Minhao Cheng

CONTACT INFORMATION

Department of Computer Science
UCLA

Tel: (1) 530-601-8331
E-mail: mhcheng@cs.ucla.edu

EDUCATION

University of California Los Angeles, Los Angeles, USA
Ph.D student, Computer Science, September 2018 - Now
Expected graduation: March 2021

University of California Davis, Davis, USA
Ph.D student, Computer Science, September 2015 - August 2018

University of Electronic Science and Technology of China, Chengdu, China
B.S., Computer Science and Technology, July, 2015

RESEARCH INTERESTS

Machine learning Security, Adversarial robustness, Deep Learning, Optimization, Big Data, Recommendation System, Sequence to Sequence Neural Network, Distributed Machine Learning, AutoML.

WORK EXPERIENCE

Microsoft Research, Redmond, USA
Robust Training Method for Better Generalization **June.2020 - Sep.2020**

IBM Research, Yorktown Heights, USA
Scalable Training Method for Adversarial Robustness **June.2019 - Sep.2019**
Advisor: Dr. Pin-Yu Chen

- As a common method to improve adversarial robustness, adversarial training has been widely adopted. However, it suffers with bad model prediction accuracy and bad scalability. In this project, we use a heuristic vicinal function to help model learn better adversarial robustness and generalization

Rakuten Slice, San Mateo, USA
Hierarchical Classification Using Neural Networks **June.2017 - Sep.2017**
Advisor: Dr. Chu-Cheng Hsieh

- The hierarchical classification problem is widely existing in a lot of applications in industry. We develop a Sequence to sequence model to get a better performance over other traditional method and plan to deploy it into product in Slice Co.

PUBLICATION

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 ACL (long), 2020.

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. (* Equal Contribution)

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 AAAI Conference on Artificial Intelligence (AAAI), 2020.

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

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

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

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.

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

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

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

Minhao Cheng, Cho-Jui Hsieh. Distributed Primal-Dual Optimization for Non-uniformly Distributed Data. International Joint Conference on Artificial Intelligence (IJCAI), 2018

Huang Fang, Minhao Cheng, Cho-Jui Hsieh. A Hyperplane-based Algorithm for Semi-supervised Dimension Reduction. IEEE International Conference on Data Mining (ICDM), 2017

Dazhuang Su, Xinzheng Niu, Minhao Cheng. Intelligent Mobile Framework Based on Swarm Computation. CIT/IUCC/DASC/PICom 2015: 1000-1006.

PREPRINT

Xiangning Chen*, Ruochen Wang*, Minhao Cheng*, Xiaocheng Tang, Cho-Jui Hsieh. DrNAS: Dirichlet Neural Architecture Search.

Minhao Cheng, Qi Lei, Pin-Yu Chen, Inderjit Dhillon, Cho-Jui Hsieh. CAT: Customized Adversarial Training for Improved Robustness.

Minhao Cheng, Pin-Yu Chen, Sijia Liu, Shiyu Chang, Cho-Jui Hsieh, Payel Das. SPROUT: Self-Progressing Robust Training.

Huan Zhang, Minhao Cheng, Cho-Jui Hsieh. Enhancing Certifiable Robustness via a Deep Model Ensemble.

Liu Liu, Minhao Cheng, Cho-Jui Hsieh, Dacheng Tao. Stochastic Zeroth-order Optimization via

Variance Reduction method.

Y.T. Chan, Hongyang Chen, Minhao Cheng. Optimal Estimation and Minimum Errors in Range-Free Localization. IEEE Transactions on Aerospace and Electronic Systems.

Patents

Minhao Cheng, Pin-Yu Chen, Sijia Liu, Shiyu Chang, Payel Das. Method and System of Training Robust Machine Learning Models.

Minhao Cheng, Xiaocheng Tang, Chu-Cheng Hsieh. Hierarchical Classification Using Neural Networks.

Research Experience

UCLA, Los Angeles, USA

Adversarial attack on interactive dialog system **July.2018 - Dec.2018**

Advisor: Prof. Cho-Jui Hsieh

- Although we have shown that the sequence-to-sequence model is not robust against adversarial attack, it is still a open question about the robustness of dialog system implemented by deep neural networks. We develop two algorithms which could attack the dialog system successfully with or without knowing the deep neural network model structure.

University of California, Davis, Davis, USA

Query Efficient Hard-label Black-box Attack **March.2018 - May.2018**

Advisor: Prof. Cho-Jui Hsieh

- It has been proved that DNNs models are vulnerable to a very small human-imperceptible perturbation. However, it is a still a challenge when we could only get hard-label instead of probability output. We develop a query efficient algorithm which could apply to industrial-strength image classifiers.

Adversarial Example for Sequence to Sequence Model **Sep.2017 - Feb.2018**

Advisor: Prof. Cho-Jui Hsieh

- Recent research on DNNs has indicated ever-increasing concern on the robustness to adversarial examples. We are designing algorithms to generate adversarial example for sequence to sequence model which is widely used in machine translation, text summarization.

Low-rank Approximation of rankSVM in Recommendation System **Sep.2016 - Mar. 2017**

Advisor: Prof. Cho-Jui Hsieh

- New low-rank approximation method using for the recommendation system. Previously, it uses the ranksvm and its variations to train the data. Now we use the low-rank method to get better speed and accuracy dealing with the pair-wised data in ranksvm.

Line Search Method for Distributed Primal-Dual Optimization **Nov.2015 - May.2016**

Advisor: Prof. Cho-Jui Hsieh

- New line search method in stochastic algorithms in large-scale distributed machine learning scenario to overcome the large Primal-Dual gap in the training periods so that we can achieve a faster convergence and a better accuracy in the distributed machine learning case.

University of Electronic Science and Technology of China, Chengdu, CHINA
Avoiding collision in Intelligent Transportation System **Apr.2014 - Jan.2015**
Advisor: Prof. Yu Xiang

- *Signal process method:* Modeling and Optimizing the detection of the collision between vehicles. We explore using signal detection methods can reduce the high latency and transmission false rate and get a better performance in avoiding collision.
- *Data Mining method:* Using data mining method to eliminate the complexity of situations. Design and trained models for avoiding collision.

Accurate Localization Techniques in Cyber Physical System and Intelligence Transportation System **Nov.2013 - Apr.2014**
Advisor: Prof. Yu Xiang

- Traditional transportation system can not satisfied the growing need for a more intelligent traffic control and Cyber Physical System brought an efficient method to solve this problem. However, the accuracy of localization in the urban area is far from satisfaction. We explore that using some messages returned from satellite helps locating the traffic.

Skills

- Programming Languages: C, Python, Matlab, Julia
- Applications: Libsvm, LaTeX, MPI, Pytorch, Tensorflow

Teaching Experience

University of California, Davis, Davis, USA

- Teaching Assistant in ECS 122B Algorithms. **Mar.2017 - Jun.2017**
- Teaching Assistant in ECS 122A Algorithms. **Apr.2016 - June.2016/ Sep.2016 - Dec.2016**
- Teaching Assistant in STA 250 Optimization. **Jan.2016 - Mar.2016**