Jiefeng Chen

Curriculum Vitae

1(608)9600494
jchen662@wisc.edu

Education

Sept. 2017 – Present
Ph.D. in Computer Science, *University of Wisconsin-Madison*.
GPA: 4.0/4.0 Advisors: Yingyu Liang & Somesh Jha

Sept. 2017 – Present
M.S. in Computer Science, *University of Wisconsin-Madison*.
GPA: 4.0/4.0

Sept. 2013 – Jul. 2017
B.S. in Computer Science & Technology, *Shanghai Jiao Tong University*.
Major GPA: 91.24/100 Overall GPA: 90.47/100 Rank: 7/137

Research Mission

Machine learning has brought unprecedented changes to human society. Yet, it is currently experiencing a fundamental problem of its trustworthiness. Examples abound: How can we produce models that are robust to imperceptible perturbations? How can we train models that produce robust interpretation of their behavior? How can we prevent private user data from leaking during the large-scale, possibly distributed, training environment? How to tackle the currently inherent fairness issues? I am interested in all these aspects and want to seek for solutions to make machine learning more trustworthy.

Publications

Conference Papers

NeurIPS 2019
Robust Attribution Regularization.
Jiefeng Chen, Xi Wu, Vaibhav Rastogi, Yingyu Liang, Somesh Jha
This paper is about proposing a training framework to achieve robust IG attributions and showing its connections with previous objectives designed for robust predictions.

EuroS&P 2019
Towards Understanding Limitations of Pixel Discretization Against Adversarial Attacks.
Jiefeng Chen, Xi Wu, Vaibhav Rastogi, Yingyu Liang, Somesh Jha
This paper is about studying when pixel discretization defenses could work and when they could not work.

ICML 2018
Reinforcing Adversarial Robustness using Model Confidence Induced by Adversarial Training.
Xi Wu, Uyeong Jang, Jiefeng Chen, Lingjiao Chen, Somesh Jha
This paper is about leveraging confidence information induced by adversarial training to reinforce adversarial robustness of a given adversarially trained model.

Journal Papers

The UMAP Journal 2017
The Effects of Self-Driving Vehicles on Traffic Capacity.
Yu Shi, Jiefeng Chen, Qi Li
The outcome of Mathematical Contest in Modeling.

Manuscripts

arXiv 2019 **Concise Explanations of Neural Networks using Adversarial Training**.
Prasad Chalasani, Jiefeng Chen, Somesh Jha, Xi Wu
This paper is about showing adversarial training could induce sparse IG attributions both theoretically and empirically.

Talks

EuroS&P 2019 **Towards Understanding Limitations of Pixel Discretization Against Adversarial Attacks** *Stockholm, Sweden*

Research/Work Experiences

Sept.2017 - Present **University of Wisconsin-Madison** Advised by Yingyu Liang & Somesh Jha
Research Assistant *Madison*
I performed research on Trustworthy Machine Learning and published several papers on Top-tier Machine Learning and Security conferences.

May.2019 - Aug.2019 **Facebook Inc** Search Team
Software Engineer Intern *Bellevue*
My intern project was to build an Statistical Machine Translation (SMT) System for Query Explansion. I mainly implemented Translation Model and Decoder of SMT system. I wrote pipelines to collect data from search logs to train the translation model. I implemented IBM model training via SQL queries to fully parallelize it and allow large corpus training. I addressed several issues in alignment model to improve the model quality. I integrated the SMT system implemented into existing Query Expansion system and got end-to-end high quality results.

Jun.2018 - Aug.2018 **Facebook Inc** Ads Ranking Infra Team
Software Engineer Intern *Menlo Park*
My intern project was to add co-occurrence supervisions on Transductive User Model (TUM) and SparseNN. I implemented cross feature co-occurrence supervision on TUM and got significant improvement, about 0.41% train NE gain and 0.36% eval NE gain. I also implemented word2vec style co-occurrence training on SparseNN to get meaningful embeddings of Ads. I created a dataset to evaluate them and got about 11% AUC improvement.

Honors & Awards

2019 **NeurIPS Travel Award**.
2018 **ICML Travel Award**.
2017 **Outstanding Graduate of Shanghai Jiao Tong University**.
2017 **Outstanding Winner, Mathematical Contest in Modeling**.
2016,2015,2014 **Academic Excellence Scholarship of SJTU (Top 5%)**.
2014 **First Prize in National Undergraduate Mathematical Competition (Shanghai)**.
2012, 2011 **The First Prize of Chinese Mathematical Olympiad in Senior**.

Skills

Programming Languages **Proficent in C/C++, Python, Matlab, Bash; Familiar with Java**.

Platform & Tools **Tensorflow, PyTorch, Linux, Git, LaTeX, Caffe, Caffe2, MySQL, CUDA**.