkchang21@gmail.com • [www.cs.toronto.edu/~kingsley](http://www.cs.toronto.edu/~kingsley)

---

**Overview** I'm interested in Applied ML Research. My PhD's works are about Machine Learning in **Interpretability (XAI)**, **Robustness**, and applied reinforcement learning in **Healthcare**. I interned in Google Cloud on **Anomaly Detection**, Microsoft, and Facebook on **Causal** Ads ranking. I was a software engineer for 2 years.

---

## Education

- **Ph.D.**, Dept. of *Computer Science*, University of Toronto *Feb. 2018 – Present*
  - **M.S.**, Dept. of *Computer Science*, University of Toronto *Sep. 2016 – Jan. 2018*
  - **B.S.**, Dept. of *Electrical Engineering and Life Science*, National Taiwan University *Sep. 2010 – June. 2015*
- 

## Main Research Works

- **[Interpretability, Healthcare] Extracting Clinician's Goals by What-if Interpretable Modeling** [link](#)  
Submitted to 2022 AISTATS  
**Chun-Hao Chang**, George Alexandru Adam, Rich Caruana, Anna Goldenberg  
**TLDR:** We extract clinicians' treatment goals by interpretable GAM modeling and what-if reasoning
  - Recovered clinicians' treatment goals for hypotensive patients in a real-world ICU data.
  - Modeled the treatment reward by the future counterfactual outcomes combined with an interpretable model - Generalized Additive Model (GAM).
  - The reward matches several clinical guidelines while previous linear model often contradicts them.
- **[Interpretability] NODE-GAM: Neural Generalized Additive Model for Interpretable Deep Learning** [link](#)  
Submitted to 2022 ICLR  
**Chun-Hao Chang**, Rich Caruana, Anna Goldenberg  
**TLDR:** We develop a deep-learning version of Generalized Additive Model (GAM) and  $GA^2M$ .
  - Proposed a new deep-learning version of Generalized Additive Model (GAM) and  $GA^2M$  - an interpretable model class. Our models have higher accuracy than traditional GAMs on large datasets.
  - Demonstrated our deep-learning GAM is interpretable.
  - When labeled data is limited, our model outperforms other GAMs by self-supervised pretraining.
- **[Interpretability] How Interpretable and Trustworthy are GAMs?** [link](#)  
Published in 2021 SIGKDD (Conference on Knowledge Discovery and Data Mining)  
**Chun-Hao Chang**, Sarah Tan, Ben Lengerich, Anna Goldenberg, Rich Caruana  
**TLDR:** We compared total 9 different GAMs and showed which GAM is more trustworthy.
  - Compared total 9 kinds of inherently explainable models (Generalized Additive Models, GAMs) in terms of human explainability, inductive biases and performances across multiple real-world datasets.
  - Found GAMs with high feature sparsity (only using a few variables to make predictions) tend to hide patterns in the data and be unfair to rare subpopulations.
  - Found inductive bias plays a crucial role in model explanations and tree-based GAMs are recommended for their low feature sparsity and high data fidelity which make them more trustworthy.
- **[Robustness] Towards robust imaging model by counterfactual generation and adversarial generation** [link](#)  
Published in 2021 CVPR  
**Chun-Hao Chang**, George Alexandru Adam, Anna Goldenberg  
**TLDR:** We proposed two data augmentations to make our models robust to data distribution shifts.
  - Proposed counterfactual and adversarial data augmentations to make imaging models robust to spurious correlation that may not hold in test time.

- Improved both generalization accuracy and out-of-class detection in multiple challenging datasets whose distributions differ from the training distributions.
- **[Robustness, Healthcare] Hidden Risks of Machine Learning Applied to Healthcare: Unintended Feedback Loops Between Models and Future Data Causing Model Degradation** [link](#)  
Published in *2020 MLHC Machine Learning for Healthcare Conference*  
George Alexandru Adam, **Chun-Hao Kingsley Chang**, Benjamin Haibe-Kains, Anna Goldenberg  
**TLDR:** We characterize a feedback loop problem that clinicians changing their decisions based on an imperfect ML system that changes the future data distribution.
  - Introduced the feedback loop problem that the future data are corrupted by human interventions caused by model's predictions; If continually updated from new data, models degrade almost exponentially.
  - Proposed various sample removal techniques to minimize the effects of the feedback loop.
- **[Interpretability] Explaining Image Classifiers by Counterfactual Generation** [link](#)  
Published in *2019 ICLR International Conference on Learning Representations*  
**Chun-Hao Chang**, Elliot Creager, Anna Goldenberg, David Duvenaud  
**TLDR:** We propose using generative models to ask counterfactual questions to interpret a black-box model.
  - Proposed a novel framework that explains neural network decisions by efficiently searching for important features that respect the distribution of the input data by generative model.
  - Our method produces more compact and relevant pixels of explanations with fewer artifacts; demonstrated our method's superiority over several baselines quantitatively on ImageNet.
- **[Healthcare] Dynamic Measurement Scheduling for Events Forecasting by Deep RL** [link](#)  
Published in *2019 ICML International Conference of Machine Learning*  
**Chun-Hao Chang\***, Mingjie Mai\*, Anna Goldenberg  
**TLDR:** We propose a reinforcement learning approach to better allocate healthcare resources for measurements.
  - Designed an automatic diagnostic machine in the hospital to prioritize the measurements of patients by Deep Q network (DQN).
  - To handle the large combinatorial action spaces ( $2^{40}$ ) in the multi-action setting, we developed a sequential action approach to learn it more efficiently; we show it outperforms other baselines.
  - Applied to a real-world clinical ICU dataset (MIMIC3) and increase 3 times informative gain or reduce 31% measurement costs compared to physicians.

## Main Work Experience

- **Student Researcher**, Google Cloud AI team *Aug. 2021 – Jan. 2022*  
Host: Jinsung Yoon
- **Research Intern**, Microsoft Research AI team *Jun. 2019 – Sep. 2019*  
Host: Rich Caruana
- **PhD Machine Learning Intern**, Ads Ranking team, Facebook UK *Jun. 2018 – Aug. 2018*
  - Modeled counterfactual inference of the ads data to measure lift effects of ads exposure on users
  - Visualized and analyzed the data and removed the outliers using SQL scripts and C++ programming.
  - Investigated various causal approach such as nearest neighbor, causal trees and multi-task learning.
- **Machine Learning Teaching Assistant**, University of Toronto / Vector Institute *6 times from 2017-2020*  
*Vector Certificate, CSC412 Neural Networks, CSC411 Machine Learning, CSC311 Intro to ML*
- **Software Engineer**, Political Warfare Bureau, R.O.C. Armed Forces *Aug. 2015 – July 2016*
  - Designed and programmed a search engine (ElasticSearch) with distributed computation to speed up.
  - Led a team of 5 people including technical and non-technical personnel (designer/front-end engineers)
- **Android Engineer Intern**, TripNotice Inc. *Jan. 2014 – Mar. 2015*
  - Independently designed and implemented a trip-booking Android app (TripNotice) in Java from scratch.
- **Conference Reviewers** (ICLR, NeurIPS, ICML, CVPR, AISTATS) *8 times from 2018*
- **Appraiser on UofT CS Graduate Admissions** *2019, 2020, 2021*

## Miscellaneous

- **MiniMBA course completion, Sickkids-GMCA miniMBA program** *Jan 2017*  
Took a 12-week class for MBA classes and won the 2nd place out of 12 groups in the final presentation.
  - **Service:** Activities director, Toronto Taiwanese Graduate Student Association *2017, 2020*
- 

## References

- **Anna Goldenberg**, Associate Professor  
Department of Computer Science at the University of Toronto  
Relationship: PhD advisor  
anna.goldenberg@utoronto.ca
- **Rich Caruana**, Senior Principal Researcher  
Microsoft Research  
Relationship: intern host and a research collaborator  
rcaruana@microsoft.com
- **David Duvenaud**, Assistant Professor  
Department of Computer Science at the University of Toronto  
Relationship: research collaborator  
duvenaud@cs.toronto.edu
- **Marzyeh Ghassemi**, Assistant Professor  
MIT in Electrical Engineering and Computer Science (EECS) and Institute for Medical Engineering (IMES)  
Relationship: the thesis committee member  
mgassem@mit.edu