Chen Dan
PhD Student, Department of Computer Science, Carnegie Mellon University
Homepage: `http://cs.cmu.edu/~cdan` Tel: (+1) 412-961-1405
Email: cdan@alumni.cmu.edu

Research

My research interest is in machine learning theory and algorithms, with a focus on:

- Statistical Learning Theory
- Robustness in Machine Learning

Work

Toyota Technological Institute at Chicago *2023 - Present*
Postdoctral Researcher working with Prof. Avrim Blum

Education

Carnegie Mellon University, Computer Science Department *2016 - 2022*
Ph.D in Computer Science, Advised by Prof. Pradeep Ravikumar
Thesis: Statistical Learning Under Adversarial Distribution Shift
Committee: Pradeep Ravikumar (Chair), Avrim Blum, Zico Kolter, Yuting Wei, Zack Lipton

Toyota Technological Institute at Chicago *Summer 2018*
Visiting Student hosted by Prof. Avrim Blum

Peking University, School of EECS *2012 - 2016*
Bachelor of Science summa cum laude in Machine Intelligence, Thesis Advisor: Prof. Liwei Wang

- Overall GPA 3.72/4.0, Major GPA 3.75/4.0; Ranking 2/46
- **Bachelor Thesis:** On Low Rank Approximation of Binary Matrices
  Top 10 Bachelor Thesis Award in School of EECS, Peking University

Skills

C/C++, Matlab, Python, LaTeX

Publications

* Alphabetical Order or Equal Contribution

1. **Chen Dan**, Yuting Wei, Pradeep Ravikumar
   Sharp Statistical Guarantees for Adversarially Robust Gaussian Classification
   37th International Conference on Machine Learning (ICML 2020)
2. Runtian Zhai*, **Chen Dan***, Pradeep Ravikumar, Zico Kolter
   DORO: Distributional and Outlier Robust Optimization
   38th International Conference on Machine Learning (ICML 2021)
3. Runtian Zhai*, **Chen Dan***, Di He*, Huan Zhang, Boqing Gong, Pradeep Ravikumar, Cho-Jui Hsieh, Liwei Wang
   MACER: Attack-free and Scalable Robust Training via Maximizing Certified Radius
   2020 International Conference on Learning Representations (ICLR 2020)
4. Bryon Aragam, **Chen Dan**, Pradeep Ravikumar, Eric Xing
   Identifiability of Nonparametric Mixture Models and Bayes Optimal Clustering,
   Annals of Statistics 2020
5. **Chen Dan**, Liu Leqi, Bryon Aragam, Pradeep Ravikumar, Eric P. Xing
   The Sample Complexity of Semi-Supervised Learning with Nonparametric Mixture Models
   32nd Conference on Neural Information Processing Systems (NeurIPS 2018)

6. **Chen Dan**, Hong Wang*, Hongyang Zhang*, Yuchen Zhou*, Pradeep Ravikumar
Optimal Analysis of Subset-Selection Based L_p Low Rank Approximation
33rd Conference on Neural Information Processing Systems (NeurIPS 2019)
7. Han Zhao*, **Chen Dan***, Bryon Aragam, Tommi Jaakkola, Geoff Gordon, Pradeep Ravikumar
Fundamental Limits and Tradeoffs in Invariant Representation Learning
Journal of Machine Learning Research (JMLR 2022).
8. Avrim Blum*, **Chen Dan***, Saeed Seddighin*
Learning Complexity of Simulated Annealing
24th International Conference on Artificial Intelligence and Statistics (AISTATS 2021)
9. Haris Angelidakis*, Pranjal Awasthi*, Avrim Blum*, Vaggos Chatziafratis*, **Chen Dan***
Bilu-Linial Stability, Certified Algorithms and the Independent Set Problem
27th Annual European Symposium on Algorithms (ESA 2019)
10. **Chen Dan**, Kristoffer Arnsfelt Hansen, He Jiang, Liwei Wang and Yuchen Zhou,
On Low Rank Approximation of Binary Matrices
43rd International Symposium on Mathematical Foundations of Computer Science (MFCS 2018)
11. Runtian Zhai, **Chen Dan**, Zico Kolter, Pradeep Ravikumar
Understanding Why Generalized Reweighting Does Not Improve Over ERM
2023 International Conference on Learning Representations (ICLR 2023)
12. Runtian Zhai, **Chen Dan**, Arun Sai Suggala, Zico Kolter, Pradeep Ravikumar
Boosted CVaR Classification
35th Conference on Neural Information Processing Systems (NeurIPS 2021)
13. Ziyu Neil Xu, **Chen Dan**, Justin Khim, Pradeep Ravikumar
Class-Weighted Classification: Trade-offs and Robust Approaches
37th International Conference on Machine Learning (ICML 2020)
14. Xun Zheng, **Chen Dan**, Bryon Aragam, Pradeep Ravikumar, Eric P. Xing
Learning Sparse Nonparametric DAGs
23rd International Conference on Artificial Intelligence and Statistics (AISTATS 2020)
15. Runtian Zhai, **Chen Dan**, Zico Kolter, Pradeep Ravikumar
MSG: Margin Sensitive Group Risk
In Submission
16. Runtian Zhai*, Tianle Cai*, Di He, **Chen Dan**, Kun He, John Hopcroft, Liwei Wang,
Adversarially Robust Generalization Just Requires More Unlabeled Data
Preprint, arXiv 1906.00555

Courses

- Advanced Introduction to Machine Learning
- A Theorist's Toolkit
- Statistical Learning Theory
- Convex Optimization
- Statistics Meets Optimization: Randomized Sketching Methods
- Graduate Artificial Intelligence
- Multimedia Database and Data Mining
- Computer Architecture

Teaching

- TA for Convex Optimization (Fall 2019), Instructor: Ryan Tibshirani.
- TA for Practical Data Science (Spring 2021), Instructor: Zico Kolter.

Academic Services

Reviewer of:

- NeurIPS 2019 (Top 50% reviewer), 2020, 2021, 2022
- ICML 2020 (Top 33% reviewer), 2021 (Expert Reviewer), 2022
- ICLR 2021, 2022
- AISTATS 2021, 2022
- AAAI 2021
- ICALP 2021
- SODA 2019
- ITCS 2019
- IEEE Transactions on Information Theory

Student Committee member for PhD Applications,
Computer Science Department, Carnegie Mellon University, 2020.

Awards

- NeurIPS Travel Award, 2019
- Top 10 Bachelor Thesis Award in School of EECS, Peking University, 2016
  (10/320 in School of EECS, 1/46 in Department of Machine Intelligence)
- Outstanding Undergraduate Research Award - Second Prize, Peking University, 2015
  (3/320 in School of EECS, 1/46 in Department of Machine Intelligence)
- May Fourth Scholarship, 2015
- 8508 Alumni Scholarship, 2014
- Suzhou Industrial Park Scholarship, 2013