Cho-Jui Hsieh

404 Westwood Plaza. Engineering VI.
Los Angeles,
CA, 90095-1596

chojui.hsieh@gmail.com
Cell Phone: 5129645633
http://www.cs.ucla.edu/∼chohsieh

Work Experience

- **Assistant Professor, Department of Computer Science**

  University of California, Los Angeles 11/2018–

- **Assistant Professor, Department of Computer Science & Department of Statistics**

  University of California, Davis 10/2015–10/2018

Education

- **Ph.D. Dept. of Computer Science, University of Texas at Austin** 8/2010 – 8/2015

  Advisor: Inderjit Dhillon

- **M.S. Dept. of Computer Science, National Taiwan University** 9/2007 – 6/2009

  Advisor: Chih-Jen Lin

- **B.S., Dept. of Computer Science, National Taiwan University** 9/2003 – 6/2007

- **B.S., Dept. of Mathematics, National Taiwan University** 9/2004 – 6/2007

Research Interests

My main research focus is on developing new algorithms and optimization techniques for large-scale machine learning problems. Currently, I am working on the following problems:

- Adversarial machine learning.
- Fast and scalable algorithms for large-scale training.
- Fast Prediction and Model Compression for big ML models.
- Recommender systems.
- AutoML: neural architecture search and hyper-parameter tuning.

Selected Honors

- Samsung AI Researcher of the Year, 2020
  (Five top AI researchers under 35 selected by committee led by Prof. Yoshua Bengio)
- Best Paper Award Finalist, Recsys, 2020:
- Best Paper Award Candidate, ICDM, 2019.
- Best Student Paper Finalist, SC, 2019.
- Best Paper Award, ICPP 2018.
- Best Paper Award Finalist, AISec 2017.
- Best Paper Award, ICDM 2012.
- Best Research Paper Award, KDD 2010.
- IBM PhD Fellowships: 2013–2014, 2014–2015.
- Best Master's Thesis Award, Institute of Information and Computing Machinery, 2009.

PhD Students

Current Ph.D. Students:
Huan Zhang, Minhao Cheng, Xuanqing Liu, Patrick H. Chen, Xiangning Chen, Yuanhao Xiong, Li-Cheng Lan, Bhairav Chidambaram, Qin Ding (co-advised with James Sharpnack), Lifeng Wei (co-advised with James Sharpnack).

Former Ph.D. Students:
Yao Li (Assistant Professor, UNC Chapel Hill)
Puyudi Yang (Facebook)
Liwei Wu (LinedIn)
Franco Liang (LinedIn)

Publications

Google Scholar Profile: Number of Citations=16000+; h-index = 46, i10-index = 86. Details available at `http://scholar.google.com/citations?user=Wy89g4IAAAAJ&hl=en&oi=ao`

Conference Publications

1. Huan Zhang*, Hongge Chen*, Chaowei Xiao, Bo Li, Mingyan Liu, Duane Boning, Cho-Jui Hsieh (* Equal Contribution). Robust Deep Reinforcement Learning against Adversarial Perturbations on State Observations. In *Neural Information Processing Systems (NeurIPS)*, 2020. **Spotlight presentation**.

2. Hongge Chen, Si Si, Yang Li, Ciprian Chelba, Sanjiv Kumar, Duane Boning, Cho-Jui Hsieh. Multi-Stage Influence Function. In *Neural Information Processing Systems (NeurIPS)*, 2020.

3. Lu Wang, Xuanqing Liu, Jinfeng Yi, Yuan Jiang, Cho-Jui Hsieh. Provably Robust Metric Learning. In *Neural Information Processing Systems (NeurIPS)*, 2020.

4. Chong Zhang, Huan Zhang, Cho-Jui Hsieh. An Efficient Adversarial Attack for Tree Ensembles. In *Neural Information Processing Systems (NeurIPS)*, 2020.

5. Kaidi Xu*, Zhouxing Shi*, Huan Zhang*, Yihan Wang, Kai-Wei Chang, Minlie Huang, Bhavya Kailkhura, Xue Lin, Cho-Jui Hsieh (* Equal Contribution). Provable, Scalable and Automatic Perturbation Analysis on General Computational Graphs. In *Neural Information Processing Systems (NeurIPS)*, 2020.

6. Utkarsh Ojha, Krishna Kumar Singh, Cho-Jui Hsieh, Yong Jae Lee. Generative Modeling of Factorized Representations in Class-Imbalanced Data. In *Neural Information Processing Systems (NeurIPS)*, 2020.

7. Jun-Ho Choi, Huan Zhang, Jun-Hyuk Kim, Cho-Jui Hsieh, Jong-Seok Lee. Adversarially Robust Deep Image Super-Resolution using Entropy Regularization. In *Asian Conference on Computer Vision (ACCV)*, 2020.

8. Liwei Wu, Shuqing Li, Cho-Jui Hsieh, James Sharpnack. SSE-PT: Sequential Recommendation Via Personalized Transformer. In *ACM Recommender Systems conference (Recsys)*, 2020. **Best long paper finalist**.

9. Yuanhao Xiong, Cho-Jui Hsieh. Improved Adversarial Training via Learned Optimizer. In *European Conference on Computer Vision (ECCV)*, 2020.

10. Benlin Liu, Yongming Rao, Jiwen Lu, Jie Zhou, Cho-Jui Hsieh. MetaDistiller: Network Self-boosting via Meta-learned Top-down Distillation. In *European Conference on Computer Vision (ECCV)*, 2020.

11. Xuanqing Liu, Hsiang-Fu Yu, Inderjit Dhillon, Cho-Jui Hsieh. Learning to Encode Position for Transformer with Continuous Dynamical Model. In *International Conference on Machine Learning (ICML)*, 2020.

12. Xiangning Chen, Cho-Jui Hsieh. Stabilizing Differentiable Architecture Search via Perturbation-based Regularization. In *International Conference on Machine Learning (ICML)*, 2020.

13. Yihan Wang, Huan Zhang, Hongge Chen, Duane Boning, Cho-Jui Hsieh. On Lp-norm Robustness of Ensemble Decision Stumps and Trees. In *International Conference on Machine Learning (ICML)*, 2020.

14. Xiaoqing Zheng, Jiehang Zeng, Yi Zhou, Cho-Jui Hsieh, Minhao Cheng, Xuanjing Huang. Evaluating and enhancing the robustness of neural network-based dependency parsing models with adversarial examples. In *Association for Computational Linguistics (ACL)*, 2020. (long).

15. Liunian Harold Li, Mark Yatskar, Da Yin, Cho-Jui Hsieh, Kai-Wei Chang. What Does BERT with Vision Look At? In *Association for Computational Linguistics (ACL)*, 2020. (short).

16. Xuanqing Liu, Tesi Xiao, Si Si, Qin Cao, Sanjiv Kumar, Cho-Jui Hsieh. How Does Noise Help Robustness? Explanation and Exploration under the Neural SDE Framework. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. **Oral presentation**.

17. Jyun-Yu Jiang*, Patrick H. Chen*, Cho-Jui Hsieh and Wei Wang. (* Equal Contribution). Clustering and Constructing User Coresets to Accelerate Large-scale Top-K Recommender Systems. In *ACM WWW International conference on World Wide Web (WWW)*, 2020.

18. Quanming Yao*, Xiangning Chen*, James T. Kwok, Yong Li, Cho-Jui Hsieh (* Equal Contribution). Efficient Neural Interaction Functions Search for Collaborative Filtering. In *ACM WWW International conference on World Wide Web (WWW)*, 2020.

19. Liwei Wu, Hsiang-Fu Yu, Nikhil Rao, James Sharpnack, Cho-Jui Hsieh. Graph DNA: Deep Neighborhood Aware Graph Encoding for Collaborative Filtering. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2020.

20. Minhao Cheng*, Simranjit Singh*, Patrick H. Chen, Pin-Yu Chen, Sijia Liu, Cho-Jui Hsieh. Sign-OPT: A Query-Efficient Hard-label Adversarial Attack. In *International Conference on Learning Representations (ICLR)*, 2020.

21. Zhouxing Shi, Huan Zhang, Kai-Wei Chang, Minlie Huang, Cho-Jui Hsieh. Robustness Verification for Transformers. In *International Conference on Learning Representations (ICLR)*, 2020.

22. Yangjun Ruan, Yuanhao Xiong, Sashank Reddi, Sanjiv Kumar, Cho-Jui Hsieh. Learning to Learn by Zeroth-Order Oracle. In *International Conference on Learning Representations (ICLR)*, 2020.

23. Huan Zhang, Hongge Chen, Chaowei Xiao, Sven Gowal, Robert Stanforth, Bo Li, Duane Boning, Cho-Jui Hsieh. Towards Stable and Efficient Training of Verifiably Robust Neural Networks. In *International Conference on Learning Representations (ICLR)*, 2020.

24. 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. In *International Conference on Learning Representations (ICLR)*, 2020.

25. Yang You, Jing Li, Sashank Reddi, Jonathan Hseu, Sanjiv Kumar, Srinadh Bhojanapalli, Xiaodan Song, James Demmel, Cho-Jui Hsieh. Large Batch Optimization for Deep Learning: Training BERT in 76 minutes. In *International Conference on Learning Representations (ICLR)*, 2020.

26. Minhao Cheng, Jinfeng Yi, Huan Zhang, Pin-Yu Chen, Cho-Jui Hsieh. Seq2Sick: Evaluating the Robustness of Sequence-to-Sequence Models with Adversarial Examples. In *the AAAI Conference on Artificial Intelligence (AAAI)*, 2020.

27. Puyudi Yang, Jianbo Chen, Cho-Jui Hsieh, Jane-Ling Wang, Michael I. Jordan. ML-LOO: Detecting Adversarial Examples with Feature Attribution. In *the AAAI Conference on Artificial Intelligence (AAAI)*, 2020.

28. Hongge Chen*, Huan Zhang*, Si Si, Yang Li, Duane Boing, Cho-Jui Hsieh. (* Equal contributio). Robustness Verification of Tree-based Models. In *Neural Information Processing Systems (NeurIPS)*, 2019.

29. Xuanqing Liu, Si Si, Xiaojin Zhu, Yang Li, and Cho-Jui Hsieh. A Unified Framework for Data Poisoning Attack to Graph-based Semi-supervised Learning. In *Neural Information Processing Systems (NeurIPS)*, 2019.

30. Liwei Wu, Shuqing Li, Cho-Jui Hsieh, James Sharpnack. Stochastic Shared Embeddings: Data-driven Regularization of Embedding Layers. In *Neural Information Processing Systems (NeurIPS)*, 2019.

31. Ruiqi Gao, Tianle Cai, Haochuan Li, Liwei Wang, Cho-Jui Hsieh, Jason D. Lee. Convergence of Adversarial Training in Overparameterized Networks. In *Neural Information Processing Systems (NeurIPS)*, 2019. **Spotlight presentation**.

32. Hadi Salman, Greg Yang, Huan Zhang, Cho-Jui Hsieh, Pengchuan Zhang. A Convex Relaxation Barrier to Tight Robustness Verification of Neural Networks. In *Neural Information Processing Systems (NeurIPS)*, 2019.

33. Yukun Ma*, Patrick H. Chen*, Cho-Jui Hsieh (* Equal contributio). MulCode: A Multiplicative Multi-way Model for Compressing Neural Language Model. In *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2019.

34. Jun-Ho Choi, Huan Zhang, Jun-Hyuk Kim, Cho-Jui Hsieh, Jong-Seok Lee. Evaluating Robustness of Deep Image Super-Resolution Against Adversarial Attacks. In *International Conference on Computer Vision (ICCV)*, 2019.

35. Y. You, Y. He, S. Rajbhandari, W. Wang, C.-J. Hsieh, K. Keutzer, J. Demmel. Fast LSTM Inference by Dynamic Decomposition on Cloud Systems. In *IEEE International Conference on Data Mining (ICDM)*, 2019. **Best paper candidate**.

36. Yang You, Jonathan Hseu, Chris Ying, James Demmel, Kurt Keutzer, Cho-Jui Hsieh. Large-batch Training for LSTM and Beyond. In *International Conference for High Performance Computing, Networking, Storage, and Analysis (SC)*, 2019. **Best Student Paper Award Finalist**.

37. Yu-Lun Hsieh, Minhao Cheng, Da-Cheng Juan, Wei Wei, Wen-Lian Hsu, Cho-Jui Hsieh. On the Robustness of Self-Attentive Models. In *Association for Computational Linguistics (ACL)*, 2019.

38. Wei-Lin Chiang, Xuanqing Liu, Si Si, Yang Li, Samy Bengio, Cho-Jui Hsieh. Cluster-GCN: An Efficient Algorithm for Training Deep and Large Graph Convolutional Networks. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2019.

39. Hongge Chen, Huan Zhang, Duane Boning, Cho-Jui Hsieh. Robust Decision Trees Against Adversarial Examples. In *International Conference on Machine Learning (ICML)*, 2019.

40. Moustafa Alzantot, Yash Sharma, Supriyo Chakraborty, Huan Zhang, Cho-Jui Hsieh, Mani Srivastava. GenAttack: Practical Black-box Attacks with Gradient-Free Optimization. In *Genetic and Evolutionary Computation Conference (GECCO)*, 2019.

41. Xuanqing Liu, Cho-Jui Hsieh. Rob-GAN: Generator, Discriminator, and Adversarial Attacker. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

42. Minhao Cheng, Wei Wei, Cho-Jui Hsieh. Evaluating and Enhancing the Robustness of Dialogue Systems: A Case Study on a Negotiation Agent. In *North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, 2019.

43. Patrick H. Chen, Si Si, Sanjiv Kumar, Yang Li, Cho-Jui Hsieh. Learning to Screen for Fast Softmax Inference on Large Vocabulary Neural Networks. In *International Conference on Learning Representations (ICLR)*, 2019.

44. Minhao Cheng, Thong Le, Pin-Yu Chen, Huan Zhang, Jinfeng Yi, Cho-Jui Hsieh. Query-Efficient Hard-label Black-box Attack: An Optimization-based Approach. In *International Conference on Learning Representations (ICLR)*, 2019.

45. Huan Zhang*, Hongge Chen*, Zhao Song, Duane Boning, Inderjit Dhillon, Cho-Jui Hsieh. The Limitations of Adversarial Training and the Blind-Spot Attack. In *International Conference on Learning Representations (ICLR)*, 2019.

46. Xuanqing Liu, Yao Li*, Chongruo Wu*, Cho-Jui Hsieh. Adv-BNN: Improved Adversarial Defense through Robust Bayesian Neural Network. In *International Conference on Learning Representations (ICLR)*, 2019.

47. Huang Fang, Minhao Cheng, Cho-Jui Hsieh, Michael Friedlander. Fast Training for Large-Scale One-versus-All Linear Classifiers using Tree-Structured Initialization. In *SIAM International Conference on Data Mining (SDM)*, 2019.

48. Qin Ding, Hsiang-Fu Yu, Cho-Jui Hsieh. A Fast Sampling Algorithm for Maximum Inner Product Search. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2019.

49. Hsiang-Fu Yu, Cho-Jui Hsieh, Inderjit Dhillon. Parallel Asynchronous Stochastic Coordinate Descent with Auxiliary Variables. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2019.

50. Huan Zhang, Pengchuan Zhang, Cho-Jui Hsieh. RecurJac: An Efficient Recursive Algorithm for Bounding Jacobian Matrix of Neural Networks and Its Applications. In *the AAAI Conference on Artificial Intelligence (AAAI)*, 2019.

51. Chun-Chen Tu, Paishun Ting, Pin-Yu Chen, Sijia Liu, Huan Zhang, Jinfeng Yi, Cho-Jui Hsieh. AutoZOOM: Autoencoder-based Zeroth Order Optimization Method for Attacking Black-box Neural Networks. In *the AAAI Conference on Artificial Intelligence (AAAI)*, 2019.

52. Patrick Chen, Si Si, Yang Li, Ciprian Chelba, Cho-Jui Hsieh. GroupReduce: Block-Wise Low-Rank Approximation for Neural Language Model Shrinking. In *Neural Information Processing Systems (NIPS)*, 2018.

53. Huan Zhang*, Lily Weng*, Pin-Yu Chen, Cho-Jui Hsieh, Luca Daniel. Efficient Neural Network Robustness Certification with General Activation Functions. In *Neural Information Processing Systems (NIPS)*, 2018.

54. Yao Li, Minhao Cheng, Kevin Fujii, Fushing Hsieh, Cho-Jui Hsieh. Learning from Group Comparisons: Exploiting Higher-Order Interactions. In *Neural Information Processing Systems (NIPS)*, 2018.

55. Xuanqing Liu, Minhao Cheng, Huan Zhang, Cho-Jui Hsieh. Towards Robust Neural Networks via Random Self-ensemble. In *European Conference on Computer Vision (ECCV)*, 2018.

56. Yang You, Zhao Zhang, Cho-Jui Hsieh, James Demmel, Kurt Keutzer. ImageNet Training in Minutes. In *International Conference on Parallel Processing (ICPP)*, 2018.

57. Xuanqing Liu, Cho-Jui Hsieh. Fast Variance Reduction Method with Stochastic Batch Size. In *International Conference on Machine Learning (ICML)*, 2018.

58. Minhao Cheng, Ian Davidson, Cho-Jui Hsieh. Extreme Learning to Rank via Low Rank Assumption. In *International Conference on Machine Learning (ICML)*, 2018.

59. Liwei Wu, Cho-Jui Hsieh, James Sharpnack. SQL-Rank: A Listwise Approach to Collaborative Ranking. In *International Conference on Machine Learning (ICML)*, 2018.

60. Tsui-Wei Weng*, Huan Zhang*, Hongge Chen, Zhao Song, Cho-Jui Hsieh, Duane Boning, Inderjit Dhillon, Luca Daniel. Towards Fast Computation of Certified Robustness for ReLU Networks. In *International Conference on Machine Learning (ICML)*, 2018.

61. Yang You, James Demmel, Cho-Jui Hsieh, Richard Vuduc. Accurate, Fast and Scalable Kernel Ridge Regression on Parallel and Distributed Systems. In *International Conference on Supercomputing (ICS)*, 2018.

62. H. Chen*, H. Zhang*, P.-Y. Chen, J. Yi, C.-J. Hsieh. Attacking Visual Language Grounding with Adversarial Examples: A Case Study on Neural Image Captioning. In *Association for Computational Linguistics (ACL)*, 2018.

63. M. Cheng, C.-J. Hsieh. Distributed Primal-Dual Optimization for Non-uniformly Distributed Data. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 2018.

64. C. Jiang, H.-F. Yu, C.-J. Hsieh, K.-W. Chang. Learning Word Embeddings for Low-resource Languages by PU Learning. In *North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, 2018.

65. T. Weng*, H. Zhang*, P.-Y. Chen, J. Yi, D. Su, Y. Gao, C.-J. Hsieh, L. Daniel. Evaluating the Robustness of Neural Networks: An Extreme Value Theory Approach. In *International Conference on Learning Representations (ICLR)*, 2018.

66. J. Wang, C.-J. Hsieh. NLRR++: Scalable Subspace Clustering via Non-Convex Block Coordinate Descent. In *SIAM International Conference on Data Mining (SDM)*, 2018.

67. P.-Y. Chen, Y. Sharma, H. Zhang, J. Yi, C.-J. Hsieh. Elastic-Net Attacks to Deep Neural Networks via Adversarial Examples. In *the AAAI Conference on Artificial Intelligence (AAAI)*, 2018.

68. X. Lian, C. Zhang, H. Zhang, C.-J. Hsieh, J. Liu. Can Decentralized Algorithms Outperform Centralized Algorithms? A Case Study for Decentralized Parallel Stochastic Gradient Descent. In *Neural Information Processing Systems(NIPS)*, 2017. **Oral presentation, 1.3% acceptance rate**.

69. J. Yi, C.-J. Hsieh, K. R. Varshney, L. Zhang, Y. Li. Scalable Demand-Aware Recommendation. In *Neural Information Processing Systems(NIPS)*, 2017.

70. H. Yu, C.-J. Hsieh, Q. Lei, I. S. Dhillon. A Greedy Approach for Budgeted Maximum Inner Product Search. In *Neural Information Processing Systems(NIPS)*, 2017.

71. P.-Y Chen*, H. Zhang*, Y. Sharma, J. Yi, C.-J. Hsieh. ZOO: Zeroth Order Optimization based Black-box Attacks to Deep Neural Networks without Training Substitute Models. In *ACM Conference on Computer and Communications Security (CCS) Workshop on Artificial Intelligence and Security (AISec)*, 2017. **Best Paper Award Finalist**.

72. H. Fang, M. Cheng, C.-J. Hsieh. A Hyperplane-based Algorithm for Semi-supervised Dimension Reduction. In *IEEE International Conference on Data Mining (ICDM)*, 2017. Full paper.

73. L. Wu, C.-J. Hsieh, J. Sharpnack. Large-scale Collaborative Ranking in Near-Linear Time. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD)*, 2017.

74. C.-J. Hsieh, S. Si, I. S. Dhillon. Communication-Efficient Distributed Block Minimization for Nonlinear Kernel Machines. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD)*, 2017.

75. S. Si, H. Zhang, S. Keerthi, D. Mahajan, I. S. Dhillon, C.-J. Hsieh. Gradient Boosted Decision Trees for High Dimensional Sparse Output. In *International Conference on Machine Learning(ICML)*, 2017.

76. H. Fang, Z. Zhen, Y. Shao, C.-J. Hsieh. Improved Bounded Matrix Completion for Large-scale Recommender Systems. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 2017.

77. K.-Y. Chiang, C.-J. Hsieh, I. S. Dhillon. Rank Aggregation and Prediction with Item Features. In *AI and Statistics (AISTATS)*, 2017.

78. H. Zhang, C.-J. Hsieh. Fixing the Convergence Problems in Parallel Asynchronous Dual Coordinate Descent. In *IEEE International Conference on Data Mining (ICDM)*, 2016. Full paper.

79. H. Zhang, C.-J. Hsieh, V. Akella. HogWild++: A New Mechanism for Decentralized Asynchronous Stochastic Gradient Descent. In *IEEE International Conference on Data Mining (ICDM)*, 2016. Full paper.

80. Y. You, X. Lian, J. Liu, H.-F. Yu, I. S. Dhillon, J. Demmel, C.-J. Hsieh. Asynchronous Parallel Greedy Coordinate Descent. In *Neural Information Processing Systems(NIPS)*, 2016.

81. X. Lian, H. Zhang, C.-J. Hsieh, Y. Huang, J. Liu. A Comprehensive Linear Speedup Analysis for Asynchronous Stochastic Parallel Optimization from Zeroth-Order to First-Order. In *Neural Information Processing Systems(NIPS)*, 2016.

82. S. Si, K.-Y. Chiang, C.-J. Hsieh, N. Rao, I. S. Dhillon. Goal-Directed Inductive Matrix Completion. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD)*, 2016.

83. S. Si, C.-J. Hsieh, I. S. Dhillon. Computationally Efficient Nystrom Approximation using Fast Transforms. In *International Conference on Machine Learning(ICML)*, 2016.

84. K.-Y. Chiang, C.-J. Hsieh, I. S. Dhillon. Robust Principal Component Analysis with Side Information. In *International Conference on Machine Learning(ICML)*, 2016.

85. K.-Y. Chiang, C.-J. Hsieh, I. S. Dhillon. Matrix Completion with Noisy Side Information. In *Neural Information Processing Systems(NIPS)*, 2015. **Spotlight presentation**.

86. I. Yen, K. Zhong, C.-J. Hsieh, P. Ravikumar, I. S. Dhillon. Sparse Linear Programming via Primal and Dual Augmented Coordinate Descent. In *Neural Information Processing Systems (NIPS)*, 2015.

87. C.-J. Hsieh, H.-F. Yu, I. S. Dhillon. PASSCoDe: Parallel ASynchronous Stochastic dual Co-ordinate Descent. In *International Conference on Machine Learning(ICML)*, 2015.

88. C.-J. Hsieh, N. Natarajan, I. S. Dhillon. PU Learning for Matrix Completion. In *International Conference on Machine Learning(ICML)*, 2015.

89. H.-F. Yu, C.-J. Hsieh, H. Yun, S. Vishwanathan, I. S. Dhillon. A Scalable Asynchronous Distributed Algorithm for Topic Modeling. In *ACM WWW International conference on World Wide Web(WWW)*, 2015.

90. C.-J. Hsieh, I. S. Dhillon, P. Ravikumar, S. Becker, P. A. Olsen. QUIC & DIRTY: A Quadratic Approximation Approach for Dirty Statistical Models. In *Neural Information Processing Systems(NIPS)*, 2014.

91. C.-J. Hsieh, S. Si, I. S. Dhillon. Fast Prediction for Large-Scale Kernel Machines. In *Neural Information Processing Systems(NIPS)*, 2014.

92. E.-H. Yen, C.-J. Hsieh, P. Ravikumar, I. S. Dhillon. Constant Nullspace Strong Convexity and Fast Convergence of Proximal Methods under High-Dimensional Settings. In *Neural Information Processing Systems (NIPS)*, 2014.

93. C.-J. Hsieh, S. Si, I. S. Dhillon. A Divide-and-Conquer Solver for Kernel Support Vector Machines. In *International Conference on Machine Learning(ICML)*, 2014.

94. S. Si, C.-J. Hsieh, I. S. Dhillon. Memory Efficient Kernel Approximation. In *International Conference on Machine Learning(ICML)*, 2014. **Recommended for JMLR Fast Track, 18 out of 1260+**.

95. C.-J. Hsieh, P. A. Olsen. Nuclear Norm Minimization via Active Subspace Selection. In *International Conference on Machine Learning(ICML)*, 2014.

96. C.-J. Hsieh, M. A. Sustik, I. S. Dhillon, P. Ravikumar, R. A. Poldrack. BIG & QUIC: Sparse Inverse Covariance Estimation for a Million Variables. In *Neural Information Processing Systems(NIPS)*, 2013. **Oral presentation, 1.5% acceptance rate**.

97. H. Wang, A. Banerjee, C.-J. Hsieh, P. Ravikumar, I. S. Dhillon. Large Scale Distributed Sparse Precision Estimation. In *Neural Information Processing Systems (NIPS)*, 2013.

98. C.-J. Hsieh, M. Tiwari, S. Shah, D. Agarwal. Organizational Overlap on Social Networks and its Applications. In *ACM WWW International conference on World Wide Web(WWW)*, 2013.

99. H.-F. Yu, C.-J. Hsieh, S. Si, and I. S. Dhillon. Scalable Coordinate Descent Approaches to Parallel Matrix Factorization for Recommender Systems. In *IEEE International Conference on Data Mining(ICDM)*, 2012. **Best Paper Award**.

100. C.-J. Hsieh, I. S. Dhillon, P. Ravikumar and A. Banerjee. A Divide-and-Conquer Method for Sparse Inverse Covariance Estimation. In *Neural Information Processing Systems(NIPS)*, 2012.

101. C.-J. Hsieh, K.-Y. Chiang and I. S. Dhillon, Low-Rank Modeling of Signed Networks. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD)*, 2012.

102. C.-J. Hsieh, M. A. Sustik, I. S. Dhillon and P. Ravikumar. Sparse Inverse Covariance Matrix Estimation Using Quadratic Approximation. In *Neural Information Processing Systems(NIPS)*, 2011.

103. C.-J. Hsieh and I. S. Dhillon. Fast Coordinate Descent Methods with Variable Selection for Non-negative Matrix Factorization. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD)*, 2011.

104. H.-F. Yu, C.-J. Hsieh, K.-W. Chang, and C.-J. Lin. Large linear classification when data cannot fit in memory. In *International Joint Conference on Artificial Intelligence(IJCAI)*, 2011. The Best Paper Track.

105. H.-F. Yu, C.-J. Hsieh, K.-W. Chang, and C.-J. Lin. Large linear classification when data cannot fit in memory. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD)*, 2010. **Best Research Paper Award**.

106. F.-L. Huang, C.-J. Hsieh, K.-W. Chang, and C.-J. Lin. Iterative scaling and coordinate descent method for maximum entropy models. In *Association for Computational Linguistics(ACL)*, 2009. Short paper.

107. C.-J. Hsieh, K.-W. Chang, C.-J. Lin, S. Sathiya Keerthi, and S. Sundararajan. A Dual Coordinate Descent Method for Large-scale Linear SVM. In *International Conference on Machine Learning(ICML)*, 2008.

108. S. S. Keerthi, S. Sundararajan. K.-W. Chang, C.-J. Hsieh, and C.-J. Lin. A sequential dual method for large scale multi-class linear SVMs. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD)*, 2008.

Journal Publications

1. Puyudi Yang, Jianbo Chen, Cho-Jui Hsieh, Jane-Ling Wang, Michael I. Jordan. Greedy Attack and Gumbel Attack: Generating Adversarial Examples for Discrete Data. *Journal of Machine Learning Research (JMLR)*, 2020.

2. Lu Wang, Huan Zhang, Jinfeng Yi, Cho-Jui Hsieh, Yuan Jiang. Spanning Attack: Reinforce Black-box Attacks with Unlabeled Data. *Machine Learning*, 2020.

3. Liunian Li, Patrick H. Chen, Cho-Jui Hsieh, Kai-Wei Chang. Efficient Contextual Representation Learning Without Softmax Layer. *Transactions of the Association for Computational Linguistics (TACL)*, 2019.

4. Yang You, Zhao Zhang, Cho-Jui Hsieh, James Demmel, Kurt Keutzer. Fast Deep Neural Network Training on Distributed Systems and Cloud TPUs. *IEEE Transactions on Parallel and Distributed Systems*, 2019.

5. Jiarui Fang, Haohuan Fu, Guangwen Yang, Cho-Jui Hsieh. RedSync: Reducing synchronization bandwidth for distributed deep learning training system. In *Journal of Parallel and Distributed Computing*, 2019.

6. K.-Y. Chiang, C.-J. Hsieh and I. S. Dhillon. Using Side Information to Reliably Learn Low-Rank Matrices from Missing and Corrupted Observations. *Journal of Machine Learning Research (JMLR)*, 2018.

7. S. Si, C.-J. Hsieh and I. S. Dhillon. Memory Efficient Kernel Approximation. *Journal of Machine Learning Research (JMLR)*, 2017.

8. F. Hsieh, K. Fujii and C.-J. Hsieh. Machine Learning Meliorates Computing and Robustness in Discrete Combinatorial Optimization Problems. *Frontiers in Applied Mathematics and Statistics*, 2016.

9. H.-F. Yu, C.-J. Hsieh, H. Yun, S. Vishwanathan and I. S. Dhillon. Nomadic Computing for Big Data Analytics. *IEEE Computer*, 2016.

10. C.-J. Hsieh, M. A. Sustik, I. S. Dhillon and P. Ravikumar. QUIC: Quadratic Approximation for Sparse Inverse Covariance Matrix Estimation. *Journal of Machine Learning Research(JMLR)*, 15:2911–2947, 2014.

11. K.-Y. Chiang, C.-J. Hsieh, N. Natarajan, A. Tewari, and I. S. Dhillon. Prediction and Clustering in Signed Networks: A Local to Global Perspective. *Journal of Machine Learning Research(JMLR)*, 15:1177–1213, 2014.

12. H. Yun, H.-F. Yu, C.-J. Hsieh, S. Vishwanathan, I. S. Dhillon. NOMAD: Non-locking, stOchastic Multi-machine algorithm for Asynchronous and Decentralized matrix completion. *Proceedings of the VLDB Endowment*, 7:11:975–986, 2014.

13. H.-F. Yu, C.-J. Hsieh, S. Si, I. S. Dhillon. Parallel Matrix Factorization for Recommender Systems. *Knowledge and Information Systems(KAIS)*, 2013.

14. H.-F. Yu, C.-J. Hsieh, K.-W. Chang, and C.-J. Lin. Large linear classification when data cannot fit in memory. *ACM Transactions on Knowledge Discovery from Data(TKDD)*, 5:23:1–23, 2012.

15. G.-X. Yuan, K.-W. Chang, C.-J. Hsieh, and C.-J. Lin. A comparison of optimization methods for large-scale L1-regularized linear classification. *Journal of Machine Learning Research(JMLR)*, 11:3183–3234, 2010.

16. Y.-W. Chang, C.-J. Hsieh, K.-W. Chang, Michael Ringgaard, and C.-J. Lin. Low-Degree Polynomial Mapping of Data for SVM. *Journal of Machine Learning Research(JMLR)*, 11:1471–1490, 2010.

17. F.-L. Huang, C.-J. Hsieh, K.-W. Chang, and C.-J. Lin. Iterative scaling and coordinate descent method for maximum entropy models. *Journal of Machine Learning Research(JMLR)*, 11:581–614, 2010.

18. R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research(JMLR)*, 9:1871–1874, 2008.

19. K.-W. Chang, C.-J. Hsieh, and C.-J. Lin. Coordinate Descent Method for Large-scale L2-loss Linear SVM. *Journal of Machine Learning Research(JMLR)*, 9:1369–1398, 2008.

Selected Software Packages

- LIBLINEAR – A Library for Large-scale Linear Classification
  1. `http://www.csie.ntu.edu.tw/~cjlin/liblinear`
  2. One of the main contributors.
  3. A comprehensive package containing several efficient linear classification and regression solvers.
- DC-SVM – A Divide-and-Conquer solver for kernel SVM
  1. `http://www.cs.utexas.edu/~cjhsieh/dcsvm`
  2. Solve classification problems with 0.5 million samples in 3 minutes.

- QUIC – QUadratic Inverse Covariance algorithm
  1. http://www.cs.utexas.edu/user/sustik/quic
  2. Proximal Newton method for sparse inverse covariance estimation.
  3. The extension–BigQUIC can solve 1 million dimensional problems (with 1 trillon parameters) in one day using a single machine.
- LIBPMF – A parallel matrix factorization library.
  1. http://www.cs.utexas.edu/~rofuyu/libpmf
  2. Fast and scalable matrix completion solver (on multi-core platforms).
- NMF-CD – Coordinate descent methods for non-negative matrix factorization
  1. http://www.cs.utexas.edu/~cjhsieh/nmf
  2. Coordinate descent algorithms for least squares NMF and KL-NMF.
- AMD – An automatic matrix differentiation library.
  1. https://github.com/pkambadu/AMD
  2. One of the main contributors during my internship in IBM research.
  3. Efficient automatic differentiation computation for matrix functions.
- LIBSVM – A Library for Support Vector Machines
  1. http://www.csie.ntu.edu.tw/~cjlin/libsvm
  2. Assisted Professor Chih-Jen Lin in maintaining the library and answering questions from users.

Talks

1. "On Scalability and Tunability of Neural Network Training". The 4th Workshop on Augmented Intelligent and Interaction (keynote), Nov, 2020.
2. "On Scalability and Tunability of Neural Network Training". Samsung AI Researcher of the Year, Nov, 2020.
3. "On Scalability and Efficiency of Machine Learning Systems". Baidu, Aug, 2020.
4. "Black-box Adversarial Attacks via Zeroth Order Optimization". UCLA workshop on zeroth order optimization, July, 2020.
5. "On Scalability and Efficiency of Machine Learning Systems". Alibaba, July, 2020.
6. "Black-box Adversarial Attacks via Zeroth Order Optimization". UCLA Applied Math, July, 2020.
7. "Adversarial Robustness of Discrete Machine Learning Models". CVPR Workshop on Adversarial Machine Learning (invited), June, 2020.
8. "Stability and Robustness of Machine Learning". Toyota Research, October, 2019.
9. "Evaluating and Enhancing the Robustness of Machine Learning Models". UCLA Applied Math, September, 2019.
10. "Large-scale Ranking and Recommendation". Blizzard, August, 2019.
11. "Neural Network Verification and Defense". UCLA, April, 2019.
12. "Neural Network Verification". Google Research, March, 2019.
13. "Model Compression and Fast Prediction for Large-scale Language Models". Siam CSE, March, 2019.

14. "Robustness of Deep Neural Networks to Adversarial Examples". National Taiwan University, Dec, 2018.

15. "On Extension of CLEVER: A Neural Network Robustness Evaluation Algorithm". GlobalSIP, Nov, 2018.

16. "Robustness of Deep Neural Networks to Adversarial Examples". Google Cloud, Aug, 2018.

17. "Adversarial Examples: Attacks and Defenses". Google Brain, Aug, 2018.

18. "Attacking Black-box Machine Learning Models by Zeroth Order Optimization". Informs. June, 2018.

19. "On Robustness of Deep Neural Networks". Peter Hall Conference. May, 2018.

20. "Improved training of deep neural networks under large batch and adversarial attacks". Google New York. Jan, 2018.

21. "Security for Deep Neural Networks". NTU Machine Learning Symposium, National Taiwan University. Dec, 2017.

22. "An Efficient Trust Region Method for Training Deep Neural Networks". Bay Area Scientific Computing Day, LBNL. Dec, 2017.

23. "Attack Deep Neural Networks". UC Davis. Oct, 2017.

24. "Communication-Efficient Distributed Block Minimization for Nonlinear Kernel Machines". ACM SIGKDD, Halifax. Aug, 2017.

25. "Matrix Completion meets Ranking". ICSA, Chicago. June, 2017.

26. "Learning from Comparisons". UC Davis. June, 2017.

27. "Computational and Statistical Challenges in Matrix Completion". Huawei. March, 2017.

28. "Modied Gradient Boosting Decision Tree for Extreme Classication". NIPS extreme classification workshop. Dec, 2016.

29. "PASSCoDe-Fix: Parallel Semi-Asynchronous Dual Co-ordinate Descent". MOPTA. August, 2016.

30. "Inexact Proximal Newton Methods for Composite Minimization". ICCOPT. July, 2016.

31. "Asynchronous Parallel Optimization in Machine Learning". UC Davis Statistical Sciences Symposium. April, 2016.

32. "Asynchronous Parallel Optimization in Machine Learning". UC Davis Statistical Sciences Symposium. April, 2016.

33. "Parallel Asynchronous Stochastic Co-ordinate Descent with Auxiliary Variables". Informs Optimization Society (IOS). Mar, 2016.

34. "Communication Efficient Parallel Block Minimization for Kernel Machines". UC Berkeley. Feb, 2016.

35. "On Computational and Statistical Challenges in Matrix Completion".

    - University of Science and Technology of China. Dec, 2015.
    - UC Davis. Nov, 2015.

36. "PASSCoDe: Parallel ASynchronous Stochastic dual Co-ordinate Descent".

    - Applied math seminar, UCLA. Nov, 2015.
    - DMML workshop, UC Berkeley. Oct, 2015.
    - International Conference on Machine Learning. July, 2015.

37. "Computational Challenges in Machine Learning". UC Davis. Nov, 2015.

38. "Matrix Completion with Noisy Observations and Features". ML seminar. UC Davis. Oct, 2015.

39. "PU Learning for Matrix Completion". International Conference on Machine Learning, Beijing. July, 2015.

40. "Exploiting Structure in Large-scale Machine Learning Problems".

    - Toyota Technological Institute at Chicago, April, 2015.
    - University of California, Los Angeles, April, 2015.
    - Boston University, April, 2015.
    - Microsoft Research, New York, Mar, 2015.
    - Carnegie Mellon University, Mar, 2015.
    - Cornell University, Mar, 2015.
    - Stony Brook University, Mar, 2015.
    - University of Illinois, Urbana-Champaign, Mar, 2015.
    - University of California, Davis, Mar, 2015.
    - University of California, Santa Barbara, Feb, 2015.
    - Dartmouth College, Feb, 2015.

41. "Automatic Differentiation for Matrix Functions". Machine Learning Symposium, National Taiwan University, Jan, 2015.

42. "Exploiting Structure in Large-scale Optimization for Machine Learning". Guest lecture at UT Austin. Nov, 2014.

43. "Matrix Completion – Theories, Applications, and Scalable Solvers". Appier, Taipei. July, 2014.

44. "Sparse Inverse Covariance Estimation for a Million Variables". ICML workshop on Covariance Selection. Jun, 2014.

45. "Nuclear Norm Minimization via Active Subspace Selection". International Conference on Machine Learning, Beijing. Jun, 2014.

46. "A Divide-and-Conquer Solver for Kernel Support Vector Machines". International Conference on Machine Learning, Beijing. Jun, 2014.

47. "Big & Quic: Sparse Inverse Covariance Estimation for a Million Variables". Machine Learning Symposium, National Taiwan University, Dec, 2013.

48. "Big & Quic: Sparse Inverse Covariance Estimation for a Million Variables". Neural Information Processing Systems, Dec, 2013.

49. "Automatic differentiation for matrix functions and Nuclear Norm Minimization Solvers". IBM Research, Yorktown Heights, NY. Aug, 2013.

50. "Sparse Inverse Covariance Estimation Using Quadratic Approximation". MOPTA, Aug, 2013.

51. "Sparse Inverse Covariance Estimation Using Quadratic Approximation". Machine Learning Symposium, National Taiwan University, Dec, 2012.

52. "Organizational Overlap on Social Networks". LinkedIn, Aug, 2012.

Grants

1. "RI: Small: Learning to Optimize: Designing and Improving Optimizers by ML Algorithms", National Science Foundation, $450,000, 10/01/20–09/30/23, So-PI.
2. "Situational Awareness in Extreme/Dynamic/Contested Environments", ARL, $1,192,089, co-PI (My share: $\sim 20\%$).
3. Facebook Research Award, $80,000, PI (My share: 50%).
4. Intel Research Award, $15,000, So- PI.
5. Gift fund, AITRICS, $20,000, So-PI.
6. "Fast DNN Training on Distributed Systems", Intel, $20,000, So-PI.
7. "RI: SMALL: Fast Prediction and Model Compression for Large-Scale Machine Learning", National Science Foundation, IIS-1719097, $450,000, 08/15/17–07/31/20. So-PI.

Teaching

1. Fall 2020: CS 269 Adversarial Robustness of Machine Learning Models
2. Spring 2020: CS 180 Introduction to Algorithms and Complexity
3. Winter 2019: CS 260 Machine Learning Algorithms
4. Spring 2018: STA 141C Big Data and High Performance Statistical Computing.
5. Winter 2018: ECS 171 Machine Learning.
6. Fall 2017: STA 250 Optimization.
7. Spring 2017: STA 141C Big Data and High Performance Statistical Computing.
8. Fall 2016: ECS 289G Scalable Machine Learning.
9. Winter 2016: STA 250 Optimization.
10. Fall 2015: ECS 289G Scalable Machine Learning.

Professional Activities

1. Organizer:
   - KDD 2020 workshop: Adversarial Learning Methods for Machine Learning and Data Mining.
   - KDD 2019 workshop: Adversarial Learning Methods for Machine Learning and Data Mining.
   - IEEE GlobalSIP symposium on Signal Processing for Adversarial Machine Learning, 2018 (Technical chair).
   - KDD 2018 workshop on Big Data, Streams, and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications (BigMine18).
   - UC Davis 2016 Statistical Sciences Symposium.
   - ICML 2014 workshop on Covariance Estimation and Graphical Model Structure Learning.
2. Senior PC / Area Chair: ICML '19, IJCAI '19, NeurIPS '19, AAAI '20, IJCAI '20, ICML '20, NeurIPS '20, ICLR '21.
3. Paper reviewer & Programming Committee: IEEE TNN, JMLR, IJPRAI, Neural Computation, Neural Computing, TKDD, IEEE TIT, Biometrika, TKDE, DAMI, Mathematical Programming A&B, IEEE WCCI '08, NIPS '11, NIPS '12, NIPS '13, IJCNN '13, ICML '14, AISTATS '14, KDD '14, NIPS '14, TAAI '14, AAAI '15, AISTATS '15, KDD '15, ICML '15, NIPS '15, ACML '15, AAAI '16, AISTATS '16, ICML '16, KDD '16, AISTATS '17, IJCAI '17, ICML '17, KDD '17, NIPS '17, CIKM '17, AISTATS '18, ICLR '18, ICML '18, KDD '18, AAAI '18, NIPS '18, UAI '18, ICDM '18.